Measurement of Competitive Balance in Conference and Divisional Tournament Design[#]

Liam J. A. Lenten[*]

School of Economics and Finance

La Trobe University

Abstract

The conference and divisional system has long been a staple part of tournament design in the major pro-sports leagues of North America. This popular but highly-rigid system determines on how many occasions all bilateral pairings of teams play each other during the season. Despite the virtues of this system, it necessitates removing the biases it generates in the set of win-ratios from the regular season standings prior to calculating within-season measures of competitive balance. This paper applies a modified version of a recent model, an extension that is generalizable to any unbalanced schedule design in professional sports leagues worldwide, to correct for this inherent bias for the NFL over the seasons 2002-2009, the results of which suggest the NFL is even more competitively balanced than thought previously.

JEL Classification Number: C14, L83

Keywords: Competitive Balance, Measurement Methods

[#]Research assistance on this paper was provided by Lan Nguyen. An earlier version of this paper was presented at a seminar at Temple University, Philadelphia, 3 September 2010. The author would like to thank the participants of the seminar for their comments and suggestions, especially Michael Bognanno. The author would also like to thank Damien Eldridge, Stefan Késenne and Markus Lang for their input.

[*]Contact details: School of Economics and Finance, La Trobe University, Victoria, 3086, AUSTRALIA. Tel: + 61 3 9479 3607, Fax: + 61 3 9479 1654. E-mail: l.lenten@latrobe.edu.au

1. Introduction

The conference and divisional system has long been considered the default tournament design in the National Basketball Association (NBA), National Football League (NFL), Major League Baseball (MLB) and National Hockey League (NHL), referred to henceforth as the 'major leagues'. The practice is also standard in other less popular North American leagues, such as Major League Soccer (MLS) and the Canadian Football League (CFL). Fundamentally, the conference and divisional system is a manifestation of 'unbalanced schedules', whereby not all bilateral pairings of teams play each other an equal number of times.

Other forms of unbalanced schedules exist elsewhere, like where the unbalancedness is largely randomized, for example, the Australian Football League (AFL). Another alternative is 'power-matching' like in Scottish Premier League (SPL) soccer, where teams play a fourth time *only* teams that finished in the same half of the standings following three full round-robins. Other forms are even more radical – such as the Belgian League format from 2010, involving many playoff combinations after two full round-robins, producing a season whereby all teams do not play the same number of games (an *uneven* schedule).[1] Despite these isolated examples, unbalanced schedules are rare in European soccer – see Cain and Haddock (2005) for an historical account of Trans-Atlantic sports industry differences, including tournament design.

There are numerous motivations in favor of the system. Logistically, it can be used to co-ordinate and minimize travel distances for teams. Economically, the system fosters local rivalries and arguably increases aggregate demand, measured typically by

[1] The year refers generally to the calendar year in which the regular season concludes.

television ratings or attendances. Paul (2003) and Paul, Weinbach and Melvin (2004) demonstrate this point, although the MLB results of Butler (2002) suggest that fans are keen to see their team against other teams that they have not played against in past seasons. Notwithstanding these merits, it is far from perfect – many fans do not like seeing their team miss out on a playoff spot to a team in another division with fewer wins. There is also the perennial dilemma of teams having already clinched a conference seeding or playoff spot losing games they should otherwise win due to purposely resting key players in the last few regular season games, indirectly influencing the playoff destiny of others (see Blodget, 2010). More problematically, unbalanced schedules cause biases in final regular season team standings, since they do not control for the strength of schedule, rendering invalid standard measures that sports economists use to characterize competitive balance in sports leagues (see Lenten, 2008 and 2010, for the SPL and AFL, respectively).

This study applies a modified adjustment procedure to identify the extent to which strength of schedule affects competitive balance measures in conference and divisional systems, based on a simple matrix algebra method. This issue is analogous to the ongoing industrial organization debate on the appropriateness of standard quantitative measures of market concentration to accurately reflect true competitive structure. The focus is on NFL data specifically, and raises a number of empirical questions that pertain to the nuances of the conference and divisional system as an unbalanced schedule. Firstly, this is a critical ongoing policy issue for major league tournament design – the NHL altered the composition of its schedule in advance of the 2010 season, and as the remaining leagues consider expansion from 30 to 32 teams. More generally, it is difficult to speculate for major league data whether adjusting the league standings will affect reported competitive balance measures significantly. We are also interested in whether the team-quality equalization aspect of the NFL fixture, unique to the major leagues, is successful in making the tournament more even. There is also the possibility that some teams (or even divisions) fare better or worse than others in the fixture systematically because of the fixed allocation of franchises to conferences and divisions over seasons.

The remainder of the paper proceeds according to the following structure: section 2 summarizes the relevant literature, while section 3 illustrates the strength of schedule adjustment methodology. Section 4 outlines alternative unbalanced schedule formats used in the various major leagues. The results are presented and the findings discussed extensively in section 5, followed by a brief conclusion in section 6.

2. Literature Review

The universal application of the conference and divisional system across the major leagues has made it the attention of substantial previous research, though not quite to the extent imagined. Given the largely North American context of the problem, much of this work centers on team ratings in unbalanced schedules – Fainmesser, Fershtman and Gandal (2009) is a recent example with respect to the highly idiosyncratic NCAA College Football. Lenten (2010) lists various studies utilizing a range of statistical methods for this purpose, as well as other studies that relate to the more general economic implications of tournament design for the professional sports industry. Of underlying relevance to sports economists and statisticians is the competitive balance paradigm, which extends back to Rottenberg's (1956) baseball players' labor market study. Competitive balance is defined as the degree of parity in sports leagues, and helps describe the ‘uncertainty of outcome’ hypothesis – a cornerstone of sports economics. Since demand for sport is expected to be greater when the league is competitively balanced, this represents a rare example in the economics discipline in which uncertainty is associated with higher utility.

Consequently, in sports economics there is a large volume of literature comparing leagues (major and other), such as Quirk and Fort (1992), and Vrooman (1995 and 2009), to assess relative effectiveness of labor market and revenue-sharing policies used by leagues to maintain/improve competitive balance. Single-league time-series studies, such as Lee and Fort (2005) and Fort and Lee (2007), are also useful for the same reason. Lee (2010) follows a similar direction for the NFL specifically, yielding the conclusion (recently topical again) that the 1993 collective bargaining agreement (CBA) improved competitive balance. Examining whether adjusting for unbalanced schedules affects competitive balance measures is motivated largely by its possible implications for these studies, since they invariably use these common measures for the purposes of testing. Therefore, there is a question of how sensitive their conclusions are to any possible biases in these measures.

