Pitfalls to Measuring Competitive Balance With Gini Coefficients

Joshua Utt and Rodney Fort

Journal of Sports Economics 2002; 3; 367

Pitfalls to Measuring Competitive Balance With Gini Coefficients

JOSHUA UTT
RODNEY FORT
Washington State University

League-winning percentage Gini coefficients have seen recent use as measurements of within-season competitive balance in Major League Baseball. The authors demonstrate that the zero-sum nature of league play renders past estimates inappropriate. Adjusted for league play, Gini coefficients reveal a much larger competitive balance problem than shown in previous estimates. However, additional complexities involving unbalanced schedules, interdivisional play, and now interleague play must be overcome before winning percentage Gini coefficients can give precise estimates of competitive balance. The authors suggest using the traditional measures of winning percentage standard deviations and their idealized values to analyze within-season competitive balance over time until these issues are overcome.

Competitive balance is an important element in the evaluation of sports league outcomes. Traditionally, the standard deviation of winning percentage is compared to its idealized value to measure the within-season inequality of league outcomes over time (Quirk & Fort, 1992, chap. 7). Just as typically, Gini coefficients (properly normalized to capture variation in championship opportunities over time) are used to measure inequality in league playoff outcomes (again, Quirk & Fort, 1992, chap. 7). Lately, in a curious turnabout, within-season inequality in Major League Baseball (MLB) has been analyzed using a winning percentage Gini coefficient (Schmidt, 2001; Schmidt & Berri, 2001). In this article, we demonstrate pitfalls to using winning percentage of Gini coefficients to measure competitive balance.

There are two issues. The first is easy to see using Figure 1. Complete equality of winning is the 45-degree line. The curve in the figure is the actual calculated amount of inequality by Schmidt and Berri (2001) for the National League in 1985 (the calculation is detailed shortly). Then, in the standard way, complete inequality is taken to be the lower right triangle so that the Gini index is the ratio of the areas $\frac{A}{A+B}$. The Gini index lies between 0 and 1, and the competitive balance decreases as the Gini index rises.

JOURNAL OF SPORTS ECONOMICS, Vol. 3 No. 4, November 2002 367–373
DOI: 10.1177/152700202237502

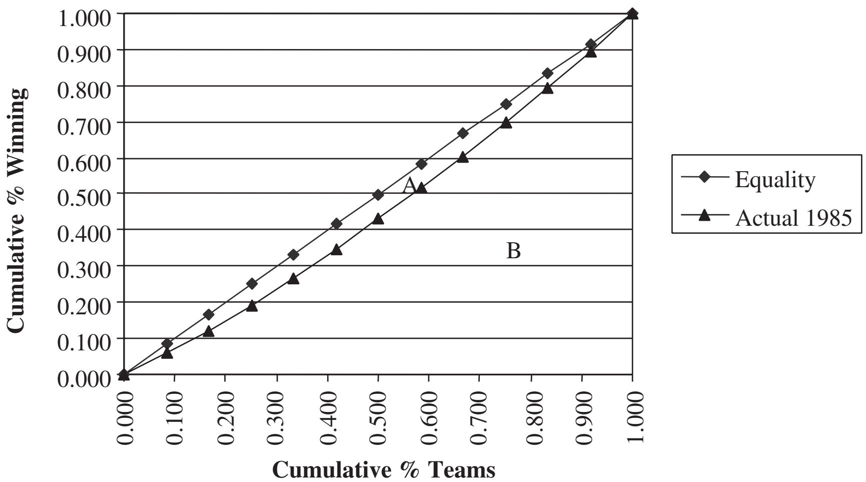

Figure 1: Gini Coefficient Derivation in Past Works

But this formulation used by other researchers incorrectly specifies the idea of a completely unequal distribution of winning percentage for league play where winning is zero sum. Indeed, the sum of the winning percentages across teams in a league equals half the number of teams. At most, one team can win all of the games it plays, a fraction of the total number of games played in the league. The most unequal outcome cannot have one team winning all the games played in the entire league so that the denominator must be smaller than A + B and analyses to date understate the level of inequality. We show that an approximation of the Gini coefficient that adjusts for league play reveals a much larger competitive balance problem than shown by the standard Gini calculation used in the literature to date. And the adjusted Gini coefficient generates some important qualitative differences for particular interesting years.

