OPTIMAL COMPETITIVE BALANCE IN A SEASON TICKET LEAGUE

RODNEY FORT and JAMES QUIRK

Theory predicts that a planner maximizing the sum of fan and owner surpluses from a league dominated by season ticket sales may prefer more, less, or the same level of competitive balance produced by a league of profit-maximizing owners. The optimal level of balance depends on the relationship between marginal impacts of talent rearrangements in larger-revenue and smaller-revenue markets. Ultimately, then, judging whether an increase in balance enhances welfare rests on careful and thorough empirical investigation. Our reading of the literature and the policy debate shows that this careful work remains to be done. (JEL D6, L83)

There is no question the level of play has decreased. Now, do games become more exciting? Are teams more evenly matched? No question. Is that good for the game or not? I don’t know. I really don’t know. I ask that question all the time.

—NFL Hall of Fame Quarterback Troy Aikman, quoted in Pedulla (2003).

I think that margin of competition, that margin of the difference between winning and losing in this league is very small, and I think that is great for the fans because every team comes in with an opportunity to win.


I. INTRODUCTION

In this article, we lay out the basic welfare foundation of optimal competitive balance for regular season play in a “closed” sports league where season tickets dominate sales. This setting best describes the National Football League (NFL), economically the most important of the four major North American leagues (NALs). We handle the case of single-ticket leagues like baseball in Fort and Quirk (2010). Our chosen focus is on balance during the regular season and details of all our modeling choices are given in the next section.

The quotes given above help to illuminate the policy issue addressed by the theory. Is the NFL too balanced as Hall of Fame Quarterback Troy Aikman’s quote suggests? Or is the level of parity on the field somehow “optimal” as suggested in Commissioner Goodell’s quote? Congressional hearings (U.S. Senate 2001) have even been convened on the subject in the remaining NALs.

For the policy-oriented literature, it is implicitly taken that more balance would be an improvement over the result generated by the leagues themselves (comprehensive reviews are in Fort 2006a; Fort and Quirk 1995; Szymanski 2003). This debate—among fans, reporters, and economists—lacks anything remotely resembling any competitive balance target, for example, an optimal level of competitive balance. From the perspective of optimality constructs, this intuitive belief that enhancing competitive balance would on net enhance general fan welfare may simply be Pareto noncomparable advocacy.

Rottenberg (1956) was the first to detail the problems associated with a lack of competitive balance. If outcomes on the field, court, or ice become too predictable, as when there are only a few very dominant teams, fans of perennially unsuccessful teams may stay away in droves and some teams in the league may actually go under. Also, even the teams that survive will have lower revenues if these disillusioned fans forsake the sport altogether. Thus, leagues have a vested interest in managing the level of competitive balance. The point of

ABBREVIATIONS
GIDP: Globally Invariant Drawing Power
NAL: North American League
NFL: National Football League
PSL: Personal Seat Licenses

Fort: Sport Management, School of Kinesiology, University of Michigan, 1402 Washington Heights, Rm. 3150, Ann Arbor, MI 48109-2013. E-mail rodfort@umich.edu
Quirk: Division of Humanities and Social Science, California Institute of Technology (Retired)

Economic Inquiry (ISSN 0095-2583)
doii:10.1111/j.1465-7295.2010.00264.x © 2010 Western Economic Association International
departure for this article is how the league’s profit-maximizing choice of competitive balance deviates from one specification of the social welfare-maximizing level of balance.

We address the question of optimal competitive balance comparing decentralized league outcomes to the level of competitive balance that maximizes the sum of consumers’ and producers’ surpluses. We realize that other possible Pareto optimal outcomes might be developed, but the surplus maximizing approach does have the virtue of utilizing theoretical tools that are conceptually amenable to relatively straightforward measurement and comparison in actual leagues. In particular, our main result is that whether the decentralized league result is too much or too little balance, relative to the surplus-maximizing level, is an empirical matter. If the marginal impacts of talent rearrangements are larger in smaller-revenue markets than in larger-revenue markets, then welfare is enhanced by rearranging talent to create more balance. But if the reverse is true, then less balance enhances social welfare. The elements required to actually assess this relationship, namely available data and careful empirical analysis, can settle the issue.

