

**SIMILAR ECONOMIES, SIMILAR HISTORIES,  
DIFFERENT STRUCTURES**  
**Transatlantic Contrasts in the Evolution of Professional Sports Leagues**

22 March 2005

Louis P. Cain and David D. Haddock \*

---

\* Cain <[lcain@luc.edu](mailto:lcain@luc.edu)> is Professor of Economics at Loyola University Chicago and Adjunct Professor of Economics at Northwestern University, while Haddock <[dhaddock@northwestern.edu](mailto:dhaddock@northwestern.edu)> is Professor of Law and of Economics at Northwestern University, and Senior Associate at PERC. We benefited from the contributions of Ran Abramitzky, Fabio Braggion, George Harmon, Charles Hickson, Brooks Kaiser, Lynne Kiesling, Robert McGuire, Chiaki Moriguchi, Roger Noll, Kirk Ross, Ian Savage, David Surdam, Peter Wardley, Samuel Williamson, Todd Zywicki, the DePaul University Economics Department Seminar, the Northwestern University Economic History Workshop, the 2004 World Cliometrics Conference in Venice, and the 2004 Southern Economics Association meetings.

## *Abstract*

*When the same industry has a substantially different structure in Europe than in North America, the differences can teach lessons about industrial evolution. For instance, the organization of professional sports allows much easier entry of teams into European than North American leagues; most European leagues utilize near-automatic promotion & relegation. In stark contrast, new teams are rarely admitted to the top North American leagues, and the exceptions are new franchises sold by the league, never a club promoted from below. This paper contrasts the histories of the English Football League and the National Baseball League of the United States—the first modern leagues on their respective continents—to explain the transatlantic difference as a consequence of compactness, the entertainment level of games, and territorial monopolies. The great distances between most U.S. city pairs were incompatible with switching between playing a game and an alternative job. Thus, well before a league formed, top-flight baseball teams employed fully professional players. The less compact U.S. also made maintenance of league parity difficult, but the innovation of rules that barred most of the large number of independent teams from competing with the National League improved parity. With higher costs, the National League buffered competition among its members to increase revenues. If the evolution of the distinctive leagues is understandable, the persistence of the North American variant into an age of increased population densities, increased real wealth, and diminished travel costs is problematic.*

Sports and other popular entertainments blossomed as the workweek shrank during the 1860s. During their early amateur and semi-professional phases, the development of sport was similar on both sides of the Atlantic. When the National (baseball) League in the U.S. and the English Football (soccer) League emerged, however, the structures diverged markedly. This is a puzzle. Both leagues were aware of developments on the other shore, so ignorance cannot account for the difference. This paper will attempt to explain how, despite similar economic histories, subtle differences in the U.S. and English economies contributed to substantial differences in the composition of costs and benefits—and in league structures.

Discussions of the economics of sports, particularly issues relating to the openness of entry or competitive balance, often contrast the promotion & relegation system common in

European leagues with the closed-entry North American structure.<sup>1</sup> Few commentators ask why the difference exists, and those few typically fall back on some *ad hoc* cultural explanation. For example, North American team owners are said to be “profit-maximizing businessmen” while European team owners are “sportsmen.” Yet even if true, sportsmen must remain solvent. Whether the goal is to maximize profits or to win championships within financial constraints, costs and benefits are important.

The tradeoffs that leagues faced were different in Europe and America, due mainly to their economic geographies. In the short run, if league members are geographically dispersed, intraleague competition is reduced, enabling teams to charge higher prices. But dispersal increases travel costs, most notably the opportunity cost of the time players had to spend away from full-time jobs during the early years. Dispersal also makes it more likely that an entrant can find a geographic niche between existing teams. Beginning by describing the league structures, we elaborate on the puzzle that divergence poses before turning to the histories of the pioneer leagues for an explanation.

### **What is the Difference in Structure?**

The peculiarities are not of a sport but of a shore.<sup>2</sup> The National League, founded in 1876, was the first permanent professional sports league in the world; the English Football League, founded

---

<sup>1</sup> In the U.S. context, this is usually associated in sports economics textbooks with a discussion of big- versus small-market teams; for example, Leeds and Allmen, *Economics of Sports*. There is a rapidly emerging literature that analyzes many aspects of the promotion & relegation system (e.g., Noll, “Economics of Promotion;” Ross and Szymanski, “Law & Economics;” and Szymanski, “Economic Design.”

<sup>2</sup> Important intercontinental distinctions now characterize even basketball, ice hockey, and soccer, though cricket seems to operate more along North American lines.

twelve years later, was the first in Europe. Each provided paradigms for other leagues on their respective continents, a distinctiveness facilitated in the United States by an antitrust exemption.

*North American Leagues*—Despite many willing entrants, top-level North American professional leagues maintain a fixed roster of teams—a “closed corporation,” as the promoters of the National League referred to their structure.<sup>3</sup> Only a fixed number of teams compete until the incumbents decide to admit others. Each team enjoys a league-enforced territorial monopoly (or duopoly).<sup>4</sup> With three exceptions today, teams in the same sport in the same metropolitan area play in different stadiums.<sup>5</sup>

Starting a new top-level team is expensive in North America. Existing teams, acting through the league, create and sell new franchises for prices vastly in excess of the value of any tangible property an entrant receives. The players, grounds, and equipment must then be assembled at additional cost. A minor league team in the expansion team’s location characteristically moves to another city or ceases operation. A rejected entrepreneur’s only recourse is to organize an entire new league. Attempts have been relatively common, but few have succeeded.<sup>6</sup>

---

<sup>3</sup> Rader, *Baseball*, p. 49.

<sup>4</sup> Only the New York City metropolitan area has had three top-league teams in a sport at the same time—three Major League Baseball teams from 1903 to 1957 and three National Hockey League teams since 1982.

<sup>5</sup> The stadiums are separated by eight to thirty miles (13 to 50 km.). . By Stigler’s survivorship standard (*Organization of Industry*, 72-74), the fact that facility sharing was once more common, but has receded to insignificance, is instructive.

<sup>6</sup> Given that entry is attempted despite a high failure rate, profitability must be perceived to be well above the norm. The most recent insurgency that ultimately succeeded in injecting all franchises into a “major” sport’s top rung was the American Football League, which started operation in 1960. No successful entry without the concurrence of the incumbents has occurred in baseball since 1901. Two recent attempts to enter against the National Basketball Association and the National Hockey League resulted in a minority of the insurgent teams ultimately being

Existing teams have relocated with league consent to new cities as far as 2800 miles away (4500 km.). An abandoned city has difficulty attracting a replacement.<sup>7</sup> Although long-distance moves generate acrimony, relocations of several dozen miles are uncontroversial—neither the Angels’ move from downtown Los Angeles to a suburb nor the Patriots’ move from downtown Boston halfway to Rhode Island generated much comment.

League rules specify roster size. They forbid loaning players to other teams at the same level, although baseball (and ice hockey) teams hire several times as many players as their active rosters can hold, assigning the surplus to “farm” teams that play elsewhere. Though teams can trade player contracts, if a cash-strapped owner attempts to capitalize the roster by selling the contracts of the team’s stars to competitors, the league will nullify the deals.

***European Leagues***—Most European leagues use nearly automatic promotion & relegation. Strong teams at one level interchange annually with the weakest teams from the next higher level. The top Football League division (known since 1992 as the Premier League) has only 20 teams, but according to League rules any of the 92 teams in four divisions could be playing there within three years.<sup>8</sup> During roughly half the years of League history, one or two top non-league teams have supplanted the bottom league teams.<sup>9</sup> Thus European leagues are open to entry

---

merged into the incumbent leagues. More commonly, all the franchises from insurgent leagues are left by the wayside.

<sup>7</sup> Washington has now regained a major league baseball team after being without one for a third of a century; the nation’s second largest market, Los Angeles, has had no professional football team for a decade.

<sup>8</sup> Throughout the article we employ the division names the Football League adopted in 1992 when all were renamed, the old First Division becoming the new Premier League, the old Second Division becoming Division One, and so on.

<sup>9</sup> Roughly 800 English and Welsh teams registered with UEFA, soccer’s European governing body, have web pages. They range from the Premier League’s reigning champion on down

through superior performance.<sup>10</sup> Few European leagues recognize territorial rights, and their member teams often play in close proximity, sometimes within walking distance of each other.<sup>11</sup> Measured as the crow flies, the stadiums of three London teams—Chelsea, Fulham, and Queens Park—are separated by less than three miles (5 km.). By North American standards, Football League teams are packed phenomenally tightly. A dozen football teams from inside London’s ring road have competed in the Premier League, and others play in distant suburbs. Smaller cities have multiple teams that have reached the Premier League—both the Manchester and Birmingham areas have 5 such teams within 20 miles (32 km.) of their centers, and each has 3 teams playing at that level this season.

A new team may locate near an existing one without permission; teams seldom relocate.<sup>12</sup> Teams rarely trade player contracts, but they often sell them for cash. If there is no vacancy on the roster, a player may be loaned to another team—possibly competing against the team that

---

(Open Directory Project). Though few of those teams have the potential for entering the Football League—now or ever—they all are linked in a hierarchy capable of promoting a team from the very bottom to the Premier League, while relegating the defending champion to the bottom. Thus, any entrepreneur with the ability to assemble a top-flight team has many platforms from which to mount the invasion.

<sup>10</sup> To avoid disappointing competitors by being unable to accommodate visiting fans or to complete the season’s schedule, leagues require that a team have an adequate stadium, which varies by division, and be financially sound. Rather than being unreasonable barriers to entry, the solvency and stadium requirements are immune to interference by competitors.