Which brings us to the second issue. Even the adjustment we use to specify a completely unequal distribution of winning percent ignores a host of other complexities, such as unbalanced schedules (teams do not play the same number of games against all opponents), league expansion, and interleague play and presents additional challenges. We caution against use of Gini coefficients in the analysis of competitive balance until these computational problems are solved.

ANALYSES TO DATE

In the works cited above that use Gini coefficients to analyze competitive balance, the Gini coefficient is calculated in the standard way using the formula in

Lambert (1993). The result is Gini coefficients that are almost always less than 0.100. Examples are 0.094 in 1985 and 0.077 in 1986. (Using Lambert's equation, we calculated the 1985 Gini at 0.094497, which appears to be very close to the value of about 0.095 shown in the chart of results in Schmidt and Berri [2001]. Lambert's equation gives the 1986 Gini at 0.077497, which also corresponds to the Schmidt and Berri charted value. Because we also were able to reproduce all of their tabled Gini coefficients for the 1990s, we assume that we have replicated their results.)

But as mentioned before, the Gini coefficient calculated in the standard way ignores that one team cannot possibly win all of the league games. Fort and Quirk (1997), in an analysis of winning percentages in college sports conferences, used the following approach to handle the problem. To normalize for league play constraints, they invented a hypothetical "most unequal distribution" of winning percentage in which there is an undefeated team, a team winning all but one game, and so on down the line to a team winning no games. But even this was only an approximation because the schedules they examined were not balanced (i.e., each team did not play every other team an equal number of times). And there are even more complications in the case of MLB that we address later.

What one discovers is that this most unequal distribution of winning percentages generates a lower bound like the "inequality" line shown in Figure 2. It is easy to see why Gini coefficients are too small in past analyses. Conceptually, Gini coefficients in past works are equal to $\frac{A}{A+B}$. But using the most unequal distribution lower bound that accounts for league play would yield a Gini equal to $\frac{A}{A+B}$. Because the area $B' < B$, adjusting for league play produces a Gini coefficient that must be larger than the values calculated in works to date.

AN INCOMPLETE ADJUSTMENT FOR ZERO-SUM WINNING PERCENTAGES

We adopt the Fort and Quirk (1997) idea and generate "adjusted" Gini coefficients and compare them to both the version of the winning percentage Gini coefficients used by other authors and the usual standard deviations of winning percentage used by most analysts. For example, for the earliest years in our comparison, both the American League and National League had 12 teams all playing 162 game unbalanced schedules. We used a hypothetical 12-team, 165-game-balanced, division-only schedule (as close to 162 games as possible for such a constraint) to derive the most unequal distribution of winning percentages for these years. This assumes that each team plays the other 15 times. The most unequal winning percentage outcome is calculated in which one team wins all 165 of its games, the second team wins all its games except the ones it lost to the undefeated team for a total of 150 wins, and so on down to the hypothetically truly hapless 0 and 165 team. For the years we will compare, over the decade of the 1990s, the number of teams in each league was not constant so that other hypothetical constructs had to be used to get the most unequal distribution bound as league sizes changed. Although we do

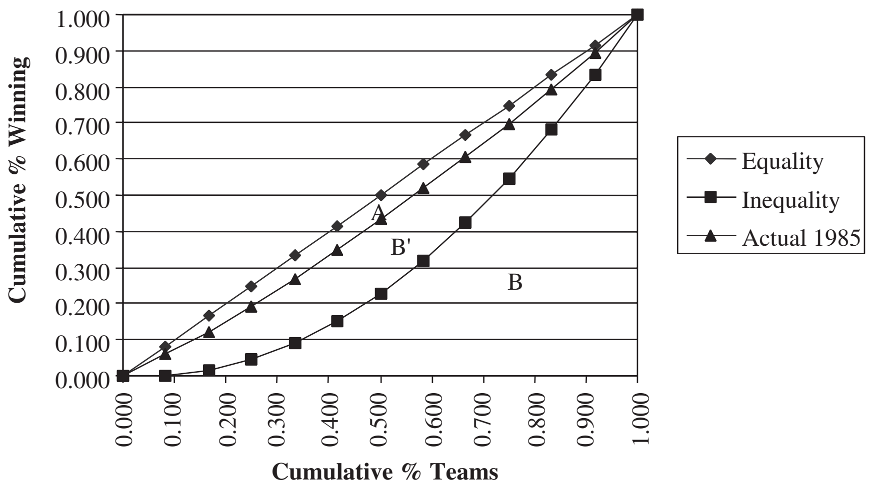

Figure 2: Gini Coefficient Adjusted for League Play

not wed ourselves to this type of adjusted Gini coefficient for a host of reasons discussed shortly, it is useful to compare it to past efforts using Gini coefficients that do not adjust for league play.