On these conclusions, numerous influential modeling studies on policy effectiveness, such as Fort and Quirk (1995) and Vrooman (1995), argue that the Coasian invariance principle holds in leagues comprised of profit-maximizing teams. That is, free agency competitive balance is invariant to the introduction or removal of labor market and revenue-sharing policies. However, other notable studies present a different view – Depken (1999) finds some (albeit mixed) empirical evidence that the removal of the reserve clause in MLB (freeing up the players’ labor market), resulted in a deterioration in competitive balance. Meanwhile, it is well-known that these interventionist policies tend to be effective in leagues of win-maximizing teams. Since the major leagues do not tend to be modeled in this way, several factors (other than team objectives) will collectively determine the (in)effectiveness of such policies. In the case of revenue sharing, these factors include the: (i) specification of team revenue functions (including or excluding absolute quality of competition); (ii) sharing arrangement (at-the-gate or pooled; equal or performance-related); and (iii) equilibrium condition (fixed-supply Walras or Nash conjectures). On these lines, Késenne (2000) concludes that revenue sharing can improve competitive balance even if teams are profit maximizers, while Szymanski and Késenne (2004) demonstrate that the contrary could occur under very different conditions. Moreover, Dietl, Lang and Werner (2009), show that revenue sharing improves competitive balance in certain types of leagues containing (heterogeneous) teams with mixed objectives.

3. Methodology

We use a modified version of the Lenten (2008) n-team unbalanced schedule league model where team i is scheduled to play some other teams k times and all remaining teams $k+1$ times. This model has the advantage of treating only the unbalancedness of the schedule as endogenous, while other factors typically included in team rating models are exogenous. This allows us to identify purely the effect arising from the policy under investigation. However, the model is inadequate to deal with the flexible range of conference and divisional systems used throughout the various major leagues, a shortcoming we aim to overcome. Therefore, the season structure is relaxed such that teams can play other teams on any number of occasions, by allowing m different values of k (not just k or $k+1$). In this setting, an initial $n \times n$ schedule (symmetric) matrix for season *t* is specified similarly to Jech (1983: p. 247), in which element $x_{i,j}$ denotes the number of times team *i* plays team *j* during the season, as

$$\mathbf{X} = \frac{1}{\sum_{j=1}^{n} x_{i,j}} \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,n} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n,1} & x_{n,2} & \cdots & x_{n,n} \end{bmatrix} \tag{1}$$

where $x_{i,j} \in \{k_1, k_2, ..., k_m\}, \forall i, j$; $k_l \in \mathbb{Z}_+ = \{k_l \in \mathbb{Z} : k_l \geq 0\}, \forall l\{1, ..., m\}$; $x_{i,i} = 0, \forall i$ and $x_{i,j} = x_{j,i}, \forall i \neq j$. One can similarly define another $n \times n$ matrix as being a full set of head-to-head *ex-post* decimalized win-loss records, $\mathbf{W}$, such that:

$$\mathbf{W} = \begin{bmatrix} w_{1,1} & w_{1,2} & \cdots & w_{1,n} \\ w_{2,1} & w_{2,2} & \cdots & w_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n,1} & w_{n,2} & \cdots & w_{n,n} \end{bmatrix} \tag{2}$$

where

$$w_{i,j} \in \left\{0, \frac{1}{k_l}, ..., \frac{k_l - 1}{k_l}, 1\right\}, \forall k_l > 0 \tag{3}$$

and where $w_{i,j} = 1 - w_{j,i}, \forall i \neq j$; and $w_{i,i} = 1/2, \forall i$.[2] For $k_l = 0$, we can assume $w_{i,j} = 1/2$ for now.[3] The interpretation of the diagonal elements is that team *i*, if hypothetically playing against itself (despite being impossible intuitively), is expected to have an equal probability to win or lose. The constraint that all teams play an equal season length is imposed to make formal representation easier, that is:

[2] The set of possible values of $w_{i,j}$ in equation (3) ignores ties – unusual in the NFL since teams have to be level after regulation time, with no further score in overtime. This occurred only twice in the sample (Pittsburgh Steelers v Atlanta Falcons, Week 10, 2002; and Cincinnati Bengals v Philadelphia Eagles, Week 11, 2008). Assuming equal allocation of competition points for a tie (as in the NFL), the set of possible values can instead be generalized as: $w_{i,j} \in \left\{0, \frac{1}{2k_l}, ..., \frac{2k_l - 1}{2k_l}, 1\right\}, \forall k_l > 0$.

[3] This is possible only because we are interested purely in the diagonal elements from the product matrix, and these elements are quantitatively unaffected by the presence of this assumption. Nevertheless, this point is revisited in the latter stages of the appendix.

$\sum_{j=1}^{n} x_{i,j} = \sum_{l=1}^{m} k_l \cdot n_{k_l}, \forall i$ (where n_{k_l} is the number of teams that team *i* plays k_l times), though it is not essential to the validity of the procedure. This also ensures that $\det[\mathbf{X}]$ is fixed for all *i*.

In assessing within-season competitive balance, the usual starting point is constructing a column vector of observed team 'strengths', to resemble a league table (a concept familiar to fans). Strength is measured simply as the win-ratio (games won divided by games played), representing the actual final season record of ordering of all *n* teams in the league from all games played. For simplicity, the associated $n \times 1$ vector, $\mathbf{V}$, can be expressed as

$$\mathbf{V}' = \begin{bmatrix} w_1 & w_2 & \dots & w_n \end{bmatrix} \tag{4}$$

where $w_i = \sum_{j=1}^{n} w_{i,j} \cdot x_{j,i}$. These elements are identical to the diagonal elements of the product of the matrices from (1) and (2). While valid for any ordering of teams, it is easiest to visualize if the ordering is identical to the *ex-post* ranking of teams.

If it were assumed that this is identical to the (unobservable) *ex-ante* ordering of teams, then such an ordering also allows us to define both extremes against which observed competitive balance measures can be deflated. This could be either: (i) a perfectly unbalanced league (closest outcome possible to a monopoly) as one in which $\mathbf{W}$ is an upper triangular matrix where $w_{i,j} = 1, \forall i < j$; or (ii) a perfectly balanced league, whereby $w_{i,j} = 1/2, \forall i, j$. A third possible benchmark is that whereby match outcomes are according to the binomial distribution: $\Pr(i \text{ wins}) = \Pr(j \text{ wins}) = 0.5$, in which case, for any $k_l > 0$ in (3) and ignoring ties, can be generalized as

$$\Pr\left(w_{i,j} = \frac{z}{k_l}\right) = \Pr\left(w_{i,j} = 1 - \frac{z}{k_l}\right) = \binom{k_l}{z} 0.5^z \cdot 0.5^{k_l - z} \tag{5}$$

where $z \in \{1,2,...,k_l\}$ is an index. The distributional properties of the dispersion of the resulting set of w_i in (4) can then be calculated (or if impractical, simulated).

In order to correct the unbalanced schedule bias in the end-of-season standings, a logical way to proceed is calculating implied win-probabilities (using only actual standings) of theoretical unplayed matches required to reset each $x_{i,j}(i \neq j)$ equal to $\max\{x_{i,j}\}$. The simple logit-style rule used by Lenten (2008) produced a solution containing a number of attractive numerical properties, such as

$$\sum_{i=1}^{n} w_{i,t} = \sum_{i=1}^{n} \tilde{w}_{i,t} \qquad \forall t\{1,...,T\} \tag{6}$$

where $\tilde{w}_i$ is the adjusted win-ratio in season *t*, which is to say that the adjustment procedure should not affect the sum of wins across all teams in any given season.