<sup>11</sup> Among the twenty teams within London’s ring road that have reached the top non-league level, the greatest straight-line separation—Crystal Palace to Watford—is about 25 miles (40 km.). Liverpool plays 0.6 miles (1 kilometer) from Everton, 5 miles (8 km.) from Tranmere, and 16 miles (26 km.) from Wigan. Manchester United also plays 16 miles from Wigan, as well as 3 miles (4 km.) from Manchester City, 7 miles (11 km.) from Stockport County, and 9 miles (14 km.) from Bolton, from Bury, and from Oldham.

<sup>12</sup> Wimbledon’s unprecedented 50-mile (80-km.) move from a deteriorated stadium to a new one in Milton Keynes sparked bitter controversy. A semi-professional team, also named Wimbledon, formed in 2002 and plays less than 8 miles (13 km.) from 4 league teams, [http://en.wikipedia.org/wiki/AFC\\_Wimbledon](http://en.wikipedia.org/wiki/AFC_Wimbledon).

owns his contract—or compete on a reserve team whose home games are played in the same stadium. No farm teams play elsewhere.

### **Why is this a Puzzle?**

When leagues form, it is advantageous for existing teams to adopt the closed cartel now common in North America, however aspiring entrants and many paying customers are advantaged by the promotion & relegation system, the European open cartel. With the European structure, (a) teams must compete for ticket sales and (recently) licensing of local broadcast rights; (b) incumbents receive no fees from entrants, who may locate literally next door; and (c) there is little competition among political units for teams since European leagues do not deny promotion to teams playing at any site, while North American teams often lease stadiums owned by their home city for far less than the amortization cost.<sup>13</sup> The only advantage for existing members of an open cartel is season-end interest among fans of teams attempting to avoid relegation. Why then did the Football League decide against a closed cartel? What can we learn from their comparative histories?

*North America*—Before the Civil War, something resembling modern baseball evolved in New York City from an informal children’s game called rounders. New Yorkers taught their fellow soldiers to play during the War, and the game spread quickly when troops returned home.<sup>14</sup> Young men soon began to form clubs whose main function was to rent fields where members could play intramurally. Eventually club "all-star" nines began to travel to play other all-stars,

---

<sup>13</sup> Okner, “Subsidies;” Baade and Sanderson, “Field of Fantasies;” Knautz, “Franchise Free Agency;” and Poitras and Hadley, “Do New Major League Ballparks...”

<sup>14</sup> Rader, *Baseball*, pp 17-18.

but there was no consolidated schedule or championship. In 1862 William Cammeyer made his Brooklyn property available free to local teams if they agreed to let him charge spectators for watching the games.<sup>15</sup> Cammeyer's innovation spread rapidly. When a team's receipts began to exceed expenses, the surplus was used to entice "crack" players away from other teams.<sup>16</sup>

The National Association of Professional Base Ball Players broke away from an earlier and staunchly amateur National Association of Base Ball Players in 1871.<sup>17</sup> Membership was unrestricted; any team could compete for the payment of an annual fee. As would be true for nearly a century, the top professional teams operated only in the Northeast and Old Northwest, east of the Mississippi Valley. Scheduling protocols meant that if a "western" team was supposed to play a series in the east, but could find no intervening opponents with open dates to help defray travel costs, the trip might be cancelled. Having made the trip, a team might even discover that its opponent had disbanded. The best-of-five format meant that a three-game sweep would precede two days of unwanted idleness, two days without paying fans. In the second year of the professional group's operation, two teams played 54 games while one played only nine. In addition to the variability of games, the number, identity, and skill levels of professional teams varied between seasons.<sup>18</sup>

---

<sup>15</sup> Kirsch, *The Creation*, p. 235).

<sup>16</sup> Within a year of Cammeyer's innovation the *Brooklyn Eagle* reported, "ballplaying has become quite a moneymaking business, many finding it to pay well to play well" (Rader, *Baseball*, p. 27).

<sup>17</sup> The National Association of Professional Base Ball Players was a proto-league, because its self-selected member clubs were merely instructed to arrange, on their own, best-of-five series against each of the others. Rader, *Baseball*, p. 45.

<sup>18</sup> Mergers and name changes create ambiguity, but at least 23 different teams competed during the professional association's five seasons, though never more than thirteen during any one. One team lost all eleven of its contests one season, while another lost all of its six the next. During those same seasons, other teams played 54 and 59 games.



The National Association of Professional Base Ball Players failed miserably to develop an orderly process for teams to contest a championship and balance their books, so the *Chicago Tribune* urged a restructuring. The result was "a new baseball association based on territorial monopolies and a restricted number of strong, financially healthy clubs."<sup>19</sup> The first modern league, the National League, excluded weak teams that had imposed external costs on their fellow Association members. In addition to centralized scheduling, the new league hoped to facilitate extended, multi-game trips that economized on travel costs by having four western teams (Chicago, Cincinnati, Louisville, St. Louis) and four eastern (Boston, Hartford, New York, Philadelphia). One club represented each city. Two opponents could block a new entrant; two supporters could block expulsion of an existing member, making involuntary relegation unlikely, barring serious malfeasance.<sup>20</sup> Eventually, in the 1920s, the US Supreme Court legitimized baseball's closed cartel by granting an antitrust exemption.

There were three early attempts to circumvent National League exclusion. The American Association was established in 1882, mainly in cities recently abandoned by the National League. Over the next ten years, the two leagues competed head-to-head in eight cities. The American Association had lost three of those confrontations and the National League but one, when, in 1892, the National League persuaded the four strongest American Association members to join a twelve-team National League, effectively killing the rival league. By 1900 the National League had bought out four of its member teams and was back to eight. The second attempt was

---

<sup>19</sup> Goldstein, *Playing for Keeps*, p. 147.

<sup>20</sup> E. Woodrow Eckard, "Team Promotion," describes the flux during the years before the National League-American Association "merger" as a "quasi team-promotion system."

the Union Association of 1884. The National League formed reserve teams of younger players in a successful effort to starve this league.<sup>21</sup>

The only effective challenge came from the minor Western League. Renaming itself the American League in 1901, this upstart moved its small-market teams into three cities the National League had just abandoned and went head-to-head with the National League in three others. Over the next two years the American League transferred two more franchises and was in direct competition with six National League clubs. Perceiving a credible threat, the National League grudgingly accepted the American League as co-equal via the Major League Agreement of 1903.<sup>22</sup> The major leagues now had not merely the teams but the precise locations that would persist for half a century.

*Europe*—The evolution of English football was similar. Informal games resembling modern football trace back millennia, but interscholastic competition between English schools led to the first codification of rules, the Cambridge Rules of 1848.<sup>23</sup> Graduating students formed clubs in their home cities. Working-class men imitated the club structure as their leisure hours grew. Eventually all-star elevens began to compete. As in early baseball, there was no central schedule, no championship.

In 1863 the Football Association was formed to refine the code and resolve disputes, incidentally severing soccer from rugby. Eight years later, the year that baseball's National

---

<sup>21</sup> Subsequently the reserve teams have comprised the minor leagues.

<sup>22</sup> It is beyond our scope to account for the American League's success after the American Association's failure, given the overlap of cities. Those in baseball were impressed with the entrepreneurial genius of Western League/American League commissioner Ban Johnson. A number of other leagues have attempted entry since 1903, but none of those clubs survived.

<sup>23</sup> Young, *British Football*, pp.74-76, and Harris, *Sport in Britain*, p. 107.

Association of Professional Base Ball Players was formed, the Football Association established the FA Cup, an open-entry knockout tournament that still exists.<sup>24</sup> The following year, the Football Association initiated international matches, initially among England, Ireland, Scotland, and Wales.<sup>25</sup> Well before the Football League emerged, the FA Cup and the internationals had grown popular, which influenced the divergent evolution of the leagues—whenever the Football League has placed too great a distance between itself and non-league members, the Football Association has threatened to exclude League teams from the FA Cup competition and their players from the internationals. Were they to do so, league clubs would have a difficult time hiring players.

Oxbridge graduates, over-represented both in the Football Association administration and the football-playing population in and around London, refused to play against anyone who was paid to play.<sup>26</sup> Industrial worker-players, concentrated in the Midlands and the north of England, could afford trips to distant games only if compensated for “broken time” on their jobs.<sup>27</sup> In 1885, three years before the Football League formed, the Football Association voted to allow payment for wages lost while playing and traveling, but it refused to condone professionalism. Consequently, the Football League initially was semi-professional. As we use the term, a semi-professional is someone who, even during the playing season, receives a regular

---

<sup>24</sup> The Football Association’s cup tournament was copied throughout Europe (remaining separate from, though simultaneous with, national league seasons).

<sup>25</sup> Even today, the principal components of the United Kingdom play as independent entities in international competitions such as the World Cup.

<sup>26</sup> Several working-class football clubs from industrial regions were temporarily suspended or otherwise punished by the FA for paying a few players (Vamplew, *Pay Up*, pp. 191-92).

<sup>27</sup> Well-heeled amateur Old Etonians traveled north to take on working-class Darwen, but Darwen’s players felt they could not forego the wages that would be lost during a return match in London even though the Old Etonians offered to pay their transportation. Harris, *Sport in Britain*, pp. 111-12.

paycheck from a job that is unrelated to the game.<sup>28</sup> Since some locales could generate greater revenues to pay stars, even limited semi-professionalism increased the variance of ability across teams, leading to one-sided contests unappealing to fans. Teams began to seek parity (opponents of a similar caliber).