Table 1 shows the adjusted Gini coefficients, the Gini coefficients from Schmidt and Berri (2001), and the standard deviations of winning percentage all for the 1990s plus the first 2 years of the new decade. Table 2 is used for comparisons, showing annual percentage changes for each of the three measurements. Schmidt and Berri offered the following conclusions from their analysis:

> Although the latter half of the decade saw an increase in the Gini coefficient and therefore a decrease in competitive balance relative to the minimum values achieved in the early 1990s, major league baseball appears to have been more competitive in the last decade of the century than the typical result achieved over the past 100 years. . . . Despite the domination by the Yankees and the Braves, major league baseball was more competitive in the 1990s than in any other period in history. (p. 149)

Actually, all measures show an average decrease in competitive balance both prior to 1995 and after for the National League, whereas the American League held reasonably steady after 1995, even though the level in each league averaged over the 1990s is lower than in previous decades. And that is precisely what Fort (2001) found, using traditional measures of competitive balance (he also found that playoffs always have been dominated by larger revenue market teams, as they are now, but that the set of larger revenue market teams had simply changed over time). And there are other interesting points of comparison among the three measures of competitive balance.

TABLE 1: Competitive Imbalance Measurement Comparisons in Major League Baseball, 1990s

	S & B Gini		*Adjusted Gini*		SD *Winning Percentage*	
Year	*AL*	*NL*	*AL*	*NL*	*AL*	*NL*
1990	0.059	0.062	0.117	0.123	0.057	0.057
1991	0.059	0.065	0.119	0.131	0.061	0.061
1992	0.068	0.072	0.136	0.143	0.063	0.066
1993	0.060	0.103	0.119	0.206	0.055	0.093
1994	0.071	0.075	0.047	0.049	0.029	0.031
1995	0.088	0.070	0.177	0.128	0.083	0.060
1996	0.071	0.061	0.144	0.123	0.069	0.056
1997	0.067	0.063	0.135	0.127	0.062	0.059
1998	0.083	0.097	0.168	0.193	0.081	0.088
1999	0.084	0.085	0.168	0.169	0.076	0.079
Average	0.071	0.075	0.133	0.139	0.064	0.065
2000	—	—	0.118	0.155	0.054	0.069
2001	—	—	0.209	0.142	0.098	0.065

SOURCE: The S & B Gini coefficients are from Schmidt and Berri (2001, Table 2, p. 148). The adjusted Gini coefficients and standard deviations of winning percentages are our calculations. AL = American League; NL = National League.

First, the Gini coefficients adjusted for league play suggest that the decrease in the level of competitive balance, except for the strike year of 1994, was always about twice the size found in recent works that use the standard Gini coefficient. As pointed out by Quirk and Fort (1992), play always has been far from competitively balanced in baseball.

Second, the standard deviations of winning percentages reveal that the increase over the 1990s was quite probably three times larger than Schmidt and Berri (2001) found for the American League and twice as large as they found for the National League. And the Gini coefficients adjusted for league play show an even larger increase in each league over the 1990s. Even though the decade averages in Table 1 are smaller than for previous decades, Table 2 shows a decrease in competitive balance over that decade that the standard Gini coefficient minimizes relative to other measures.

Finally, the Gini coefficients used in past works appear to have dramatically missed the mark during the interesting strike episode of 1994 to 1995. Indeed, in the American League, past works show that competitive balance declined from 1993 to 1994, whereas the other two measures show very large competitive balance improvements. It is easy to explain why competitive balance would fall during a shortened year, but no explanation of an increase in the American League during the 1994 season occurs to us. Other sign conflicts between the standard Gini coefficients and the other two measures of competitive balance occur for the American League in 1998 to 1999 and in the National League for 1994 to 1995.