However, a more direct way to account for the win-ratio bias is to compare the strengths of team *i*'s actual full-season schedule to that from a hypothetical schedule in which every game were against an opponent of average strength of all teams (excluding itself). The former, $\mathbf{A}$, is simply a column vector of sum products of the number of games against each opponent and the corresponding win-ratios, hence

$$\mathbf{A} = \sum_{j=1}^{n} x_{i,j} \bullet \mathbf{XV} \tag{7}$$

Meanwhile, the latter, $\mathbf{Y}$ (also a column vector), can similarly be computed as

$$\mathbf{Y}' = \frac{\sum_{j=1}^{n} x_{i,j}}{n-1} \left[\frac{n}{2} - w_1 \quad \frac{n}{2} - w_2 \quad \cdots \quad \frac{n}{2} - w_N \right] \tag{8}$$

Using only these two pieces of information to determine the adjustment means that the distribution of individual head-to-head records do not matter, which is also considered to be a desirable property.[4] In vector form, taking equation (4) as a base, the adjusted set of win-ratios arising from equations (7) and (8) for all *n* teams can be expressed as follows

$$\begin{aligned}\tilde{\mathbf{V}}' &= \begin{bmatrix}\tilde{w}_1 & \tilde{w}_2 & \dots & \tilde{w}_n\end{bmatrix} \\ &= \mathbf{V}' + \left(1 \Big/ \sum_{j=1}^{n} x_{i,j}\right) \bullet \left[\mathbf{A}' - \mathbf{Y}'\right]\end{aligned} \tag{9}$$

where each element in $\tilde{\mathbf{V}}$, for any single team, *i*, can alternatively be calculated according to the following function relating the adjusted win-ratio to the unadjusted win-ratio directly, expressed in scalar form as

$$\tilde{w}_i = w_i + \left[\left(\sum_{l=1}^{m}\sum_{k_l=1}^{n_{k_l}} w_{k_l} \cdot n_{k_l}\right) - \sum_{j=1}^{n} x_{i,j}\left(\frac{n/2 - w_i}{n-1}\right)\right] \Big/ \sum_{j=1}^{n} x_{i,j} \tag{10}$$

for any ordering of inequalities for $k_1, k_2, ..., k_m$. Equation (10) has the additional desirable property, noted by Lenten (2010), that it is invariant to the assumed number of full round-robins, which is crucial for NFL data, since it contains zeros and ones.

4. Tournament Formats

Correction for unbalanced schedules is identical irrespective of the nature of the tournament design; however, the application of the proposed procedure is potentially profoundly different. A quick glance at table 1 demonstrates the similarities and differences between the major league season structures. The middle two columns, which reveal how many times team *i* plays all other teams, are ordered vertically such that $k_1 < k_2 < .. < k_{m-1} < k_m$. The current NFL system was introduced upon expansion

[4] The appendix presents an alternative procedure in which these head-to-head records *do* matter.

to 32 teams (16 in both conferences and 4 in each division). This system sees each team playing teams in their own division twice per season, automatically allocating opponents in 6 out of 16 matches. They also play all teams in one other division within their own conference, plus all teams from one division in the other conference, allocating a further 8 matches. The other two matches in the season are allocated according to a (lagged) power-matching rule – they are against teams in the two divisions in the same conference they would not play otherwise, that finished in the same divisional rank-position in the previous season.

To formalize the earlier notation, this means collectively that under the current system, $n = 32$, $k_1 = 0$ $(n_{k_1} = 18)$, $k_2 = 1$ $(n_{k_2} = 10)$, and $k_3 = 2$ $(n_{k_3} = 3)$. From 2002, the idea was for an 8-year rotational system, in which each team would play every other team at least twice – once home, once away (even inter-conference teams). This full rotational period defines the sample, extending over the regular seasons from 2002-2009.[5] This involves the placing and aggregation of the pairing of teams and match results from 2,048 regular season matches into the construction of 8 separate 32×32 matrix observations for both **X** and **W**.

A look at the remainder of table 1 reveals the varying nature of tournament designs, mostly on the basis of season length, as all other listed examples are based on 30-team leagues. In a comparative sense, the 1995-1998 NFL tournament design was similar, except with slightly greater weighting on the intra-divisional and power-balancing aspects. Despite an identical divisional composition (6×5) and season length of 82 games, the NHL and NBA differ markedly in terms of schedule concentration, with

[5] From 2010, a new rotational system applies, similar to the previous system, though altered slightly to reduce the overall travel burden on East Coast teams.

NHL teams playing a significantly higher proportional number of games against intra-conference teams than in the NBA. The previous (2001-2009) NHL system was even more highly concentrated, with values of k_l ranging from 0 to 8. The MLB system is the most difficult to describe, with general differences between divisions, depending on how many teams in both that division and the corresponding conference ('league' in MLB terminology), and further differences still on a team-specific basis.

A general way to describe the concentration of schedules is to use the Herfindahl index from industrial organization. The results are outlined for each of the major leagues in table 2. Initially, the index is deflated by the index from a theoretical 'idealized' schedule, involving the most even allocation of games (between opponents) possible holding both *n* and season length fixed. On this basis, the MLB schedule is the most concentrated, due to the long season and there being little inter-league play. As examples from the 2009 season, the index ratio was 2.15 for the Boston Red Sox (AL) and 2.02 for the Philadelphia Phillies (NL). However, if the index is instead deflated by the benchmark index from a truly balanced schedule, then the NFL provides the most concentrated schedule out of the major leagues. While this is somewhat generated merely by the short season, it demonstrates that the NFL presents the most interesting test case for the adjustment procedure, as it provides the greatest scope for identification of large adjustments should they be present.

While applicable to any type of unbalanced schedule, the NFL is easily the most suitable candidate of the major leagues for this technique. In addition to the concentrated schedule, it also overcomes the other major leagues for numerous compelling reasons. Firstly, the allocation of competition points is equal for all games

– unlike the NHL, where the total allocation of points is greater if the match extends to overtime, creating drawbacks for some within-season competitive balance measures. Another problem with the NHL is discontinuity of the sample arising from the 2005 season cancellation due to the players' strike. Secondly, MLB is the least suitable candidate, owing to the non-uniform nature of schedules between teams discussed earlier combined with the scarcity of inter-league play (only about 6% of matches, compared to 25% in the NFL). Finally, the comparative analysis is more powerful when the composition of teams is identical throughout the sample. This is the case for the NFL; with the same 32 teams since 2002 (even the NBA has had two relocations since 2005).

