Book review – National Pastime by Stefan Szymanski and Andrew Zimbalist

Reviewed by Phil Miller

As a graduate student in the mid 1990's, I began work on a project on the economics of baseball for a research methods class. In researching the existing literature, I read Andrew Zimbalist's 1994 classic Baseball and Billions, a detailed examination of the American baseball industry from the eyes of an economist, a book which I reference to this day.

Although it is the world's game, soccer does not have the appeal in the United States that it has in the rest of the world. Baseball, on the other hand, has limited appeal outside the US. Yet soccer and baseball share common origins, and developed roughly parallel to one another in time. A comparative study of the economic factors shaping the two sports has been begging for consideration for some time. Such an analysis has just been delivered by Zimbalist and one of Europe's top sports economists, Stefan Szymanski, in the new book National Pastime: How Americans Play Baseball and the Rest of the World Plays Soccer.

Szymanski and Zimbalist describe the origins of the sports' organizational structure and their development to the present day. As one has come to expect from the authors, the book is logically sound, and loaded with facts and illustrative anecdotes. The prose steers clear of professional jargon while remaining true to the authors' training in economics.

The result is an entertaining and informative book. It will appeal to lay readers interested in the common origins and different growth paths of soccer and baseball, as well as professionally-trained economists interested in the economics of sports.

Both sports owe their organizational origins to England's 18^{th} century aristocratic pastime: cricket. As so-often happens with athletic competitions, gambling interests began oozing into the game, undermining the credibility of the play on the field. But when the British public school system adopted the game and its students began to play it, cricket was given new life and it spread to a wider audience. As its popularity spread, commercial interests began entering the game.

But in 19^{th} century England, keeping the proper social order was of utmost importance. This was reflected in the distinction between amateurs and professionals in sports. Men who participated in cricket matches as a leisure activity without compensation were termed "gentlemen" while those who participated for compensation were merely called "players." Even though differences in social status were important, cricketers from different social classes competed against one another in matches. Cricketers from the middle classes competed in order to move up the social ladder, competitions that eventually became the promotion and relegation systems used worldwide in soccer leagues.

Even though cricket reached a wide audience, the length of a cricket match could last several days, limiting its commercial appeal. In part because of this limited appeal, soccer and baseball - with games completed in a couple of hours - developed and spread.

Like cricket, both baseball and soccer began as a gentleman's game and, as their popularity spread, commercial interests entered the picture. But due to differences between American and English social and business cultures, they grew and spread in different directions. In America, generating profits through private enterprise was the primary driving force behind the growth of baseball while spreading the popularity of and participation in soccer was a primary driving force behind soccer's growth.

In the late 1800's, the economic and political reach of England dwarfed that of the United States. Soccer's global appeal today derives directly from England's global reach in the 19th century. Two primary factors limited baseball's global appeal and stunted its worldwide growth. First, American social and business culture had a relatively limited global reach which kept international interest in the game limited. Second, baseball was developed as a business, and its teams were developed as profit-seeking enterprises with territorial monopolies protected through league rules. Although some of the early pioneers of baseball attempted to promote the game internationally, their main focus was not on encouraging participation in the sport but instead was on capturing an international audience for the game. Except for Japan and Latin America, this tactic was largely unsuccessful, and baseball did not spread internationally.

Since both sports were originally developed as gentlemen's games, the labor markets of both sports are rooted in leisured amateurism. As each sport spread in popularity, the governing bodies and team owners in each sport sought to place limits on player mobility and compensation. These limits were manifested in baseball's reserve system and in soccer's retain and transfer system as well as its maximum wage rule.

In baseball's reserve system, each team put its players "on reserve:" no other club can negotiate with those players. When a player (and his contract) was transferred to another team, the team either received another player or a cash transfer in exchange. Of course, generating profits, not preserving leisured amateurism and the gentlemanly pursuit of the game, was the goal of the reserve system. Owing in part to baseball's antitrust exemption given, oddly, because baseball competitions were not considered to constitute interstate commerce, the reserve system existed until the mid 1960's when it was modified first by the player draft and then in the mid 1970's by free agency.