TABLE 2: Percentage Change Comparisons, 1990s

Year	*S & B*	*AL Adjusted*	SD	*S & B*	*NL Adjusted*	SD
1990-1991	0.0	1.7	7.0	4.8	6.5	7.0
1991-1992	15.3	14.3	3.3	10.8	9.2	8.2
1992-1993	–11.8	–12.5	–12.7	43.1	44.1	40.9
1993-1994	18.3	–60.5	–47.3	–27.2	–76.2	–66.7
1994-1995	23.9	276.6	186.2	–6.7	161.2	93.5
1995-1996	–19.3	–18.6	–16.9	–12.9	–3.9	–6.7
1996-1997	–5.6	–6.2	–10.1	3.3	3.3	5.4
1997-1998	23.9	24.4	30.6	54.0	52.0	49.2
1998-1999	1.2	0.0	–6.2	–12.4	–12.4	–10.2
Average to 1995	9.2	43.9	27.3	5.0	28.9	16.6
Average after 1995	0.0	–0.1	–0.6	8.0	9.7	9.4
Decade average	5.1	24.3	14.9	6.3	20.4	13.4
1999-2000	—	–11.3	–15.6	—	11.5	6.2
2000-2001	—	77.1	81.5	—	–8.4	–5.8

NOTE: AL = American League; NL = National League.

TO GINI OR NOT TO GINI?

Failing to adjust Gini coefficients for league play dramatically overstates the level of competitive balance in baseball and provides some puzzling qualitative results. But there are additional issues that arise beyond the adjustment used by Fort and Quirk (1997) in their study of college football to capture the zero-sum nature of league play. In addition to division play, the winning percentage used to crown MLB division champions includes unbalanced schedules, interdivisional play, and interleague games. Normalization on all these dimensions would be needed to implement successfully a Gini coefficient comparison of competitive balance.

We could not develop an approximate Gini coefficient to adjust for all of these complications simultaneously. Without such a measure, our practice into the future will be to use the standard deviation of winning percentage and its idealized value in making competitive balance comparisons over time, as in Fort (2001).

CONCLUSIONS

Winning percentage Gini coefficients have recently been used to analyze competitive balance over time both within season and for championship outcomes. Of course, adjusted for championship opportunities, Gini coefficients work just fine for the analysis of championship balance. But there are two problems with this approach to within-season competitive balance. First, under league play, one team cannot win all of the league's games, only all of its own games. We use an adjustment following Fort and Quirk (1997) to cover the zero-sum nature of league play and show that past works overstate the level of competitive balance in MLB and miss the mark qualitatively for some years in the 1990s. Competitive balance does behave in predictable ways relative to real-world occurrences such as player strikes, and although far from balanced in absolute terms, competitive balance in MLB has been improving over time.

But there remains a host of complexities even after adjusting for league play-unbalanced schedules, interdivisional play, and interleague play for within-season competitive balance. Normalizing Gini coefficients for these factors was beyond our ability. And until a remedy for these complications is found, we will stick with the tried and true standard deviation of winning percentages (and their idealized values) for within-season competitive balance analysis of winning percentages.

REFERENCES

Fort, R. (2001). Revenue disparity and competitive balance in Major League Baseball. In *Baseball's revenue gap: Pennant for sale? Hearing before the Subcommittee on Antitrust, Business Rights, and Competition of the Committee on the Judiciary*, United States Senate, 106th Congress, 2nd Session, November 21, 2000. pp. 42-52.

Fort, R., & Quirk, J. (1997). Introducing a competitive economic environment into professional sports. In W. Hendricks (Ed.), *Advances in the economics of sports* (Vol. 2). Greenwich, CT: JAI Press.

Lambert, P. J. (1993). *The distribution and redistribution of income: A mathematical analysis*. Manchester, UK: Manchester University Press.

Quirk, J., & Fort, R. D. (1992). *Pay dirt: The business of professional team sports*. Princeton, NJ: Princeton University Press.

Schmidt, M. B. (2001). Competition in Major League Baseball: The impact of expansion. *Applied Economic Letters*, *8*, 21-26.

Schmidt, M. B., & Berri, D. J. (2001). Competitive balance and attendance: The case of Major League Baseball. *Journal of Sports Economics*, *2*(2), 145-167.

Joshua Utt is a graduate student in the Department of Economics at Washington State University.

Rodney Fort is a professor in the Department of Economics at Washington State University.