5. Results

The descriptive statistics of the raw win-ratios for all 32 teams over the sample are displayed in table 3. As shown, there is considerable turnover of standings over the sample (indicative of between-season competitive balance), with only the Detroit Lions failing to have a winning season. Inversely, the New England Patriots and Indianapolis Colts were the only teams that did not have a losing season. These three teams, and the Oakland Raiders, are also the only teams to have a mean win-ratio over the eight seasons outside the bounds of $1/3$ and $2/3$. The team-specific standard error of win-ratios over the sample – useful as a between-season measure of competitive balance according to Humphreys (2002) – supports this contention. A look at the middle column of table 4 shows that the mean value of the team-specific standard errors is higher than the analogous figures for the other major leagues.

As for within-season, one may use a range of measures to describe NFL competitive balance. Firstly, we consider the most widely-used measure – the ratio of the actual standard deviation to the idealized (binomial distribution of match results as in (5)) standard deviation (Noll, 1988; Scully, 1989), represented as

$$\text{ASD/ISD} = \sqrt{\sum_{i=1}^{n}(w_i - 0.5)^2 \Big/ n} \times 2\sqrt{\sum_{j=1}^{n} x_{i,j}} \qquad (11)$$

The right-hand column of table 4 shows that the mean dispersion of NFL win-ratios is the lowest (most balanced) of the major leagues, according to ASD/ISD. Recalling the section 2 debate on labor market and revenue-sharing restrictions influencing competitive balance, high NFL balancedness (both within- and between seasons) is due arguably to the considerable equalizing effect of the restrictive suite of policies (relative to other major leagues) used by the NFL. However, this has recently become far less so, with the abolition of the salary cap in March 2010.

As a means of robustifying the results to alternative measures, three others are also reported, as formalized in equations (12)-(14). Respectively, these measures are: (i) the Herfindahl index of competitive balance (HICB), similar to the baseline Herfindahl index, except allowing *n* to be time-varying; (ii) the concentration index of competitive balance, C12ICB, measuring the proportion of total wins by the top twelve teams in that season – the number that qualifies for the playoffs;[6] and (iii) the standard Gini coefficient.[7] For (ii) and (iii), w_i are assumed rank-ordered.

[6] The twelve teams assigned by simple rank-order is typically not the precise set of teams that make the playoffs. This is another distinctive feature of the conference and divisional system.

[7] While GINI is variant to changes in n and $\sum_{j=1}^{n} x_{i,j}$ over time, neither changes during the sample, circumventing this crucial concern. Otherwise we would be unable to apply the Utt and Fort (2002) methodology, as it depends on the (unobservable but assumed) *ex-ante* rank-order of teams under NFL tournament design, of which there are 32! $\left(\approx 2.63 \times 10^{35}\right)$ possible combinations.

$$\mathrm{HICB} = 4\sum_{i=1}^{n} w_i^2 \Big/ n \tag{12}$$

$$\mathrm{C12ICB} = \sum_{i=1}^{12} w_i / 6 \tag{13}$$

$$\mathrm{GINI} = \left(n^2 \Big/ 4\sum_{h=1}^{n}\sum_{i=1}^{h} w_i - n \right) - 1 \tag{14}$$

Initially of curiosity are the magnitudinal changes to the win-ratios arising from the adjustment. It is shown in table 5 that the outliers are quite large, because of the high schedule concentration noted previously. Out of 256 adjustments (for 32 teams over 8 seasons), a total of 11 are greater in absolute magnitude than the value of a win (0.0625), of which 6 are negative and 5 positive, and a further 58 are greater than the absolute value of half-a-win. The largest single adjustments are -0.0819 (Seattle Seahawks in 2007) and 0.0818 (Cleveland Browns in 2004). The adjustments are also scatter plotted against the original win-ratios in figure 1. Fortunately, there appears to be little systematic bias, providing technical backing for the methodology.

Normally of acute concern is the extent to which hypothetical adjustments create qualitative differences, in terms of changes to team rank-order. This is not quite as much of an issue in the conference and divisional system, since it is recognized that not necessarily the best teams will always make the playoffs. Nevertheless, under the criteria specified for playoff qualification, it is still of interest to see whether different teams would have qualified notionally under adjusted win-ratios.

Specifically, there are a total of only nine cases where this occurs over the sample, eight of which involve AFC teams. One case amazingly involves three teams (AFC East in 2002), in which New York Jets fall from first to third, promoting New

England Patriots from second to the division title. The eight remaining cases involve only two teams swapping positions. Of these, four have no bearing on playoff positions, while a further two (AFC North in 2005 and NFC East in 2009) involve top-two teams in cases where second earns a wildcard, affecting only seedings. The two remaining cases would have altered playoff qualification, with the Denver Broncos beating Kansas City Chiefs for an AFC wildcard in 2006 and New England Patriots topping Miami Dolphins for the AFC East 2008 division title. In addition, there are three further cases (again all AFC) in which adjusted standings would have altered wildcard places, with Denver Broncos replacing Cleveland Browns in 2002, and Miami Dolphins (2003) and Baltimore Ravens (2004) both replacing the Broncos. The latter case is the only one (out of twelve) in which the combined adjustments of the two teams involved overcome the value of a win.[8] In most cases involving teams tied on wins (14 out of 22), the adjustment assigns an ordering not inconsistent with that derived from the secondary and tertiary criteria specified currently by the NFL to separate teams tied on wins.

Aspects of the adjustments at an aggregate level are also of intuitive appeal to fans. Table 6 provides a profile of possible favoritism in the fixture. Beginning with the right-hand column, it is seen that the mean adjustment for AFC teams is 0.0088, a sizable figure over 128 observations. This results from AFC teams winning the majority of inter-conference games over the sample. Furthermore, the divisional adjustments are positive across the board in the AFC, along with NFC East. By team, Tennessee Titans are the only AFC team with an easier than average schedule over the 8 seasons, whereas only 4 NFC teams have positive mean adjustments, three of

[8] In numerous other cases, the adjustment changes playoff seedings, however, this is of less concern.

them in East division. The mean adjustment is a descriptive outlier – greater than two standard errors using mean adjustments from all teams – for Oakland Raiders (positive) and Seattle Seahawks (negative), with the difference between them of 0.0643 greater than one full win per season. The Seahawks case demonstrates nicely the outcomes arising from the procedure. While their raw win-ratio of 0.5234 is slightly above average over the sample, they play a disproportionate number of games against three teams (the other NFC West teams) that have performed weaker than average – the adjusted win-ratio (0.4876) takes this into account.

Returning momentarily to figure 1 and the line of best fit, further inspection also reveals a noticeable negative correlation $(\rho = -0.3706)$, indicating that teams achieving a higher win-ratio on average received more favorable fixtures, though this factor accounts for a small fraction of the variation. This result is totally intuitive, since a team with a fixed level of talent and performance will (all else equal) achieve a higher win-ratio when allocated an easier schedule. When mean adjustments by divisional place finish are computed, as in table 7, it is found that the means are increasing montonically in rank-order, as expected. If, however, the adjustments are based on rank-order from the previous season (one season of observations lost), the correlation changes sign $(\rho = 0.1102)$, so do the signs of all the mean adjustments in table 7, though all are close to zero. The interpretation is that the power-balancing aspect to the fixture evens up the competition, though to a lesser degree than intended since changes in standings from one season to the next are difficult to predict.