The Football League formed in 1888 when Birmingham's Aston Villa club proposed that a few top semi-professional teams annually play one another home and away, and eleven others agreed.

With [semi-] professionalism "good gates had become imperative" to the leading clubs, but the multiplicity of local cup competitions, often with one-sided early rounds, could not guarantee this and they also interfered with more remunerative friendly fixtures against other major teams.<sup>29</sup>

Because the best London clubs were steadfastly amateur, the Football League made no pretense of being national; it kept games near its players' main jobs. Six Lancashire teams, three in-and-around Birmingham, and three in-between comprised the initial league.<sup>30</sup> Two years later the league expanded north and east to industrial Durham, Yorkshire, and Lincolnshire, not south toward London.

---

<sup>28</sup> A semi-professional plays during off-hours from that job and receives less from his sport than would have to be paid to a professional player. A person paid to play for an industrial team used for promotional purposes by a non-sports firm is a professional, unless the player is expected to participate meaningfully in that firm's normal business during the season.

<sup>29</sup> Vamplew, *Pay Up*, p. 125.

<sup>30</sup> After five seasons Accrington withdrew from the Football League rather than accept relegation and folded two seasons later (<http://accringtonstanley.has.it/>). All eleven of the other original teams still play. Aston Villa (Birmingham), Blackburn, Bolton, Everton (Liverpool), and West Bromwich (Birmingham) are in the Premier League at present. Burnley, Derby County, Preston, Stoke City, and Wolverhampton are in Division One (or Championship as of 2004), with Notts County in Division Three (the new League Two).

From the start the league was willing to vary its membership, requiring reapplications by the four clubs with the poorest won-lost records. The stated motivation was to assure that league teams were and remained the best available.<sup>31</sup> Adding a lower division and annually interchanging a few teams between the two divisions following test matches made maintaining league quality less contentious and more systematic, the election system remaining only for the bottom teams of the lower division. That enabled the Football League to absorb a rival league one year its junior—the Football Alliance—initially at an inferior level.<sup>32</sup> In 1898 annual promotion & relegation replaced test matches, as the tests produced anomalous results.<sup>33</sup>

In 1893, Arsenal became the Football League’s first London-based team, but the league didn’t move south with alacrity until after World War I, when the upper division of the Southern League was absorbed as a third division.<sup>34</sup> Twenty small northern teams were added the following year, but this did not result in a four-division Football League until 1958. In the

---

<sup>31</sup> Rail connections weighed heavily, “What has been the first point urged in any application for membership ... ? The ease with which visiting clubs could get to the new centre and away from it.” Journalist James Cotton, as quoted in Taylor and Coyle, “The Election.” p. 9.

<sup>32</sup> Arnold, “Rich man, poor man,” p. 52. The Football Alliance brought prominent teams to the Football League with seven eventually reaching the Premier League—Birmingham City, Grimsby, Manchester City, Manchester United, Nottingham Forest, Sheffield Wednesday, and Sheffield United.

<sup>33</sup> Following the first season, the upper division’s worst team retained its rung while two teams that had finished well above it during the full season did not; the lower division’s best team failed to win promotion while the second and third best teams did

<sup>34</sup> Arnold, “Rich man, poor man,” p. 52). Five additional London teams followed Arsenal into the Football League piecemeal as the league gradually expanded between 1893 and World War I. The Southern League merger in 1920 brought six more London teams. Bristol City and Luton had been the only non-London southern teams in the pre-war Football League, but the merger added fourteen more, and for the first time there were teams in Wales.

interim, the third level was divided into northern and southern units, with no inter-division matches.<sup>35</sup> By 1921, the number of league teams had almost reached its present level.<sup>36</sup>

***The Inadequacy of a Cultural Explanation***—The difference in league structures is sometimes alleged to result from “sportsmen” running European teams, while “profit-maximizers” run those in North America.<sup>37</sup> Methodologically, the allegation relies on an unobservable phenomenon—the observer’s supposition regarding entrepreneurial motivation. Moreover, it is unnecessary. As will be shown, a consideration of costs and benefits generates a satisfactory account of intercontinental differences between leagues.

Empirically, history reveals that the first teams on both shores were informal recreation among amateurs—teams run for and by the players. When spectators betrayed a willingness to pay to watch, a few teams began to offer a share of the receipts to entice some players to join. This predated league formation on both continents. When the strongest of those teams began to absorb more player time during the season and pay them commensurately, they had to train a more critical eye on the bottom line. Professional managers, as opposed to “sportsmen,” grew paramount as teams on both shores were forced to become profit-oriented.<sup>38</sup> Baseball teams in

---

<sup>35</sup> We believe that the 1958 realignment was the result of falling travel costs, as was the 1979 formation of a nationwide league for the top semi-professional teams skimmed from regional units.

<sup>36</sup> From 12 teams to 86, the Football League’s average expansion rate between 1888 and 1921 was well over 2 teams per season. Over a comparable period following its birth, Major League Baseball expanded from 8 teams to 16, less than a quarter of a team per season. The season-by-season Football League records reported throughout this article are found at [http://www.the-english-football-archive.com/football\\_league.htm](http://www.the-english-football-archive.com/football_league.htm).

<sup>37</sup> This is a sports-related example of a more general claim that there are innate attitudinal differences between British and American businessmen. See Chandler, *Scale and Scope*.

<sup>38</sup> On the transition, see Goldstein, *Playing for Keeps*, regarding baseball, and Vamplew, *Pay Up*, regarding soccer.

the United States focused intently on covering expenses before football teams did so in England, but that reflected the higher costs that U.S. teams incurred rather than a difference in outlook.

### **What can explain this puzzle?**

On both continents, as a few semi-professional teams grew stronger than their local competitors, they began to look farther for worthy opponents. But why, despite similar origins, did the teams that formed the Football League opt for promotion & relegation and forego the benefits of a closed cartel that appeared to benefit the National League? Our explanation focuses on compactness (economic geography) and its impact on travel on the cost side, and the excitement level of games and territorial monopolies on the benefit side.

*Compactness*— For top-flight teams, uncoordinated scheduling proved easier in densely populated England than in sparsely populated America, accounting perhaps for the Football League's later formation. Working-class English clubs had more opponents nearby. They could easily arrange games against clubs at locations that permitted round-trip travel within a day, requiring little disruption of the players' work schedules. A similar environment existed for many clubs in urban America, but the best ones discovered that they could opt to travel short distances and play what were often one-sided games against inferior opponents before few fans or travel long distances to meet comparably strong teams. Long-distance travel disrupted players' work at their full-time jobs, so the top echelon of U.S. baseball became professional much earlier than any English football club.<sup>39</sup>

---

<sup>39</sup> Members of some early baseball teams held jobs, but those permitted latitude for considerable time away from work, if indeed any time at work was actually expected. Many New York Mutuels held patronage positions with the city, while nearly all Washington Nationals were U.S. government employees.

Old England is substantially smaller than New England, as shown in Figure 1 where England and Wales are shown at the same scale as the U.S. The area of England and Wales together is 58,309 square miles; the area of New England is 66,672 square miles. The average of the 48 contiguous states is 62,944 square miles. According to the 1890 U.S. and 1891 British censuses, Old England contained roughly half as many people as the entire United States. Even today England and Wales hold more than a fifth of the U.S. population in an area almost the size of Michigan. So the Football League plays in an area that would comprise an average state.

The Football League covers only a middling area among the major soccer leagues of Europe. England and Wales are roughly the size of Greece (51,146 sq. mi.). They are roughly twice the size of Scotland (30,045 sq. mi.), but less than half the size of Italy (116,314 sq. mi.). Even the French league, the one that spans the largest range (211,208 sq. mi.), covers an area less than half that of the northeast quadrant of the U.S. (429,989 sq. mi.), where the National League began play.<sup>40</sup> Due to their greater population densities, European nations were blessed with a denser rail network than the United States.

To illustrate that team travel costs would be less in England, we defined a sample of every team that ever played in the Premier League. There have been 61 Premier League teams in 42 different cities.<sup>41</sup> Today 30 teams play Major League Baseball in 26 different cities. The average distance between the English cities in the sample, 70 miles (110 km.), barely exceeds ten percent of the 660 miles (1060 km.) separating the American cities.<sup>42</sup> The average distance

---

<sup>40</sup> France has roughly the area of California or Texas. While Russia is the largest European nation by area, all notable Russian soccer teams play in Moscow.

<sup>41</sup> One is now disbanded, while four others are playing this season below the four Football League levels.

<sup>42</sup> English distances are rounded to the nearest five miles, while American distances are rounded to the nearest ten.



between the 12 cities in the original Football League was 30 miles (50 km.), while the average distance among the original National League cities was ten times greater. The *maximum* separation possible between two Football League teams is less than the *average* separation between teams in Major League Baseball, though with today's travel costs it barely matters.

The furthest that two English teams could travel for a league game is the 400 miles (650 km.) between Plymouth and Newcastle, though those teams have played in the same division during only 14 of the Football League's 105 seasons. That distance is about the same as from Boston to Baltimore, or Los Angeles to San Francisco. Manchester and Liverpool are 35 miles (55 km.) apart. Each of those cities is within 100 miles (160 km.) of Birmingham, similar to the distance between Chicago and Milwaukee but less than the minimum separation between National League teams in 1876.<sup>43</sup> Shorter English distances meant lower costs for moving a Football League team and lower costs for the fans who traveled with it.