Soccer's labor market developed along a similar path. English soccer officials instituted a maximum wage that effectively limited player compensation, making it more comparable to general workers than one might expect when viewed from today's eyes. To restrict the mobility of players, soccer leagues developed a "retain and transfer" system in which each club listed a set of players it would retain and a set of players it would transfer for a fee. But unlike baseball, when a soccer player transferred between teams, his contract was not transferred and he was free to negotiate a new contract with his new team. To buttress the retain and transfer system and the maximum wage rule, contractual agreements arose between the English football league and leagues in other nations to further restrict the movement of players between teams.

Before the emergence of the European Union, labor mobility between European countries was restricted. But as the EU developed and the sport emerged in Asia and Africa, legal and contractual barriers to player mobility diminished. This brought more international competition for soccer players, thereby putting upward pressure on player wages.

The authors then discuss the motivations of team owners and the influence of broadcast revenues in each sport. Since baseball teams were organized as businesses, the rules of business preside. Although baseball owners surely obtain non-monetary satisfaction from the game, it is generally accepted that they operate their teams to either generate profits per se, or to enhance the overall value of their asset portfolio.

Soccer clubs on the other hand, do not tend to be owned for profit, but rather to provide satisfaction to the owners. This owes, in large part, to the market for soccer clubs and due to the organization of leagues. In baseball, if a person wants to start a new team, she must first get the approval of the current team owners and must pay an expansion fee. Soccer clubs, on the other hand, can be organized relatively easily and do not require the approval of current team owners. In addition, the promotion and relegation system makes it possible for new teams in lower divisions to move to different divisions. Consequently, the competitive pressures on owners are intense and the ability of most clubs to turn a profit is restricted. Therefore, it is generally accepted that soccer owners run their clubs to maximize wins, subject to the constraints imposed by commercial and on-the-field competition.

Even though owners of teams in the different sports may have different objectives, revenues are obviously critical to clubs in both sports. First the telegraph, then radio and television, became important means by which games of each sport could reach a wider audience. However, since gate revenue was the primary income source of baseball and soccer teams, the broadcasting of games was initially frowned upon for fear that such broadcasts would cut into attendance and team revenue. But leaders in both sports have since realized that while broadcasts can cut into attendance, they also allow each sport to reach a much more sizable audience.

One of the primary differences between broadcasting in America and in the rest of the world is that American broadcasting is primarily a private concern. In most other countries, broadcasting has been, until recently, a public concern. In the 1980's with the general wave of privatization and market orientation that swept the world and with the emergence of the European Union, more commercialism began to creep into the broadcasting of soccer. This commercialism and its accompanying increase in revenues has put upward pressure on soccer salaries.

In their next-to-last chapter, the authors tackle the complex questions of uncertainty of outcome and competitive balance within a league. As mentioned above, one of the strengths of the present book is its readability, in part due to the authors having mostly steered clear of economic and statistical jargon. However, this is not the case in this chapter.

One of the appeals of sports is that no-one knows how games and seasons will play out – there is always some sort of uncertainty present. The authors note two types of uncertainty associated with sporting outcomes – within-season uncertainty and between-season uncertainty. They do not specifically define the types of uncertainty, but instead intuitively describe what is meant by these terms. Between-season uncertainty is relatively straightforward - it generally refers to the spread of championships among various teams over time. If there are a few perpetually dominant teams within a league, there is less uncertainty between seasons and, consequently, there is less competitive balance between seasons within the league.

Within-season uncertainty is a bit more complicated. As the authors use the term, it generally refers to the closeness of play between teams throughout a season as generally evidenced by the spread of team winning percentages. If team winning percentages exhibit little spread, then there is more uncertainty within the season and competition within the season is more balanced.

Researchers in the economics of sport have identified another type of uncertainty of outcome that occurs throughout a give season – the uncertainty of outcome associated with individual matches. Arguably, the two types of uncertainty of outcome owe their roots to the uncertainty of outcome of individual matches.

The authors correctly note that one of the allures of sports is that no-one knows how particular competitions will play out. When there is little doubt about the outcome of a match, interest in that match is muted. If there is little doubt about the results of individual matches over a season, there is less uncertainty regarding how the season will play out and interest in the season will be muted. When there is little doubt about matches between seasons, season-to-season interest in the sport will be muted. Indeed, when gambling interests infiltrate a league (as happened with cricket), they will seek to fix the outcome of individual matches. If fans become aware of the fixing, the uncertainty of match outcome (as well as the game's credibility) will be damaged, and fan interest will wane.