Most crucially, the calculated unadjusted measures of competitive balance are shown in the top portion of table 8. It is observed (on the basis of ASD/ISD) that the 2002 season was comfortably the most competitively balanced of the sample (in fact, most even since 1995), and while 2006 was also more even than average, all other seasons produced a similar competitive balance levels. This finding is reinforced by the other various measures reported in table 8. The adjusted competitive balance measures and adjustments are also computed and exhibited in the middle and bottom portions of table 8, respectively. The information from table 8 is replicated graphically in figure 2. The most striking result is that these adjustments are negative for all seasons across all measures, although in some years the magnitude of the adjustments are relatively large (most noticeably in 2003), and close to zero in other years (2004, 2006 and 2007). These figures indicate in a quantitative sense that the NFL is considerably more competitively balanced when the unbalanced schedule is accounted for.

Naturally, the issue of statistical significance of differences between unadjusted and adjusted arises. While any basic nonparametric sign test or rank-sum test would definitively reject the null of no statistical difference, we seek further support or otherwise to substantiate this. To this end, the more powerful parametric difference in means *t*-test for paired samples is applied. Since there is no *a priori* reason to presume that the adjusted measures should be lower, the two-tailed version is used. The *t*-statistics for ASD/ISD and HICB of 3.7842 and 3.8697 respectively are both easily above the 5% critical value of 2.3646, and while the values of 2.9991 for C12ICB and 3.3911 for GINI are also significant at 5%, they are insignificant at 1%. Despite this apparently strong result, caution is exercised for two reasons: (i) small sample size; and (ii) the distribution of competitive balance measures may be non-normal. One possible reason for our result may be that the level of ability of a team converges on the level of other teams that they play more often. However, even if true, it is difficult to conclude outright that the conference and divisional tournament design plays any part in generating that result. The same application to data from the CFL – a league with only 8 teams but with a similar (albeit reduced-form) conference-style unbalanced schedule, produces a difference of means *t*-test statistic for the ASD/ISD measure of merely -0.5707, making it comfortably insignificant.[9]

6. Conclusion

This paper has presented a modified unbalanced schedule model that can be generalized to any league in which teams play other teams any number of times. The model was then applied to NFL data over the 2002-2009 regular seasons. The major conclusion is that, quantitatively, the adjusted competitive balance metrics imply the league to be more competitively balanced than indicated by unadjusted measures in every season, and significantly so for the sample means. Qualitatively, the findings, while consistent, are less categorical. The logically following question of whether adjusted or unadjusted measures more accurately reflect the influence of uncertainty of outcome on demand for American football is one that is difficult to answer using average season attendances, since virtually all NFL games are sell-outs. Regardless, adjusted measures are still of much consequence to sports economists. Especially so for economists using microeconomic models to make inferences about the effectiveness of interventionist league policies, and using within-season competitive balance metrics from different sports leagues as a means of comparative analysis to test those inferences empirically. This is because the use of adjusted measures could alter conclusions about which leagues are more competitively balanced.

[9] The CFL is more comparable to the NFL than the other major leagues. Not only does it involve virtually the same sport and an almost identical season length (18 games), but data from seasons 1998-2001 and 2006-2009 is used, ensuring identical sample length. To again guarantee the same set of teams through the sample, seasons 2002-2005 are excluded, when the ill-fated Ottawa Renegades also competed. The sample ASD/ISD means are 1.4041 (unadjusted) and 1.4233 (adjusted).

There was also an appealing story behind the adjustments generated from the procedure. At an aggregate level, in terms of both by team and by rank-order in division standings, the means were sufficiently large to raise concerns about the high concentration of the schedule. While it is improbable that the league's constituent teams would collectively favor overhauling the NFL tournament design radically in the short run, they would likely be open to the idea of at least debating the issue upon being made aware of this story. The findings presented and discussed here have crucial implications for the various major league commissions, in terms of their ongoing reforms to their labor market and revenue-sharing policies, as well as tournament design itself.

Table 1: Finer Details of Tournament Design in the Major Leagues

League	n	Divisions × Teams	k_l	n_{k_l}	Season Length	Notes
NFL 2002-present	32	8 × 4	0 1 2	18 10 3	16	Power-balancing aspect 2 out of 16 games. Out-of divisional games rotational
NFL 1995-1998	30	6 × 5	0 1 2	17 8 4	16	Power-balancing based on 4/16 games. Only inter-conference games rotational
NHL 2010-present	30	6 × 5	1 2 4 6	12 3 10 4	82	Mix of inter-conference opponents between playing once and twice, completely rotational
NHL 2001-2009	30	6 × 5	0 1 4 8	5 10 10 4	82	Mix of inter-conference opponents between playing not at all or once, completely rotational
NBA 2005-present	30	6 × 5	2 3 4	15 4 10	82	Mix of intra-conference but inter-divisional teams play 3 or 4 times rotational
MLB 1997-present	30	1 × 6 4 × 5 1 × 4	Not uniform		162	Not highly structured like other leagues - teams tend to play 6-19 games against teams in their own division. They also play 15-18 games against inter-league teams

Table 2: Concentration Measures of Tournament Design of Major Leagues

League/Period	Actual Herfindahl Index Divided by Theoretical from: Idealized Schedule	Balanced Schedule
NFL: 2002-present	1.3750	3.0345
NFL: 1995-1998	1.5000	2.9095
NHL: 2010-present	1.3898	1.5139
NHL: 2001-2009	1.8051	1.9662
NBA: 2005-present	1.0847	1.1816
MLB: 1997-present	≈ 2.00-2.20	≈ 2.15-2.40

Table 3: Descriptive Win-Ratio Statistics of All Teams in the NFL (2002-2009)

Team	Mean	Standard Error	Maximum	Minimum
Bills	0.4297	0.0779	0.5625	0.3125
Dolphins	0.4453	0.2098	0.6875	0.0625
Patriots	0.7500	0.1531	1.0000	0.5625
Jets	0.4766	0.1600	0.6250	0.2500
Ravens	0.5469	0.1662	0.8125	0.3125
Bengals	0.4570	0.1806	0.6875	0.1250
Browns	0.3672	0.1473	0.6250	0.2500
Steelers	0.6367	0.1686	0.9375	0.3750
Texans	0.3828	0.1473	0.5625	0.1250
Colts	0.7734	0.0814	0.8750	0.6250
Jaguars	0.4922	0.1651	0.7500	0.3125
Titans	0.5547	0.2017	0.8125	0.2500
Broncos	0.5781	0.1145	0.8125	0.4375
Chiefs	0.4453	0.2277	0.8125	0.1250
Raiders	0.3125	0.1637	0.6875	0.1250
Chargers	0.6172	0.2044	0.8750	0.2500
Cowboys	0.5625	0.1602	0.8125	0.3125
Giants	0.5391	0.1668	0.7500	0.2500
Eagles	0.6367	0.1456	0.8125	0.3750
Redskins	0.4219	0.1326	0.6250	0.2500
Bears	0.4922	0.1873	0.8125	0.2500
Lions	0.2422	0.1434	0.4375	0.0000
Packers	0.5781	0.1912	0.8125	0.2500
Vikings	0.5313	0.1250	0.7500	0.3750
Falcons	0.5039	0.1623	0.6875	0.2500
Panthers	0.5547	0.1313	0.7500	0.4375
Saints	0.5156	0.1760	0.8125	0.1875
Buccaneers	0.4688	0.2059	0.7500	0.1875
Cardinals	0.4063	0.1377	0.6250	0.2500
Rams	0.3672	0.2301	0.7500	0.0625
49ers	0.3906	0.1558	0.6250	0.1250
Seahawks	0.5234	0.1828	0.8125	0.2500