We also are not blind to the fact that some advocacy of particular mechanisms to enhance balance actually may be thinly veiled attempts to redistribute wealth from players to owners and among owners. But that is another virtue of our exercise. Future assessments of the distributional consequences of various approaches to competitive balance can now proceed with a firm grasp of the optimal target.

The article proceeds as follows. In Section II, we compare the decentralized league model and planner’s optimum for leagues that heavily utilize season ticket sales. Our analysis identifies the conditions that determine whether the planner would prefer more or less balance than this type of league will produce in its decentralized equilibrium. Section III lays out the policy implications. All our findings suggest that whether more balance is preferred to less rests on empirical questions that have yet to be assessed. Conclusions and suggestions for future research are given in Section IV.

II. OPTIMAL COMPETITIVE BALANCE IN A “SEASON TICKET” LEAGUE

In some leagues, season ticket sales dominate team revenue functions. For example, the NFL has only eight home games and two preseason games to sell. Team Marketing Report (2008) tabulates average seat prices weighted by the proportions of different types of seats in stadiums for all NFL teams. The highest of these is about $118 per game suggesting a season ticket price of approximately $1,180. If the team performs below expectations, fans have only lost the value of the few games they then choose not to attend. So, football owners are able to do what every owner would like to do, namely, transfer the risk that the team performs below expectations to fans. Fans confronted primarily with season ticket options must make their estimate of the value of that purchase primarily on the quality of the home team. Further, in terms of postseason chances, every game in the NFL is more important to fans, even those against poorer opponents, dampening the importance of visiting team quality in the fan purchase decision. Our first modeling choice is to focus on a season ticket league and we will assume that demand depends upon own ticket price and the team’s own winning percent.

For our second modeling choice, the season ticket league is analyzed using the “closed league,” competitive talent equilibrium model (originally, El Hodiri and Quirk 1971). Members of a closed league essentially face a completely inelastic supply of talent; “open league” members might increase their talent by importing it from some other league. The NFL, in particular, is distinguished on the closed league basis from other world leagues (e.g., world football). Recently, there has been some international talent migration in the other NALs.

The competitive talent market distinction is best portrayed by the classical Walrasian tatonnement referee. Using all the information on the impacts of one team’s talent choice on the other teams in the league, the referee’s price comes to rest where no league member would change their talent choice. While we find this competitive process acceptable for our needs, especially for NALs, we note that the veracity of the competitive talent market choice is currently under contention (see Fort 2006b; Fort and Quirk 2007; Szymanski 2004; Szymanski and Kesenne 2004). Fort and Quirk (2007) also provide an argument to the effect that the common knowledge requirement of rational expectations equilibrium is satisfied for closed, competitive talent market sports leagues. Formulating our welfare problem in a rational expectations framework would simply add more complications to what is
a very complicated problem (even in the present Walrasian case). Further, there is no need to delve into any mechanisms used to alter the league outcome, such as revenue sharing, in order to derive our comparisons between the league and the planner. The impacts of revenue sharing in a rational expectations equilibrium for this type of model are dealt with in Fort and Quirk (2007).

We also model owners as choosing a single price, rather than as price discriminators. As is well known from welfare economics, perfect price discrimination would lead the planner and the league to choose the same talent distribution, although the distribution of wealth would likely be different in the two cases. In the NFL, suspicions of price discrimination might be fueled by the presence of personal seat licenses (PSLs), for example. A PSL is a lump-sum payment that guarantees its holder the right to purchase a particular season ticket into perpetuity. Thus, PSLs resemble two-part prices that might be devices extracting consumers’ surpluses. Additional suspicions are raised when PSLs are auctioned since auctions also can be surplus extraction devices. Recently, the New York Jets were able to auction front-row seats around the 45 yd line for $82,000 each (Waszak 2008).

We ignore price discrimination for a number of reasons (and discuss auctions shortly). First, the degree to which PSLs really are a price discrimination mechanism is debatable. If selling rights to consecutive season tickets really is an attempt at price discrimination then, just as with magazine subscriptions, we would expect to see some form of inter-temporal “two-for one deal” marketing approach, but we do not.

Second, just as a season ticket sells a different product than a single-game ticket, namely, rights to seat location for the entire season, PSLs also sell a different product than a standard season ticket. Under a PSL, the rights to a particular seat, in a particular location, are lifetime rights. If the PSL replaces a lifetime-rights season ticket, there is no additional value. But if the PSL adds a lifetime right (as it clearly did for the New York Giants, given previous ticket-holder laments) then value is created. Further, depending on the package, the PSL grants other access to food and services.