It should be clear that compactness led to lower travel times and costs, especially when those costs include the opportunity cost of the labor. Assume the earnings foregone while actually playing one game of football (as distinct from the earnings foregone while traveling to and from it) are the same as for playing one game of baseball. Foregone earnings swamp direct travel costs, but information assembled by Michael Hauptert and Kenneth Winter from the account books of the New York Yankees suggests that even the direct travel costs for that team in 1915, the first year data are available, were equivalent to the annual salaries of five players.<sup>44</sup> We have calculated the total time for a team to travel round-trip to another city and play an average game. We took the intercity mileage, assumed average train speeds of 35 mph, then

---

<sup>43</sup> Subsequently the average separation between major league baseball teams has gone up substantially, not down.

<sup>44</sup> Hauptert and Winter, "Pay Ball."

added 6 hours (3 hours for getting to and from the stations at each end, dressing and warm-ups, and undressing and treating injuries, plus 3 hours for playing the game).<sup>45</sup> Of the 67 city pairs in the Football League, 54 of them were such that a round trip could be completed in 12 hours or less, while another 12 required 14 hours or less. Only one, Plymouth and Newcastle, required as much as 30 hours. For the National League, of the 28 city pairs, only 4 involved a round trip of 14 hours or less (13 hours being the minimum) while 21 required a trip of at least 24 hours.

The average round trip in the Football League (including game time) was 10.9 hours; the average in the National League, 41 hours. Since strong competitors were nearby, English football teams were able to travel to an away game on Saturday morning, participate in the game during the afternoon, and return home Saturday evening. Their fans could do the same. When the Football League formed, there was a great advantage to eschewing the advantages of a closed cartel of territorial monopolies. Separating the teams to form those monopolies would have meant paying the players more, since they would have had to work less at their jobs or quit them altogether. Separating the teams into territorial monopolies would have prevented many fans from traveling with a team, reducing team revenues while increasing costs (wages and travel). By contrast, baseball teams had to travel so far that their players became fully professional; they could no longer maintain outside employment during the season. Few supporters could travel to away games, so attendance was almost exclusively local fans. Thus, when the National League formed, there was an advantage to accepting territorial monopoly.

European fans have always attended away games in large numbers, and from the beginning most Football League games were scheduled to enable fans to leave work Saturday at noon (the end of the British workweek until the 1950s), and return home before bedtime. Seldom

---

<sup>45</sup> This assumption is based on timetables in *Travellers' official railway guide* for the United States (1899) and *Bradshaw's guide* for Great Britain (1882).

does a Football League match require a night away from home. In contrast to all North American sports, night games remain rare in the Football League. With air travel, the direct travel cost difference between the two continents has fallen substantially, but even today U.S. teams and fans generally must incur at least an overnight stay.

*Excitement level of games*—The excitement level of games is a function of parity and variety/rivalry. Parity is a virtue. Its benefits have been discussed for some time—the success of any league depends on staging games with uncertain outcomes.<sup>46</sup> If outcomes become too predictable, games become boring, and few fans will bother to attend. On the other hand, sports fans prefer that their team win. We characterize parity as having positive but decreasing marginal value. Most fans would prefer that the probability of the home team winning any particular game is greater than 50%, but as the home team’s advantage grows, attendance begins to wane.<sup>47</sup> Consider one team in a two-team market. The other local team offers low travel costs, but two teams playing repeatedly expose parity differences, which hurts attendance. So, variety is also a virtue, again with positive but decreasing marginal value; a team with too many opponents plays none often enough. Variety and rivalry are negatively correlated. If rivalry also has positive but decreasing marginal value, an optimal league size will exist where adding a team leads to a loss of passion that overwhelms the weak benefit of added variety, but the increased rivalry due to subtracting a team fails to compensate for the ennui from repetition.

To produce games with uncertain outcomes in the sparsely populated, relatively rail-poor U.S., the National League restricted membership to one team in each of the eight cities shown on

---

<sup>46</sup> See, for example, Canes, “Social Benefits.”

Figure 1. The National League might have been begun with even fewer teams, but that would have involved too much reduction in variety.<sup>48</sup> Although the Football League began in the midsection of England with twelve teams spanning less than 120 miles (210 km.), parity in that league tended to be higher than the National league had been at a similar stage of development. Even before admitting a second division, the Football League had expelled Stoke City (the loser in 29 of the 44 games it played across two seasons) and replaced it with Sunderland (the champion during its second season in the league). With far more opponents of roughly equal ability within a given radius, variety was less costly to English teams, so, until reaching long-run equilibrium, any division of the Football League would be predicted to have more teams than would baseball. This requires a relatively strong assumption for cross-section analysis—omitted variables have similar impacts on baseball and football.<sup>49</sup> Testing that hypothesis is trivial. The Football League began with a larger membership and grew to 28 teams within four years. Despite a much larger U.S. population, baseball was well over a century old before 28 teams were eligible to compete for its championship. Even today Major League Baseball has only 30 clubs, of which 22 teams have won a World Series title. 105 teams have played in the Football League, 92 are there at present, 61 of which have reached the top division at some point, and 23 have won its championship. From 1888 to the present, the inverse Herfindahl yields an

---

<sup>47</sup> David Surdam, “The American ‘Not-So-Socialist’ League,” has shown that revenue sharing benefited the New York Yankees in their post-World War II glory days as they were a far better road than home draw.

<sup>48</sup> Cincinnati won less than one game of every seven they played during the National League’s inaugural season, while Chicago won nearly four of every five. The prospect of watching Chicago play Cincinnati must have held little excitement for fans in either city. The league contracted to six teams for its second and third seasons.

<sup>49</sup> The assumption is unnecessary for time-series analysis. The model yields a great number of predictions that we have had insufficient time to quantify. Football League teams, for instance, should rarely play long, uninterrupted series of games at home or away, whereas baseball teams would play them as a matter of course. cursory examination indicates that such is the case.

equivalent of 35.9 teams and 11.3 champions for the Premier League, but only 23.7 teams and 9.4 champions for Major League Baseball.<sup>50</sup>

Given our strong assumption that omitted variables have similar impacts, parity was less costly in England at the outset, so the Football League would have been the more balanced league.<sup>51</sup> Even without such an assumption, falling travel costs would raise parity on both continents, though more strongly for baseball where the initial costs were greater. Testing the parity prediction is more challenging than testing the variety prediction. The measure of parity we employ is the proportion of clubs that win between 40% and 60% of the maximum points (football) or games (baseball) available to a team in a season. For consistency we adjusted the official tally after 1981 when the Football League began awarding three rather than two points for a win. Draws have always been worth one point in soccer and impossible in American baseball, though Japanese baseball leagues permit tied games.<sup>52</sup> Figure 2 contrasts Premier

---

<sup>50</sup> The inverse Herfindahl as defined by George Stigler, *Organization of Industry*, p. 31, is “the equivalent number of firms of equal size” that would produce that Herfindahl. For the number of teams, the share is the number of years that a given team has competed divided by total team-years of competition. For the number of championships, the share is the number of championships a given team has won divided by the total number of championships awarded. The data used for Major League Baseball includes ten seasons of American Association data from 1888 until its demise, thus biasing the computation against confirmation of our hypothesis.

<sup>51</sup> Eckard, “Team Promotion,” argues that parity in Major League Baseball was greater in the years before the National League-American Association “merger,” and that the closed cartel was adopted “by owners seeking to protect themselves from that particular type of competition” (146). This raises the question why Football League owners did not seek similar protection, a question that this article addresses, but Eckard does not. Cain and Haddock have examined long-run parity in professional baseball from the emergence of the National Associate to the present. As Figure 1 of “Measuring Parity” shows, parity has generally trended upward, though with substantial variance. Parity began the decline that Eckard mentions not in 1903 but with the American League’s transition from minor to major league in 1901. Parity continued to decline until 1907, but then began a recovery, reaching an unprecedented level in 1915. Major league parity has never since fallen as low as it was in 1903. A similar transitory dive in parity had followed the creation of the American Association in 1882.

<sup>52</sup> The idealized standard deviation, devised to study baseball (Quirk and Fort, *Pay Dirt*), is a commonly used measure of parity. Cain and Haddock, “Measuring Parity,” explore the

League and Major League Baseball parities at similar league ages. The curves are offset by 17 years, the difference between the 1871 inauguration of the National Association and 1888 when the Football League was formed.<sup>53</sup>

We ended this analysis with the 1955-56 Football League season due to several exogenous events that seem to have affected parity adversely—the inauguration of the Champions Cup, then the Champions League and the UEFA Cup, and/or a court decision that eliminated maximum salaries in the Football League. In the Champions Cup/Champions League and UEFA Cup the best few teams from the top division of all European leagues (plus an occasional entrant from lower divisions) compete against each other the following year, with contests staged in mid-week and interwoven with the weekend league competitions. Those European competitions provide an additional bonus for excelling—a lucky few teams play additional games against similarly skillful foreign clubs. The games are played on mid-week evenings against unusually distant clubs; fewer visiting fans attend. The extra games are more valuable to teams with strong local followings. After 1956 those teams had an added incentive to invest in team quality that was lacking in those teams cyclically promoted to the lower reaches of the Premier League or relegated to the upper reaches of Division One.<sup>54</sup>

---

heretofore-unrecognized errors that arise when applying that alternative to soccer and other sports that permit games to end in ties. Statistical analysis based on the idealized standard deviation (appropriately corrected for use with Football League statistics) leads to identical conclusions to those reported here.

<sup>53</sup> The effect remains when we use chronological dates, but is less easily discerned visually.