However, in baseball, teams play a 162 game schedule, so the importance of individual match-ups is of relatively little importance. In soccer, the number of overall matches is much less than in baseball (roughly 30 to 40 matches per season), so the uncertainty of match outcome is relatively more important. While the omission of a discussion on the uncertainty of match outcomes is not too troubling in the context of the sports being compared, its acknowledgement would make the chapter more complete.

In the discussion on baseball, the authors use the term "standard deviation of the winning percent" to make their case about within-season competitive balance. This term is very familiar to people who perform research in sports economics and while the authors provide an intuitive working definition of the term, its use seems esoteric compared to the rest of the book. Moreover, the authors do not use this statistic when discussing within-season uncertainty in soccer.

These criticisms aside, baseball has historically been the most balanced of the two sports in terms of between-season uncertainty, especially during the period between the 1960's and the 1990's. However, with the expiration of baseball's national television contract in the early 1990's, locally-generated revenues became more important to teams. Teams such as the New York Yankees and the Atlanta Braves began to dominate their divisions and, the authors argue, between-season competitive balance decreased. Certainly, these teams have dominated their divisions and the Yankees have won several World Series, but teams such as the Anaheim Angels, Florida Marlins, and Arizona Diamondbacks have won also won World Series. Moreover, the introduction of the wild card format in baseball allows more teams into the playoffs. So it is not clear that between-season competitive balance decreased during the last two-thirds of the 1990's.

In terms of within-season uncertainty, although its various leagues have very few of the competitive balance measures seen in baseball, soccer has been the more balanced sport. There are no player drafts, there are few gate sharing agreements, and sharing of national media revenues are imbalanced. Even without these remedies, the promotion and relegation system creates the necessary incentives for all teams to continue competing until the final days of the season. If a team gets relegated, it will lose enormous sums of revenues. To prevent relegation, poor teams continue to compete even when a championship is far out of reach. On the other hand, if a team gets promoted, it stands to gain enormous revenues. To enhance promotion possibilities, the top teams continue to compete even when they have championship prospects well in hand. Consequently, one does not generally see the win-loss spread between teams in soccer leagues that one may see in baseball.

In their last chapter, the authors examine current and future issues facing each sport. In baseball, the authors discuss some of the implications of the most-recent collective bargaining agreement signed in 2002, specifically the implications of the revenue sharing agreement. They argue that this system creates perverse incentives because it penalizes success. According to the agreement, each team is required to pay 34 percent of its revenues into a central pool which is then divided equally among teams. If a team invests in playing talent that leads to an improvement in its winning percentage and enhances its revenues, it must pay 34 percent of the increase into the central pool. However, regardless of how a given team performs, it receives an equal share of the central pool.

The players' union and the league negotiators seem to have realized that there may not be adequate incentive for teams to spend shared revenue on talent acquisition. Consequently, there is a stipulation included in the collective bargaining agreement that is designed to provide an incentive for teams to invest shared revenues in their teams. The agreement essentially states that any team which does not use shared revenues to improve its playing talent must answer to the commissioner. Whether this is a credible threat and an efficient use of the commissioner's office is debatable. In any case, the authors argue that the net incentive provided by the revenue sharing system penalizes success instead of rewarding it. If the revenue sharing system does succeed at improving league balance, it will most likely do so from the top down instead of improving teams from the bottom up, lessening the overall quality of play.

The authors conclude the book by describing what they view as the biggest challenge to each sport. They then provide prescriptions to help cure what ails each sport. Interestingly, the authors suggest that each sport can take something from the other to help it face its challenges. To help soccer face its challenge, the authors suggest that soccer officials take some qualities of the business model used in baseball. Baseball leaders can face their challenge by encouraging a more long-term approach in encouraging people to be fans of the game.

In soccer, the biggest threat is the financial crisis many teams face, a crisis that has arisen for four reasons. First, they argue that the strict adherence by league governing bodies to separations between national leagues blunts the incentives of teams to develop new competitions which would enhance their revenue streams.

Second, the authors point to the incentives generated by the promotion and relegation systems. The revenue differential between divisions provides an incentive to teams to continue competing although a championship may already have been decided. But this system also gives teams an incentive to invest highly in playing talent. Coupled with the increased mobility of players in the international soccer labor market, the system has led to increased team payrolls. Third, the authors point to the general lack of a profit motive in soccer leagues which makes owning a soccer team a relatively poor investment.