Table 4: Mean Competitive Balance Metrics for Major Leagues (2002-2009)

League	Win-Ratio Standard Error	ASD/ISD Ratio
NFL	0.1626	1.5339
MLB	0.0525	1.8221
NBA	0.1316	2.6305
NHL*	0.0835	1.7073

*From 2001 to cover the exclusion of the 2005 ‘lockout’ season. Based on wins and ties (up to 2004), and overtime results equivalent to regulation time (from 2006), irrespective of league points allocation.

Table 5: Adjustment Factors on Win-Ratios of All NFL Teams

Team	2002	2003	2004	2005	2006	2007	2008	2009
Bills	-0.0273	0.0663	0.0137	-0.0060	0.0175	0.0136	-0.0489	0.0116
Dolphins	0.0098	0.0158	0.0466	-0.0410	0.0155	0.0249	-0.0330	0.0566
Patriots	0.0274	-0.0035	0.0043	0.0118	0.0120	-0.0151	-0.0135	0.0197
Jets	0.0020	0.0233	0.0275	0.0193	0.0275	0.0154	-0.0273	0.0176
Ravens	0.0038	-0.0389	0.0528	0.0194	-0.0095	0.0096	0.0275	0.0255
Bengals	0.0250	-0.0430	0.0430	-0.0174	0.0078	-0.0411	0.0457	-0.0038
Browns	-0.0117	0.0330	0.0818	0.0038	-0.0003	-0.0663	0.0642	0.0057
Steelers	-0.0086	-0.0040	-0.0015	-0.0018	0.0469	-0.0428	0.0335	-0.0097
Texans	0.0095	0.0643	0.0019	0.0231	0.0194	0.0156	0.0176	0.0059
Colts	-0.0175	0.0003	0.0081	-0.0309	0.0237	0.0257	0.0061	-0.0152
Jaguars	0.0018	0.0369	0.0294	-0.0271	0.0195	0.0217	0.0311	-0.0059
Titans	-0.0154	-0.0193	0.0057	0.0037	-0.0234	0.0040	-0.0309	0.0391
Broncos	0.0294	0.0040	-0.0116	0.0101	0.0411	0.0136	-0.0430	0.0273
Chiefs	0.0273	-0.0720	0.0488	0.0079	0.0137	0.0076	0.0250	0.0076
Raiders	0.0353	0.0076	0.0643	0.0310	0.0465	0.0076	0.0135	0.0213
Chargers	-0.0078	-0.0042	-0.0154	0.0606	0.0082	0.0060	0.0156	-0.0368
Cowboys	-0.0060	-0.0350	0.0116	0.0255	-0.0292	0.0062	0.0001	-0.0057
Giants	-0.0135	0.0466	0.0116	-0.0018	-0.0039	0.0197	0.0100	0.0352
Eagles	-0.0232	-0.0154	-0.0368	0.0272	-0.0194	0.0625	0.0167	-0.0096
Redskins	0.0253	0.0252	-0.0275	0.0431	0.0174	0.0567	-0.0215	-0.0159
Bears	0.0134	-0.0137	-0.0412	-0.0369	-0.0680	0.0410	-0.0234	-0.0059
Lions	-0.0159	0.0291	-0.0079	-0.0021	-0.0023	0.0410	0.0425	0.0113
Packers	-0.0408	-0.0077	-0.0389	0.0232	-0.0117	-0.0212	-0.0001	-0.0525
Vikings	-0.0060	-0.0410	-0.0195	-0.0136	-0.0118	0.0039	0.0079	-0.0505
Falcons	-0.0028	0.0330	-0.0604	-0.0078	-0.0020	0.0076	-0.0350	0.0059
Panthers	-0.0157	-0.0486	0.0058	-0.0447	-0.0273	0.0214	-0.0037	0.0391
Saints	0.0001	0.0000	-0.0352	0.0134	-0.0311	-0.0215	-0.0039	-0.0641
Buccaneers	-0.0095	0.0058	-0.0295	-0.0447	-0.0120	-0.0292	-0.0175	0.0446
Cardinals	-0.0060	0.0349	-0.0431	0.0018	-0.0178	-0.0664	-0.0117	-0.0507
Rams	0.0058	-0.0583	-0.0117	-0.0197	-0.0156	0.0016	0.0211	0.0054
49ers	0.0079	0.0097	-0.0238	0.0310	-0.0059	-0.0412	-0.0548	-0.0234
Seahawks	0.0038	-0.0311	-0.0527	-0.0602	-0.0253	-0.0819	-0.0100	-0.0295

Table 6: Mean Aggregate Adjustment Factors in the NFL (2002-2009)

Team		Division		Conference	
Bills	0.0051	East	0.0089	AFC	0.0088
Dolphins	0.0119				
Patriots	0.0054				
Jets	0.0132				
Ravens	0.0113	North	0.0071		
Bengals	0.0020				
Browns	0.0138				
Steelers	0.0015				
Texans	0.0197	South	0.0071		
Colts	0.0000				
Jaguars	0.0134				
Titans	-0.0046				
Broncos	0.0089	West	0.0122		
Chiefs	0.0082				
Raiders	0.0284				
Chargers	0.0033				
Cowboys	-0.0041	East	0.0055	NFC	-0.0088
Giants	0.0130				
Eagles	0.0003				
Redskins	0.0129				
Bears	-0.0169	North	-0.0100		
Lions	0.0119				
Packers	-0.0187				
Vikings	-0.0163				
Falcons	-0.0077	South	-0.0116		
Panthers	-0.0092				
Saints	-0.0178				
Buccaneers	-0.0115				
Cardinals	-0.0199	West	-0.0193		
Rams	-0.0089				
49ers	-0.0126				
Seahawks	-0.0359				

Table 7: Mean Aggregate Adjustment Factor by Placing within Division

Divisional Rank-Order	Mean Adjustment Based on Rank-Order in the:	
	Current Season	Previous Season
1	-0.0178	0.0016
2	-0.0064	0.0015
3	0.0100	-0.0023
4	0.0142	-0.0009