<sup>54</sup> We ran preliminary regressions of parity against league age that included dummy variables for both the Champions Cup/Champions League and the UEFA Cup. At conventional confidence levels, those regressions suggest that both had a negative impact on parity in the Football League's Division One, while the Champions Cup/Champions League had an even stronger impact on the Premier League, though the UEFA Cup had little. Since most English participants in both competitions come from the upper half of the Premier League, the results vis-à-vis Division One are surprising. Chow tests lead to strong acceptance of the hypothesis that there

A regression of the ratio of Major League Baseball parity to Premier League parity against league age suggests convergence, the ratio increasing from 0.673 at the start to 0.910 at the end. The statistical confidence that the increase of 0.004 per year differs from zero exceeds 95 percent.

To produce a more attractive product—one with both variety and parity—the National League frowned on entry. Assuming that the league started with the appropriate teams, potential applicants either were too far away (say, in San Francisco) implying travel that was too time-consuming and costly, or were in a market that was thought to be unable to support a suitable team. Moreover, the National League frowned on exit. From the outset, games against non-league teams were impeded, later banned entirely; any team that left the National League would experience difficulty scheduling worthy opponents. Thus, the National League was pinched on the variety margin despite having included a few fringe members, as it was pinched on the parity margin because it included them. In contrast, non-league games have always been a staple of Football League seasons, even institutionalized in annual competitions such as the FA Cup.

The Football League could be blasé about the danger that an owner would weaken his team by capitalizing the contracts of star players. The problem would be dealt with promptly as the weakened team would be relegated to a division more in keeping with its reduced ability. Further, the Football League could also be relaxed about turnover because its greater density means that a defecting Football League team would be, and is, relatively easy to replace. If a North American team attempted to capitalize its roster, however, the impact on parity and the external cost on other teams could prove long-lived.

---

was no break between 1955 and 1956 for Major League Baseball, weakly rejects the hypothesis for the Premier League, and strongly rejects it for Division One.

*Territorial monopolies*—The National League seized the opportunity to form territorial monopolies even as it formed the league. It built a wall that was penetrable only at league discretion. Despite considerable turnover over the first few seasons, the league was closed at birth to a degree that the Football League never has been.<sup>55</sup> Monopoly successfully achieved, even if for meritorious purposes, may provide market power. When an additional market grows to a size adequate to support a major league caliber team, an entrepreneur there faces a hard choice—organize an entire league with opponents who are sited in initially substandard markets (risking piecemeal preemption by major league expansion as those markets grow) or convince the league to create an expansion franchise and offer it for sale.

An English or Welsh football team—even a non-league team—need only keep winning to reach the Premier League. Alternative leagues were easy to form, as the mergers of the Football League with the Football Alliance and Southern League illustrate. When top-flight English players became fully professional, the foregone wage factor evaporated, but by then the multi-divisional promotion & relegation system was firmly entrenched. By 1921 the Football League already had 86 teams, of which 33 had played in the Premier League. To want an effective closed cartel is one thing; to create one quite another. It seems that the Football League could not have succeeded, and it never tried. If the league moved to a 40-mile-radius territorial monopoly for small cities and territorial duopoly for more populous ones, the Football League would have to expel ten of the twelve League teams playing in London this season, along with four more from London’s suburbs.<sup>56</sup> 20 of the 22 teams within a 40-mile radius of Manchester

---

<sup>55</sup> Two of the very best professional baseball teams played their home games in Philadelphia. Incorporating both into the National League would have improved parity, but defeated a territorial monopoly in Philadelphia. Only one was admitted.

<sup>56</sup> The Baltimore Orioles have argued that they deserved compensation for mitigation of territorial rights for Major League Baseball’s decision to relocate the Montreal baseball team to



(including Liverpool) would similarly have to be expelled, as would seven of nine teams in and around Birmingham, six of eight in Yorkshire, and five of seven near Newcastle. An insurgent league would have more established teams to choose from than the Football League would have kept for itself.

Once a league has reached the size at which the marginal value of variety and rivalry were offsetting, further expansion becomes unattractive. Teams in markets with the strongest support benefit from promotion & relegation with little risk. Minor random year-to-year fluctuations in team quality mean that promotion & relegation will increase parity in each division. That certainly benefits continuing members of the upper division, and perhaps some in the lower divisions as well, though it imposes a cost on the relegated teams. Though the Football League is definitely fluid, with no team avoiding at least occasional relegation, a core has enjoyed frequent membership in the top division, and, while doing so, promotion & relegation has provided them with greater parity.<sup>57</sup>

As long as travel costs are significant, our analysis predicts Football League team survival in smaller markets than are necessary to support baseball teams, and collaterally English metropolitan areas that possess more teams for a given population. Given England's size, every Football League game would be against a neighbor (at least from a North American perspective). The lack of territorial monopoly increases the price elasticities Football League teams face when

---

Washington, 40 miles (65 km.) from Baltimore. The mayor of San Jose, California, hopes to gain a major league team for his city, but San Francisco, 50 miles (80 km.) to the north, was given territorial rights to San Jose when the Giants moved west in 1958.

<sup>57</sup> No team has played its entire history in the Premier League, but 32 teams have played there for more than a quarter of the Football League's 105 seasons, 19 for more than half, and five for more than three-quarters.

marketing tickets and local broadcast rights, but games draw fans from two markets.<sup>58</sup> In contrast, most of baseball's attendance is drawn from the territory over which it has a monopoly.

Inter-country differences in market size are difficult to quantify because the United Kingdom and the United States do not define metropolitan areas in analogous ways. Comparing the "conurbations" as defined in the 2001 UK Census and the "combined statistical areas" (CSA) defined in the 2000 U.S. Census makes London appear to be about the size of the San Jose-San Francisco-Oakland, CA, CSA, the U.S.'s 5th largest. When we used a road map to define our own ersatz London CSA of roughly the same area as Chicago's, London looms much larger, but is nonetheless less populous than New York City. London has six teams playing in the Premier League this season and seven other functioning teams that have done so in the past (including Wimbledon that now plays in Milton Keynes), while New York City has only two Major League baseball teams, the two other survivors that once played there have moved to California. Our imaginary Lancashire-Cheshire CSA has a smaller population than Chicago's CSA, but has six Premier League teams, as contrasted with Chicago's two Major League clubs.

Table 1 lists the conurbations and cities with teams in the Premier League and Division One by size over time. The number of teams in the original Football League, and the number of teams today give an indication of the growth of the league, with the last four columns indicating when that growth took place. The conurbations plus Nottingham had multiple teams with

---

<sup>58</sup> Though European football leagues enforce no territorial monopolies for their teams, FIFA, the world governing body, enforces territorial monopolies for leagues. A league in one nation can add a team from a different nation only if the latter has no domestic league. A violation would result in disqualification from FIFA-sanctioned events such as the World Cup for the national team and the Champions League for club teams. Thus Monaco plays in the French league and Auckland in the Australian league, but a separate Scottish league formed in 1890, and the two major Glaswegian teams, Celtic and Rangers, have been stymied in recent bids to join the English Football League despite a dearth of strong Scottish opponents. Cardiff, Swansea City, Wrexham, and a few other Welsh teams had entered the Football League before the Welsh League began play, and are grandfathered in.

Premier League experience, and there are other Football League teams in many of these areas yet to be promoted that far. The pattern is extremely varied, especially in comparison with North America's. An area with over 7 million is host to 13 teams, but an area with less than 100,000 also supports a team.

Table 2 shows the U.S. experience from the founding of the National League until the 1903 Major League Agreement, with urban areas ranked by the 1890 Census. The table includes all teams playing in the top leagues (the National League and the American Association; the latter denoted with an asterisk) on a year-by-year basis. The two important features of this table are 1) the movement of teams toward the largest markets, comparable to the Football League movement toward West Yorkshire and later London, and 2) no city had more than two teams, counting Brooklyn as a separate city through much of this period.

Table 3, where the major urban areas of the U.S. today with Major League Baseball teams are ranked at four census years corresponding with the Major League Agreement, the beginning of relocation and expansion, and the league today.<sup>59</sup> Markets with teams at each date are indicated in boldface type; those without are in italics. In this table, the process of transfer and expansion that led to teams in an increasing number of major markets is evident. Milwaukee, the smallest market with a team currently in the league, is roughly 19 times larger than Burnley, the smallest city with a team in the Premier League or Division One.

In 1900, teams were in most major markets of the northeast, conveniently on the rail network. Baltimore and Buffalo (the 8<sup>th</sup> largest area in 1900) were anomalous, but their histories

---

<sup>59</sup> We stopped with the 56 largest areas in 1900. Miami, Phoenix, San Diego, and Tampa-St. Petersburg were all smaller at that time.

militated against their value as an entrant, at least at that time.<sup>60</sup> In 1900 San Francisco and New Orleans (the 12th largest area) were too distant to be included. The sixteen teams after the Major League Agreement played in ten cities, and the smallest ranked 15th. Those teams remained in precisely the same cities for half a century, by which time the smallest market still ranked 17th.

By 1950 the minor Pacific Coast League (PCL) appeared to be on the verge of forcing its way to major league status on the basis of the region's rapidly growing population and wealth. The PCL had four teams playing in urban areas that ranked among the nation's ten largest and two more ranked in the top 25. Moreover, all of those markets were growing rapidly. Major League Baseball preempted the threat by jumping to the West Coast in 1958, transferring two of its three New York City franchises as soon as the emergence of air transport made such a move feasible. During the 1950s, the smallest of the original major league markets had fallen to 21st, while air travel between New York and Los Angeles was more convenient than travel by rail between New York and Chicago had been in 1900.<sup>61</sup>

When confronted by a growing number of major markets without teams, enough markets perhaps to comprise a third league, Major League Baseball began a program of preemption, first by relocating franchises from its multi-team cities, then by expansion. By 2000, teams were in the top 22 markets and two smaller ones. Without going head-to-head with the existing league no entrant could have a team in any market ranking higher than 23rd. Only a few large U.S.