Fourth, the authors point to the historical reliance on public support available to teams in times of financial crisis. Because of baseball's development within the sphere of private enterprise, the connection between baseball and politics has been less pronounced than in soccer. In baseball the discipline of the market generally rules the school, and while substantial government assistance for new stadium construction was forthcoming in the latter half of the 20th century, financial bailouts were not.

Soccer, however, is a different matter. In many countries, the national soccer team is a symbol of the country and there is often much nationalistic fervor when a country's national team performs well in international competition. A similar thing can be said of local soccer teams competing in local rivalries or on the national scene. While this is also true in baseball, the symbolic connection between a team and its city/nation is stronger in soccer than it is in baseball. In terms of the business of running a soccer team, if a team found itself in financial straits, its owner could often count on government assistance to keep it afloat.

In the European Union, however, public assistance has been less-readily available for many soccer clubs. Yet the competitive pressures of promotion and relegation as well as the international labor market remain. Public assistance has provided a crutch to financially imprudent team owners. When crutches are removed, it takes awhile for the person to learn how to live without them. Team owners must learn to make future business decisions while internalizing more of the risks associated with remaining competitive.

To help cure what ails them, the authors suggest that soccer officials introduce more of a sport-as-business model into the game. In addition, they argue that the development of major leagues in soccer can improve their cash flow by allowing for more competition between the biggest teams while still maintaining local rivalries. Overall, encouraging more private entrepreneurship can help soccer leagues avert the financial crisis facing many teams. However, because of the distaste many soccer fans have for commercial interests and because of the deep roots of the current organization of soccer leagues, this medicine, while useful, will likely be hard to take.

Baseball's crisis is more long term. With the 2002 collective bargaining agreement, the authors argue that the short term financial outlook is positive. But dark clouds exist on the horizon, primarily due to the relative disinterest in international baseball and the decreased interest and participation in baseball by American youths.

As noted above, when the pioneers of baseball tried to take the game internationally, their main focus was on building an immediate audience instead of encouraging participation in the sport (and thereby building a long-term audience). In America, baseball had very little competition in the sporting world, owing in part to its antitrust exemption. Today, however, youngsters have a wide variety of choices of sports to play and, unfortunately for baseball, kids are more likely to follow and play basketball, football, and soccer than baseball. While attendance levels at baseball games are currently strong, they are not likely to continue to be as strong when today's youngsters become tomorrow's adults. Thus, while the present financial situation is positive, the growth prospects for baseball look bleak.

The problem in baseball, which the authors acknowledge, is twofold. First, baseball's antitrust exemption has somewhat insulated team owners from focusing on the long term. Second, its governing body is comprised of a collection of league owners with substantial investments in their team, a collection which changes over time, and a commissioner who is an employee of the owners. Consequently, there has been little incentive to focus on the long term and little of the strength and adhesiveness in league leadership necessary to develop and implement long-term initiatives.

For example, the authors argue that allowing American major league players to compete in Olympic baseball would help increase international interest in the game. In addition, the authors argue that building a stronger bridge between professional baseball and amateur baseball (for example, by investing in Little League baseball) would increase domestic interest among youngsters. Implementing these suggestions, however, necessitates short-term sacrifices for long-term gains that major league officials have not been willing to accept.

Szymanski and Zimbalist have written a readable and well-researched book that touches on most of the important topics in the economics of sports teams and leagues: the demand for the sport; the labor market in the sport; and the relationship between sports teams (and leagues) and government. One topic that did not receive much attention was a comparison of stadium investments within the two sports. The authors argue that baseball teams' local monopoly positions have allowed many of them to receive generous subsidies for stadiums. They also argue that the lack of such monopoly positions for soccer clubs (in addition to the promotion and relegation system) has led to an underinvestment in soccer stadiums which may have led to fatal incidents at soccer stadiums. Unfortunately, the authors do not provide much of an examination of the stadium issues in the two sports. Given the importance of this issue especially in the United States, such an analysis would be very interesting and would have fit in very well with the rest of the book.

The criticisms I have of this book are relatively minor, as one might expect about a book written by two of the best sports economists in the business. Each chapter is organized so that there are separate sections on baseball and soccer. While the individual sections do not break much ground by themselves, having a discussion of one sport followed by a discussion of the other sport improves the overall flow of the book. If the reader is a fan of either or both sports (or sports in general) and is interested in the historical parallels and connections between them, this book is a solid source of information and a recommended addition to the fan's library.