Table 8: Original, Adjusted and Change in Competitive Balance Measures

	Original			
Season	ASD/ISD	HICB	C12ICB	GINI
2002	1.3005	1.1057	1.3125	0.2241
2003	1.5104	1.1426	1.4167	0.2728
2004	1.5155	1.1436	1.3958	0.2701
2005	1.6677	1.1738	1.4375	0.3107
2006	1.4252	1.1270	1.3542	0.2476
2007	1.6346	1.1670	1.4167	0.2954
2008	1.6310	1.1663	1.3906	0.2950
2009	1.5861	1.1572	1.3854	0.2881
Mean (2002-2009)	1.5339	1.1479	1.3887	0.2755
	Adjusted			
Season	ASD/ISD	HICB	C12ICB	GINI
2002	1.2705	1.1009	1.3073	0.2185
2003	1.3341	1.1112	1.3715	0.2342
2004	1.4838	1.1376	1.3910	0.2622
2005	1.5662	1.1533	1.4041	0.2861
2006	1.3933	1.1213	1.3470	0.2432
2007	1.6112	1.1623	1.4053	0.2925
2008	1.5537	1.1509	1.3883	0.2813
2009	1.4920	1.1391	1.3477	0.2638
Mean (2002-2009)	1.4631	1.1346	1.3703	0.2602
	Change			
Season	ASD/ISD	HICB	C12ICB	GINI
2002	-0.0301	-0.0048	-0.0052	-0.0057
2003	-0.1763	-0.0313	-0.0452	-0.0387
2004	-0.0317	-0.0059	-0.0048	-0.0079
2005	-0.1015	-0.0205	-0.0334	-0.0247
2006	-0.0320	-0.0056	-0.0071	-0.0045
2007	-0.0234	-0.0047	-0.0113	-0.0029
2008	-0.0773	-0.0154	-0.0024	-0.0137
2009	-0.0940	-0.0181	-0.0377	-0.0242
Mean (2002-2009)	-0.0708	-0.0133	-0.0184	-0.0153

Figure 1: Adjustment Factor Against Original Win-Ratio

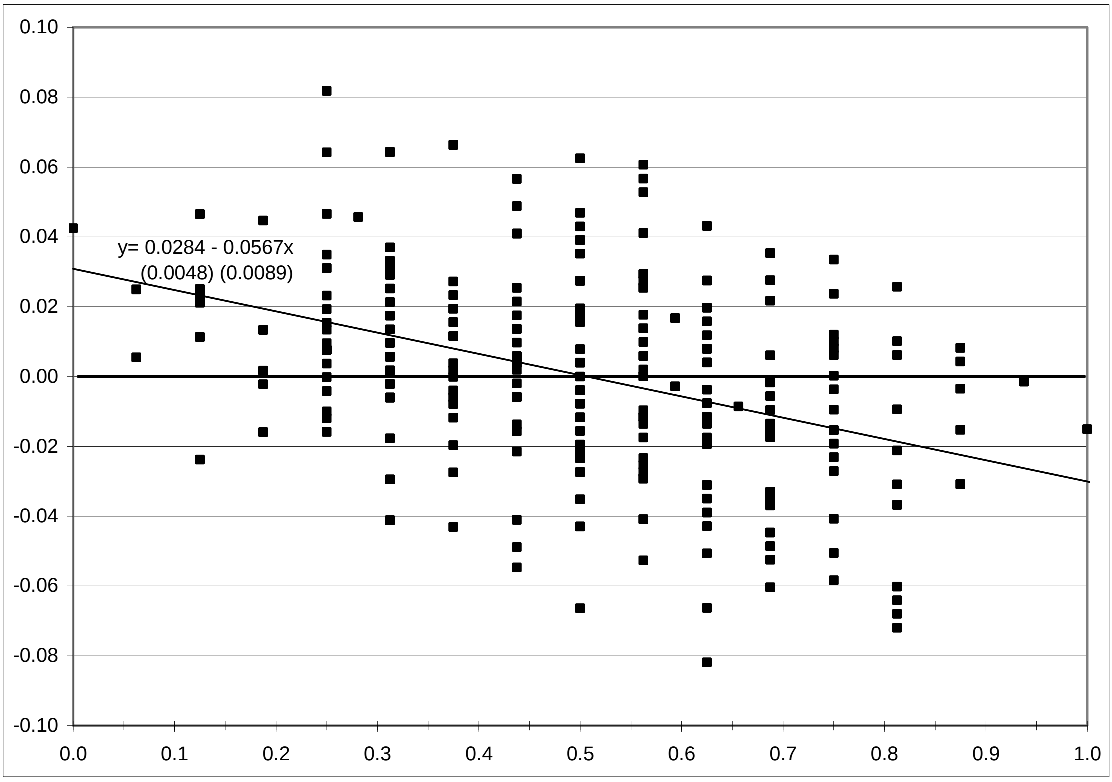
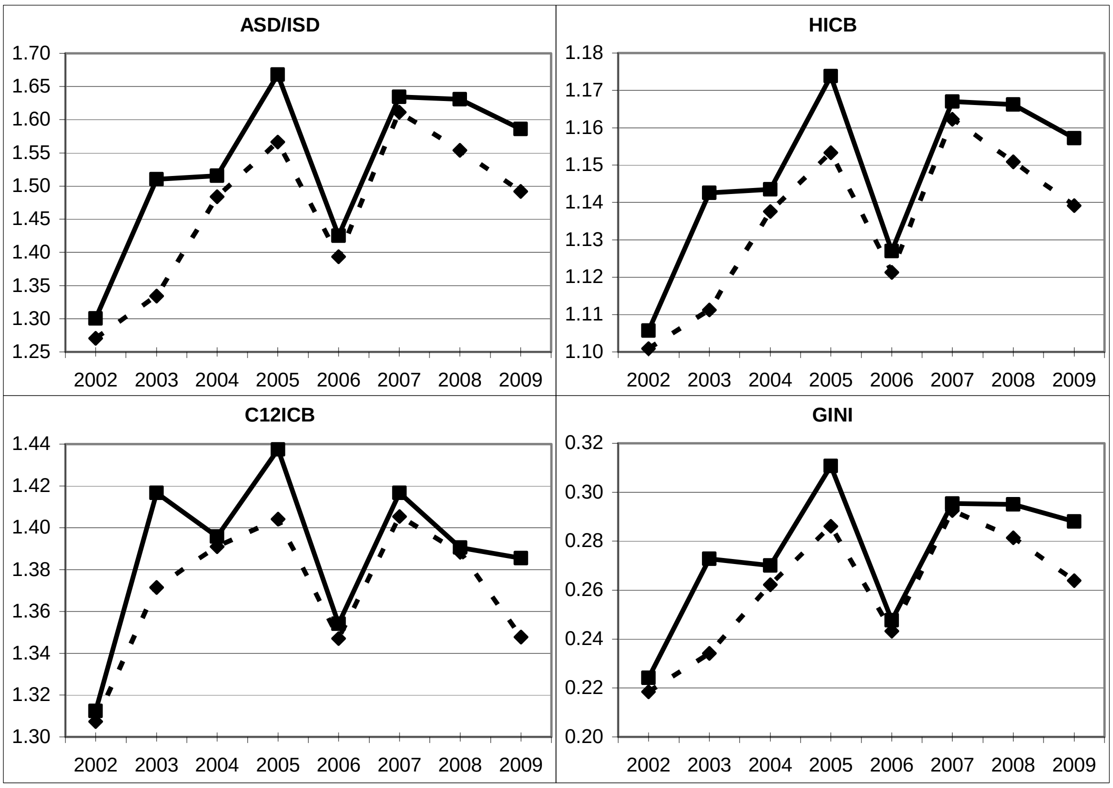