---

<sup>60</sup> Baltimore had been the home of several teams over the years, but failed to keep any; Buffalo had been given a seven-year opportunity to support a National League team, but, like Baltimore, could not keep it.

<sup>61</sup> The PCL had two teams in Los Angeles and one each in San Francisco, Oakland, Seattle, Portland, San Diego, and Sacramento, in order of their populations in 1950. Counting Anaheim with Los Angeles (as the Los Angeles CSA does), five of those seven cities now house relocated or expansion Major League Baseball teams.

markets are on the outside, hat in hand, hoping wanly for team movement or expansion, almost begging to finance fancy new facilities.

In sum, our argument is that the tradeoff between costs and benefits was different for the two leagues. Since, for the Football League, close competition required far less travel, and because compactness provided a means of controlling players' wages and increasing attendance, the potential additional revenues from territorial monopolies were less than the increased costs would have been. For the National League, a beneficial degree of compactness was unattainable, high travel costs were unavoidable, so it opted to exploit the demand-side benefits of territorial monopoly. By the time the top levels of the Football League became fully professional, the number and location of teams made reversion to territorial monopoly infeasible.

## **Conclusion**

A sports league is necessarily a cartel due to the mutual interdependence among the teams. The revenue that one team receives is correlated with the on-field performance of the others. Avoiding opportunistic behavior requires centralized decision-making to define rules of competition and cooperation, and to discipline teams that violate the rules. That, however, fails to tell whether the cartel will be open or closed. Subject to the constraints the league faces, it will attempt to enhance the revenues of member teams by promoting parity, variety, rivalry, and visits by fans to opposing stadiums, by shielding teams from competition, and by selecting a structure that reduces costs. Those motivations conflicted with each other as long as travel costs were high. Parity, variety, and rivalry were more expensive in North America than in most of Europe. The U.S. is less compact, so attracting visiting fans was nearly impossible; a league's

ability to control labor cost by using semi-professional players was weak. High travel costs were hardly an unmitigated curse because they provided the opportunity to form territorial monopolies that reduced demand elasticities. The limited number of unserved territories at any time enabled incumbent teams to extract entry fees from new franchises, and to reduce their costs by extorting stadium construction from political units due to credible threats to shift franchises.<sup>62</sup>

By the late nineteenth century, bodies existed to rule the sports on both sides of the Atlantic. The Football League wanted to cooperate with the Football Association, while the National League revolted against the National Association of Professional Base Ball Players, which only five years earlier had revolted against its amateur predecessor. In North America, the National League was formed as a closed cartel to help differentiate its product from that of the National Association and to establish a nationally recognized brand. To market it, they sought to impede rival teams' efforts to compete against the clubs the National League selected. The outcome was a league structure that was successfully challenged once, and then remained static for a half century.

The Football League began as an open cartel which allowed the teams a) to continue their involvement in the FA Cup competition and for their players to remain involved in international competition, both of which were sponsored by the Football Association, and b) to allow for the entrance of additional members—eventually including ones located in the populous south of England. While the Football League was careful regarding entry due to the potentially adverse impact on members' long-term financial viability, it was a muted concern. England's short distances meant that gate receipts included a significant portion collected from fans of the

---

<sup>62</sup> Designed as a dual use (football/baseball) stadium, the Louisiana Superdome in New Orleans was completed in 1975. A football team immediately began to play there, but a major league baseball team has yet to arrive.

visiting team. Compactness meant that the league employed semi-professional players for decades. Even today many semi-professional players compete for teams one level removed from the League. In England, the outcome was a league structure that expanded quickly; promotion & relegation added flexibility.

Our historical explanation leaves many questions for future work. Perhaps the most interesting of these is why leagues in other sports on a continent adopted the structure of the first league.<sup>63</sup> Later leagues benefited by the learning-by-doing of the pioneer leagues, but they understood the benefits and costs of both structures. The factors discussed above remained relevant, but the weights changed, particularly in North America, as there were increased population densities, increased real wealth, and diminished travel costs.

Population densities on opposite shores of the Atlantic have converged since the National League began. Since the middle of the twentieth century air travel has reduced costs dramatically. But for the *Chase* (1914) and *Federal Baseball* (1922) decisions, antitrust pressure might well have opened the closed North American cartels. G. Edward White comments,

“the [*Federal Baseball*] decision appears strikingly out of touch with its subject matter, treating baseball as an amateurish ‘game’ when it was a highly professionalized business. ... [T]he decision, viewed as an effort to characterize the actual legal and economic

---

<sup>63</sup> A cursory survey of other continents reveals a mixed protocol in Japan, where the football league has two divisions that practice promotion & relegation, while Japanese baseball leagues have kept fixed rosters since the early 1950s; unlike North American teams, they enjoy no territorial monopolies. Argentina’s football league employs promotion & relegation according to results aggregated across three seasons rather than one. Brazilian teams are widely spread across a vast area, and employ a third protocol that emphasizes knockout tournaments, while league play is regional and attracts less interest. During the latter half of the nineteenth century English football underwent an evolution similar to that now taking place in Brazil.

character of major league baseball in 1922, was remarkably myopic, almost willfully ignorant of the nature of the enterprise.”<sup>64</sup>

If professional sports leagues had evolved only after World War II, promotion & relegation might now be universal, but instead the evolution of North American sports leagues appears to provide a crystalline example of path dependence. The existing structure is valuable to those within it, so change will only come from outside. The evolution of antitrust law has precluded that; it has substantially shielded the top-rank North American leagues from entry in every sport. The story of why such continent-wide duplication took place requires a good deal more research into the specific details surrounding the formation of later leagues.

An important antitrust question arises because teams cannot enter the sports industry without opponents, but North American professional sports leagues endeavor to keep few markets open where an entrant’s teams would be adequately supported at a top-flight level. Rather than bear the parity cost of forming an insurgent league and playing unworthy opponents, most aspirants elect to pay a substantial franchise fee to enter the existing cartel or else turn to non-sports pursuits. Potential entrants are a revenue source for incumbents rather than a competitive threat. Economic competition—as distinct from game competition—is between different sports rather than between different teams playing the same sport. Different sports often have different fans and are weak substitutes, but would continue to compete for fans if promotion & relegation were adopted in North America.

---

<sup>64</sup> White, *Creating the National Pastime*, p. 70. *American League Baseball Club v. Chase*, 149 N.Y.Supp. 6 (1914); *Federal Baseball Club of Baltimore, Inc. v. National League of Professional Baseball Clubs, et al.*, 259 U.S. 200 (1922), affirming 269 Fed. 681 (App. D.C. 1921), which had reversed the trial court. With but a handful of exceptions, North America’s closed cartels are and always have been outliers among the world’s professional sports leagues.



There are also important public choice questions because entry into and relocation within a North American league require the approval of those already there. A team's threat to relocate, leaving its present region indefinitely bereft in that sport, leaves North American governments vulnerable to rent seeking. Many teams enjoy publicly financed stadiums and parking concessions for little or no rent, though the teams could bear the cost of constructing their own, just as most European teams do. No North American professional league—not even soccer—employs promotion & relegation, though both major and minor league teams readily relocate (often repeatedly) to distant cities.

No analysis of the contrasting industrial structures can be complete without understanding the historical origins of the entrenchment of one system on one side of the Atlantic while it has been so energetically rejected on the other.

## Bibliography

- Arnold, Tony. "Rich man, poor man: economic arrangements in the Football League." In *British Football and Social Change: Getting into Europe*, edited by John Williams and Stephen Wagg, 48-63. London: Leicester University Press, 1991.
- Travellers' official railway guide for the United States, Canada and Mexico*. Edited by W.F. Allen. New York: National Railway Publication Co., December 1899.
- Baade, Robert, and Alan Sanderson. "Field of Fantasies," *Intellectual Ammunition*, <http://www.heartland.org/archives/ia/marapr96/cover.htm> (1996).
- Bradshaw's general railway and steam navigation guide, for Great Britain and Ireland. London: W. J. Adams & Sons, August 1882.
- Cain, Louis P., and David D. Haddock. "Measuring Parity: Tying Into the 'Idealized' Standard Deviation." *Journal of Sports Economics* forthcoming.
- Canes, Michael E. "The Social Benefits of Restrictions on Team Quality." In *Government and the Sports Business*, edited by Roger G. Noll, 81-114. Washington: The Brookings Institution, 1974.
- Chandler, Alfred D., Jr. *Scale and Scope: the Dynamics of Industrial Capitalism*. Cambridge: The Belknap Press of Harvard University Press, 1990.
- Davis, Lance E. "Self-Regulation in Baseball, 1909-1971." In *Government and the Sports Business*, edited by Roger G. Noll, 349-386. Washington: The Brookings Institution, 1974.
- Eckard, E. Woodrow, "Team Promotion in Early Major League Baseball and the Origin of the Closed Sports League," *Explorations in Economic History* 42, no. 1 (2005): 122-152.
- Goldstein, Warren. *Playing for Keeps: A History of Early Baseball*. Ithaca: Cornell University Press, 1989.
- Harris, Harold Arthur. *Sport in Britain: Its Origins and Development*. London: Stanley Paul, 1975.
- Hauptert, Michael, and Kenneth Winter, "Pay Ball: Estimating the Profitability of the New York Yankees, 1915-1937" (working paper, 2004).
- Infoplease, <http://search.netscape.com/> (2004).
- Kirsch, George B. *The Creation of American Team Sports: Baseball and Cricket, 1838-72*. Urbana: University of Illinois Press, 1989.
- Knautz, Robert, editor. "Franchise Free Agency." *Policy Spotlight*, <http://www.free-market.net/features/spotlight/9702b.html> (1997).
- Leeds, Michael and Peter von Allemn. *The Economics of Sports*. Boston: Addison Wesley, 2002.
- Mokyr, Joel. *The Lever of Riches: Technological Creativity and Economic Progress*. Oxford: Oxford University Press, 1990.
- Noll, Roger G. "The Economics of Promotion and Relegation in Sports Leagues," *Journal of Sports Economics* 3, no. 2 (2002): 169-203.
- Okner, Benjamin A. "Subsidies of Stadiums and Arenas." In *Government and the Sports Business*, edited by Roger G. Noll, 325-348. Washington: The Brookings Institution, 1974.
- Poitras, Marc, and Lawrence Hadley. "Do New Major League Ballparks Pay for Themselves?" *Journal of Business*, forthcoming.
- Quirk, James, and Mohamed El Hodiri, "The Economic Theory of a Professional Sports League." In *Government and the Sports Business*, edited by Roger G. Noll, 33-80. Washington: The Brookings Institution, 1974.
- Quirk, James, and Fort, Rodney, *Pay Dirt: The Business of Professional Team Sports*. Princeton: Princeton University Press, 1992.
- Rader, Benjamin G. *Baseball: A History of America's Game*, second edition. Urbana: University of Illinois Press, 2002.
- Rivkin, Steven R. "Sport Leagues and the Federal Antitrust Laws." In *Government and the Sports Business*, edited by Roger G. Noll, 387-410. Washington: The Brookings Institution, 1974.