Figure 2: Original (Thick Line) and Adjusted (Dashed Line) NFL Measures



References

Blodget, H. (2010), "NFL To Investigate Colts For Tanking Game Against Jets", *Business Insider*, 3 January, URL: http://www.businessinsider.com/henry-blodget-nfl-to-investigate-colts-for-tanking-game-against-jets-2010-1

Butler, M. R. (2002), "Interleague Play and Baseball Attendance", *Journal of Sports Economics*, 3 (4), 320-334.

Cain, L. P. and Haddock, D. D. (2005), "Similar Economic Histories, Different Industrial Structures: Transatlantic Contrasts in the Evolution of Professional Sports Leagues", *Journal of Economic History*, 65 (4), 1116-1147.

Depken, C. A., II (1999), "Free-Agency and the Competitiveness of Major League Baseball", *Review of Industrial Organization*, 14 (3), 205-217.

Dietl, H. M., Lang, M. and Werner, S. (2009), "Social Welfare in Sports Leagues with Profit-Maximizing and/or Win-Maximizing Clubs", *Southern Economic Journal*, 76 (2), 375-396.

Fainmesser, I. P., Fershtman, C. and Gandal, N. (2009), "A Consistent Weighted Ranking Scheme with an Application to NCAA College Football Rankings", *Journal of Sports Economics*, 10 (6), 582-600.

Fort, R. and Lee, Y. H. (2007), "Structural Change, Competitive Balance, and the Rest of the Major Leagues", *Economic Inquiry*, 45 (3), 519-532.

Fort, R. and Quirk, J. (1995), "Cross-Subsidisation, Incentives, and Outcomes in Professional Team Sports Leagues", *Journal of Economic Literature*, 33 (3), 1265-1299.

Humphreys, B. R. (2002), "Alternative Measures of Competitive Balance in Sports Leagues", *Journal of Sports Economics*, 3 (2), 133-148.

Jech, T. (1983), "The Ranking of Incomplete Tournaments: A Mathematician's Guide to Popular Sports", *American Mathematical Monthly*, 90 (4), 246-266.

Késenne, S. (2000), "Revenue Sharing and Competitive Balance in Professional Team Sports", *Journal of Sports Economics*, 1 (1), 56-65.

Lee, T. (2010), "Competitive Balance in the National Football League After the 1993 Collective Bargaining Agreement", *Journal of Sports Economics*, 11 (1), 77-88.

Lee, Y. H. and Fort, R. (2005), "Structural Change in MLB Competitive Balance: The Depression, Team Location and Integration", *Economic Inquiry*, 43 (1), 158-169.

Lenten, L. J. A. (2008), "Unbalanced Schedules and the Estimation of Competitive Balance in the Scottish Premier League", *Scottish Journal of Political Economy*, 55 (4), 488-508.

Lenten, L. J. A. (2010), "The Extent to Which Unbalanced Schedules Cause Distortions in Sports League Tables", *Economic Modelling*, doi:10.1016/j.econmod.2010.08.003.

Noll, R. G. (1988), Professional Basketball, Stanford University Studies in Industrial Economics: 144.

Paul, R. J. (2003), "Variations in NHL Attendance: The Impact of Violence, Scoring and Regional Rivalries", *American Journal of Economics and Sociology*, 62 (2), 345-364.

Paul, R. J., Weinbach, A. P. and Melvin, P. C. (2004), "The Yankees Effect: The Impact of Interleague Play and the Unbalanced Schedule on Major League Baseball Attendance", *New York Economic Review*, 35, 3-15.

Quirk, J. P. and Fort, R. D. (1992), *Pay Dirt: The Business of Professional Team Sports*, Princeton University Press, Princeton.

Rottenberg, S. (1956), "The Baseball Players' Labor Market", *Journal of Political Economy*, 64 (3), 242-258.

Scully, G. W. (1989), *The Business of Major League Baseball*, University of Chicago Press, Chicago.

Szymanski, S. and Késenne, S. (2004), "Competitive Balance and Gate Revenue Sharing in Team Sports", *Journal of Industrial Economics*, 52 (2), 165-177.

Utt, J. A. and Fort, R. D. (2002), "Pitfalls to Measuring Competitive Balance with Gini Coefficients", *Journal of Sports Economics*, 3 (4), 367-373.

Vrooman, J. (1995), "A General Theory of Professional Sports Leagues", *Southern Economic Journal*, 61 (4), 971-990.

Vrooman, J. (2009), "Theory of the Perfect Game: Competitive Balance in Monopoly Sports Leagues", *Review of Industrial Organization*, 34 (1), 5-44.

Appendix

An alternative procedure (more suited to team rating techniques) can also correct the elements in $\mathbf{V}$ for the unbalanced schedule, producing quantitatively different adjustment factors to section 3. Here, the correction factor is disaggregated by team, so that team i's respective head-to-head records against each team influence i's corrected win-ratio. Initially, a neutral season matrix based on a balanced schedule is defined, $\mathbf{B}$, in which every team plays each other on an equal number of occasions

$$\mathbf{B} = \frac{1}{n-1}\begin{bmatrix} 0 & 1 & \cdots & 1 \\ 1 & 0 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 0 \end{bmatrix} \tag{A.1}$$

Using this, one can calculate a matrix corresponding to unbiased estimates of win-ratios when corrected for unbalanced schedules, expressed neatly by invoking the (left) distributivity property as

$$\mathbf{W}(\mathbf{B} + (\mathbf{X} - \mathbf{B})) = \mathbf{WB} + \mathbf{W}(\mathbf{X} - \mathbf{B}) \tag{A.2}$$

and extracting the diagonal elements of $\mathbf{WB}$. The diagonal elements of $\mathbf{W}(\mathbf{X} - \mathbf{B})$ are the magnitudinal adjustments themselves. The off-diagonal elements also have an interpretation (albeit of little theoretical use) – if all $k_l > 0$, $(w \cdot b)_{i,j}$ is the estimated win-ratio for team i had they played team j's schedule. However, for the interpretation to be more meaningful, instead of assuming $w_{i,j} = 1/2$ for any $x_{i,j} = 0$ (as on p. 7), it can be estimated more accurately as the probability from the logistic relation

$$\ln\left(\frac{\mathrm{pr}}{1-\mathrm{pr}}\right) = \frac{(w \cdot b)_{i,i}}{(w \cdot b)_{j,j}} \tag{A.3}$$