- Ross, Stephen F., and Stefan Szymanski. "The Law & Economics of Optimal Sports League Design." University of Illinois College of Law, Illinois Public Law and Legal Theory Research Papers Series, Working Paper No. 03-14, 30 September 2003.
- Seymour, Harold. *Baseball: The Early Years*, 3 volumes. New York: Oxford University Press, 1960.
- Stigler, George. *The Organization of Industry*. Homewood, IL: Richard D. Irwin, 1968.
- Surdam, David. "The American 'Not-So-Socialist' League in the Postwar Era: The Limitations of Gate Sharing in Reducing Revenue Disparity in Baseball." *Journal of Sports Economics* 3, no. 3 (2002): 264-290.
- Szymanski, Stefan. "The Economic Design of Sporting Contests." *Journal of Economic Literature* 41, no.4 (2003): 1137-1187.
- Taylor, Matthew, and John Coyle. "The Election of Clubs to the Football League, 1888-1939." *The Sports Historian*.19, no. 2 (1999): 1-24.
- U.S. Congress 82d Congress, 2d Session, House Report No. 2002. *Organized Baseball*. Report of the Subcommittee on Study of Monopoly Power of the Committee of the Judiciary. Washington: U.S. Government Printing Office, 1952.
- Vamplew, Wray. *Pay Up and Play the Game: Professional Sport in Britain, 1875-1914*. Cambridge: Cambridge University Press, 1988.
- Van Vleck, Va Nee L. "Reassessing Technological Backwardness: Absolving The 'Silly Little Bobtailed' Coal Car." Iowa City: University of Iowa Ph.D. dissertation, 1994.
- Wagg, Stephen, ed. *Giving the Game Away": Football, Politics and Culture on five Continents*. London: Leicester University Press, 1995.
- White, G. Edward. *Creating the National Pastime: Baseball Transforms Itself, 1903-1953*. Princeton: Princeton University Press, 1996.
- Young, Percy M. *A History of British Football*. London: Stanley Paul, 1968.

Figure 1

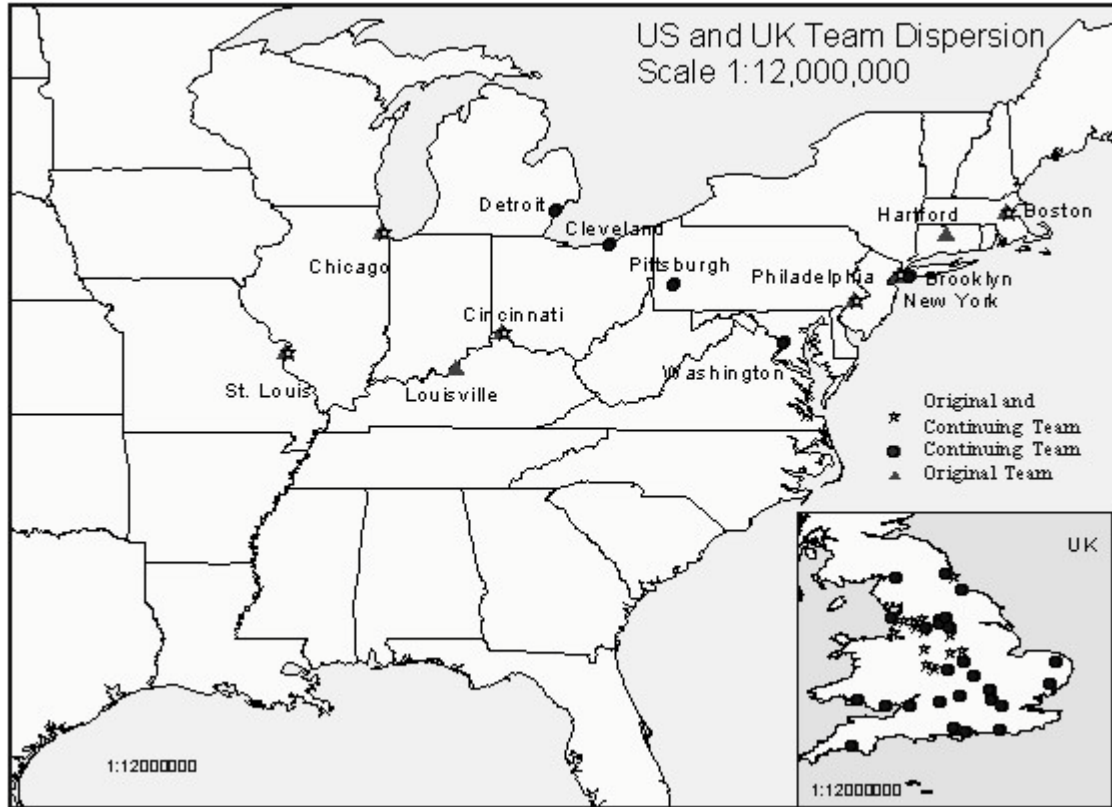
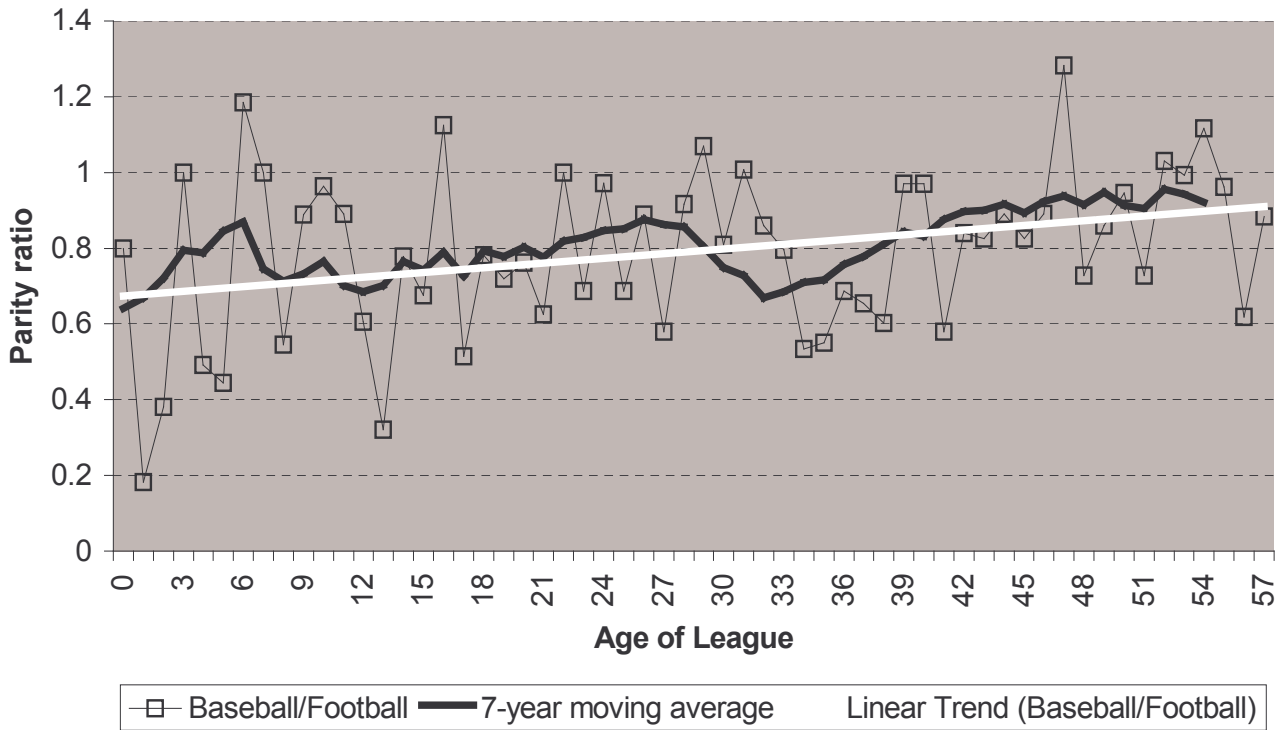


Figure 2

Parity of Baseball to Football, 1888-1956



**Table 1****Football League Teams by Area, 1888-2004**

Rank	2001 Population	Original Teams	Teams Today	1888- 1900	1901- 1915	1919- 1939	1946- 2004	
<b>Conurbations</b>								
1	London	7,172,036	0	13	1	5	10	13
2	West Midlands	2,555,596	3	5	4	4	5	5
3	Greater Manchester *	2,482,352	1	5	5	6	5	5
4	West Yorkshire	2,079,217	0	4	0	3	4	3
5	Merseyside	1,362,034	1	2	2	2	2	2
6	South Yorkshire	1,266,337	0	3	3	3	3	3
7	Tyne & Wear	1,075,979	0	2	2	2	2	2
<b>Cities</b>								
1	Bristol City	380,615	0	1	0	1	1	1
3	Cardiff	305,340	0	1	0	0	1	1
4	Leicester	279,923	0	1	1	1	1	1
5	Nottingham	266,995	1	2	2	2	2	2
7	Brighton & Hove	247,820	0	1	0	0	0	1
11	Stoke-on-Trent	240,643	1	1	1	1	1	1
12	Swansea	223,293	0	1	0	0	1	1
13	Derby	221,716	1	1	1	1	1	1
14	Southampton	217,478	0	1	0	0	1	1
16	Northampton	194,477	0	1	0	0	0	0
19	Portsmouth	186,704	0	1	0	0	1	1
20	Luton	184,390	0	1	1	0	1	1
22	Swindon	180,061	0	1	0	0	0	1
29	NE Lincolnshire	157,983	0	1	1	1	1	1
36	Blackpool	142,284	0	1	1	1	1	1
39	Blackburn w Darwen **	137,471	2	1	3	1	1	1
40	Middlesbrough	134,847	0	1	1	1	1	1
41	Oxford	134,248	0	1	0	0	0	1
45	Preston	129,642	1	1	1	1	1	1
46	Norwich	121,553	0	1	0	0	1	1
49	Ipswich	117,074	0	1	0	0	0	1
51	Carlisle	100,734	0	1	0	0	0	1
52	Burnley	89,541	1	1	1	1	1	1

\* includes Glossop West End

\*\* includes Accrington and Darwen

**Table 2**  
**Major League Baseball Teams by Area, 1897-1903**

city	1890rank																									
	1876	1877	1878	1879	1880	1881	1882*	1883*	1884*	1885*	1886*	1887*	1888*	1889*	1890*	1891*	1892	1893	1894	1895	1896	1897	1898	1899	1900	1901*
New York	1	1						2	2	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Brooklyn	1								1*	1*	1*	1*	1*	1*	2	1	1	1	1	1	1	1	1	1	1	1
Chicago	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2
Philadelphia	3	1						1*	2	2	2	2	2	2	2	2	1	1	1	1	1	1	1	1	1	2
St. Louis	4	1	1					1*	1*	1*	2	2	1*	1*	1*	1*	1	1	1	1	1	1	1	1	1	2
Boston	5	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	1	1	1	1	1	1	1	1	1	2
Baltimore	6							1*	1*	1*	1*	1*	1*	1*	1*	1*	1	1	1	1	1	1	1	1	1	1
Pittsburgh	7							1*	1*	1*	1*	1*	1	1	1	1	1	1	1	1	1	1	1	1	1	1
San Francisco	8																									
Cincinnati	9	1	1	1	1			1*	1*	1*	1*	1*	1*	1*	1	2	1	1	1	1	1	1	1	1	1	1
Cleveland	10				1	1	1	1	1				1*	1*	1	1	1	1	1	1	1	1	1	1	1	1
Buffalo	11				1	1	1	1	1	1																
New Orleans	12																									
Washington	13									1*		1	1	1	1		1*	1	1	1	1	1	1	1	1	1
Detroit	14						1	1	1	1	1	1	1	1												1
Milwaukee	15			1													1*									1
Newark	16																									
Minneapolis	17																									
Jersey City	18																									
Louisville	19	1	1					1*	1*	1*	1*	1*	1*	1*	1*	1*	1*	1	1	1	1	1	1	1	1	1
Rochester	20															1*										
St. Paul	21																									
Kansas City	22											1					1*	1*								
Providence	23			1	1	1	1	1	1	1	1															
Denver	24																									
Indianapolis	25			1						1*			1	1	1											
Albany	26																									
Columbus	27								1	1					1	1	1									
Syracuse	28				1											1*										
Worcester	29					1	1	1																		
Toledo	30									1*						1*										
Richmond	31									1*																
New Haven	32																									
Paterson NJ	33																									
Lowell MA	34																									
Nashville	35																									
Scranton	36																									
Fall River MA	37																									
Troy	38				1	1	1	1																		
Cambridge	39																									
Omaha	40																									
Atlanta	41																									
Memphis	42																									
Camden	43																									
Wilmington	44																									
Dayton	45																									
Grand Rapids	46																									
Reading PA	47																									
Trenton	48																									
Lynn MA	49																									
Charleston SC	50																									
Hartford	51	1	1																							
St. Joseph	52																									
Evansville	53																									
Los Angeles	54																									
Des Moines	55																									

Source: Seymour, vol. 1, pp. 363-364

**Table 3**  
**Major League Baseball Teams by Area, 1900-2000**

Metropolitan Areas					
Rank	1900	1950	1960	2000	2000 Population
1	<b>New York</b>	<b>New York</b>	<b>New York</b>	<b>New York</b>	21,361,797
2	<b>Chicago</b>	<b>Chicago</b>	<b>Los Angeles</b>	<b>Los Angeles</b>	16,373,645
3	<b>Philadelphia</b>	<i>Los Angeles</i>	<b>Chicago</b>	<b>Chicago</b>	9,312,255
4	<b>St. Louis</b>	<b>Philadelphia</b>	<b>Philadelphia</b>	<b>Washington-Baltimore</b>	7,538,385
5	<b>Boston</b>	<b>Detroit</b>	<b>Detroit</b>	<b>San Francisco-Oakland</b>	7,092,596
6	<i>Baltimore</i>	<b>Boston</b>	<b>San Francisco-Oakland</b>	<b>Philadelphia</b>	5,833,585
7	<b>Cleveland</b>	<i>San Francisco-Oakland</i>	<b>Boston</b>	<b>Boston</b>	5,715,698
8		<b>Pittsburgh</b>	<b>Pittsburgh</b>	<b>Detroit</b>	5,357,538
9	<i>San Francisco</i>	<b>St. Louis</b>	<b>St. Louis</b>	<b>Dallas-Fort Worth</b>	5,346,119
10	<b>Cincinnati</b>		<b>Washington</b>	<b>Houston</b>	4,815,122
11	<b>Pittsburgh</b>	<b>Cleveland</b>	<b>Cleveland</b>	<b>Miami *</b>	4,452,557
12		<b>Washington</b>	<b>Baltimore</b>	<b>Atlanta</b>	4,548,344
13	<b>Detroit</b>	<i>Baltimore</i>		<b>Seattle-Tacoma</b>	3,604,165
14	<i>Milwaukee</i>	<i>Minneapolis-St. Paul</i>	<i>Minneapolis-St. Paul</i>	<b>Minneapolis-St. Paul</b>	3,271,888
15	<b>Washington</b>			<b>Phoenix *</b>	3,043,878
16		<i>Milwaukee</i>	<i>Houston</i>	<b>Cleveland</b>	2,945,831
17		<b>Cincinnati</b>	<b>Milwaukee</b>	<b>St. Louis</b>	2,754,328
18				<b>San Diego *</b>	2,698,687
19	<i>Minneapolis</i>	<i>Seattle</i>	<i>Seattle</i>	<b>Pittsburgh</b>	2,525,730
20		<i>Kansas City</i>	<i>Dallas</i>	<b>Denver</b>	2,449,054
21		<i>Houston</i>	<b>Cincinnati</b>	<b>Tampa-St. Petersburg *</b>	2,395,997
22	<i>Kansas City</i>		<i>Kansas City</i>	<b>Cincinnati</b>	2,050,175
23	<i>St. Paul</i>	<i>Dallas</i>	<i>San Diego</i>		
24		<i>Atlanta</i>	<i>Atlanta</i>		
25	<i>Denver</i>		<i>Miami</i>	<b>Kansas City</b>	1,901,070
26			<i>Denver</i>		
27	<i>Allegheny</i>				
28		<i>Denver</i>			
29					
30					
31			<i>Tampa-St. Petersburg</i>	<b>Milwaukee</b>	1,689,572
32		<i>San Diego</i>			
33					
34					
35					
36	<i>Los Angeles</i>				
37		<i>Miami</i>	<i>Phoenix</i>		
38					
39					
40					
41					
42					
43	<i>Atlanta</i>				
44					
45					
46			<i>Fort Worth</i>		
47		<i>Tampa-St. Petersburg</i>			
48	<i>Seattle</i>				
49					
50					

Source: U.S. Census Bureau, Census 2000

**Table 8. Population in Combined Statistical Areas (CSAs) and Their Geographic Components in Alphabetical Order and Numerical and**  
Internet Release date: December 30, 2003

**\* Table 3a. Population in Metropolitan and Micropolitan Statistical Areas Ranked by 2000 Population for the United States and Puerto Rico: 1'**  
Internet Release date: December 30, 2003