Third, the variation in PSL prices across fans is based upon willingness to pay for different locations, a different experience altogether. The price of New York Giants’ PSLs ranges from $1,000 to $20,000 per seat with more than half of the seats costing at least $5,000 (Vacchiano 2008). The PSL does not include the price of the tickets.

Fourth, it ends up that season tickets purchased under PSL rights can be resold in the NFL’s official “secondary ticket market” created in 2007 (Veiga 2007). Indeed, PSLs themselves can be completely unbundled from the associated season ticket and sold separately, online (e.g., www.seasonticketrights.com). Preclusion of resale is typically thought to be a prerequisite for successful price discrimination.

The final reason for which we ignore price discrimination follows from practical observation on owner choices to centralize significant revenue portions, through their league, to be redistributed equally among owners. Especially, in the season-ticket dominated NFL, television revenues, branded team merchandise, and emerging electronic rights have all been centralized through the league. We join Kahn (2007) in observing that centralizing revenues at the league level limits the degree to which surpluses can be captured.

As for auctions, the price discrimination idea is that buyers will be lined up in order of willingness to pay so that surpluses are extracted. In addition, a PSL auction makes resale of season tickets different from the usual dissipation of consumer surplus; the surplus may not go to the reseller who bought tickets at face value without having to pay a PSL fee. Thus, if there is competition among resellers bidding for the PSLs themselves, even if the season tickets are resold, then the team, rather than the resellers, might still be capturing consumers’ surpluses.

This may occur but there are two offsetting features of PSL auctions to date. First, the New York Jets’ auction is the only one to date so the surplus extraction, if it exists, cannot be general. More typical is the simple setting of PSL rates by seat location as in the New York Giants case mentioned above. Second, resale is nationalized—the resale agreement is between the NFL (not individual teams) and the reseller. As just mentioned, this type of choice by owners reduces the chances for surplus extraction.

Our third modeling choice is to focus on regular season play. We further assume a league at a given absolute level of play, the major league level, and that all differences among teams at that level are relative differences (extensions are in Kesenne 2000; Marburger 1997; Rascher 1997). This builds Rottenberg’s (1956) outcome
in the league; with slopes downward for any revenue product of win percent, MRP the price of a unit of talent. Second, the marginal revenue product of talent is equalized across the league, (t̄ − t, w) = 0, i = 1, . . . , n. Equation (2) shows that all team owners set the marginal revenue product of talent equal to the marginal cost of talent. In Equation (2), even though t is a function of w, our Ticket Price Assumption (t is chosen to maximize revenue for any value of w so that (∂Di/∂ti)(ti/Di) = −1) ends up canceling out all terms involving dt/dw.

Let t̄ = ti(w̄), D̄ = Di(ti(w̄), w̄) and MRP̄ = MRPi(ti(w̄), w̄), i = 1, . . . , n, be the profit-maximizing ticket price, and ticket demand and talent demand evaluated at that price, respectively. There is a league profit-maximizing equilibrium at the optimal ticket price vector, talent choice vector, and price of talent, (t̄, w̄, p), if:

\[ \sum_{i=1}^{n} w_i = \frac{n}{2}. \]

Equation (3) shows that, at a league profit-maximizing equilibrium, marginal revenue product of talent is equalized across the league, that is:

\[ \sum_{i=1}^{n} w_i = \frac{n}{2}. \]

For what follows, we observe two implications from Equation (4). First, given our Ticket Price Assumption, this equilibrium also has total revenue maximized for the league as a whole. We make use of this observation in our specification of the planner’s optimum shortly. Second, in our model, it is possible for league revenues to be negative.
to be maximized, consistent with Equation (4) and the Ticket Price Assumption, at a constant vector \( w^* = 0.500 \), that is, a perfectly balanced league. However, Equation (2) makes it clear that this can only occur if \( \text{MRP}_i = \text{MRP}_j \) for all \( i, j = 1, \ldots, n \) and for any common equilibrium value of \( w \). As long as there is variation in talent demand itself, the league cannot be perfectly balanced. We make headway at this point by imposing well-ordered variation in ticket demands across the teams in the league:

**Globally Invariant Drawing Power (GIDP) Assumption:** Assume the set of teams \( I = \{1, \ldots, n\} \) is listed in order of drawing power such that:

i. \( i \geq j \), \( w_i \geq w_j \Rightarrow D_i(t, w_i) \geq D_j(t, w_j) \) for any common \( t \geq 0 \).

ii. Let \( q_D \) be the number of tickets demanded. Then, \( i \geq j \), \( w_i \geq w_j \Rightarrow D_i^{-1}(q_D, w_i) \geq D_j^{-1}(q_D, w_j) \) for any common \( q_D \).

iii. \( i > j \) implies \( t^*_i > t^*_j \) in equilibrium.

Generally speaking, this is the well-known larger- and smaller-revenue market distinction common in the analysis of sports leagues. Under the GIDP Assumption, team 1 occupies the largest-revenue market; team 2 is in the next largest-revenue market and so on down to team \( n \). This seems reasonable especially over any relevant team or league planning horizon since the location of teams helps determine their drawing power and team location is completely in the hands of the league itself.

We first look at competitive balance in the decentralized profit-maximizing equilibrium of the season ticket league (proofs of propositions are in the Appendix):

**PROPOSITION 1.** In an \( n \)-team season ticket league, with a competitive talent equilibrium, with GIDP among teams, league revenues are maximized at \((t^*, w^*)\) satisfying Equation (4) where \( t^*_i > t^*_j \) and \( w^*_i > w^*_j \), for all \( i > j \).

Proposition 1 shows that the decentralized league equilibrium exhibits competitive imbalance with larger-revenue market teams winning more than smaller-revenue market teams. As discussed earlier, the only time this will not be true is if (removing the GIDP Assumption) there are no larger- and smaller-revenue markets to begin with, that is, talent demand is identical in all markets.

We next consider the planner’s optimum. For simplicity, we have the planner take the monopoly pricing power of each team as given. Let \( C_i = \int_{t_i(w_i)}^{\infty} D_i(t_i(w_i), w_i) dt_i \) and \( R_i = t_i D_i(t_i(w_i), w_i) \) be fans’ surpluses and team’s surplus in market \( i \), respectively. [Correction added on 20 May 2010 after first publication online on 25 February 2010: in the preceding sentence, the expression \( C_i = \int_{t_i(w_i)}^{\infty} D_i(t_i(w_i), w_i) dt_i \) was corrected to \( C_i = \int_{t_i(w_i)}^{\infty} D_i(t_i(w_i), w_i) dt_i \) and \( R_i = t_i(w_i)[D_i(t_i(w_i), w_i)] \) was corrected to \( R_i = t_i D_i(t_i, w_i) \).] We use revenues for team surpluses since (noted above) decentralized profit maximization by owners in a season ticket league leads to maximization of the league’s total revenue anyway. The planner chooses the vector \( w \) (the distribution of talent and, hence, competitive balance) to maximize the sum of surpluses, accounting for the adding-up constraint.

The Lagrangean for this problem is:

\[
L = \sum_{i=1}^{n} (C_i + R_i) + \lambda \left( \sum_{i=1}^{n} w_i - \frac{n}{2} \right),
\]

and the first-order conditions are:

\[
\frac{\partial L}{\partial w_i} = \frac{\partial C_i}{\partial w_i} + \frac{\partial R_i}{\partial w_i} + \lambda = 0, i = 1, \ldots, n,
\]

and

\[
\lambda = \sum_{i=1}^{n} w_i - \frac{n}{2} = 0.
\]

Let \( t'_i = t_i(w'_i) \), \( D'_i = D_i(t_i(w'_i), w'_i) \) and \( \text{MRP}'_i = \text{MRP}_i(t_i(w'_i), w'_i) \), \( i = 1, \ldots, n \), be the welfare-maximizing ticket price, and ticket demand and talent demand evaluated at that price, respectively. The two derivatives in the first-order condition with respect to \( w_i \) are:

\[
\frac{\partial C_i}{\partial w_i} = \int_{t'_i}^{\infty} \frac{\partial D_i}{\partial w_i} dt_i,
\]

where \( t'_i = t_i(w'_i) \). [Correction added on 20 May 2010 after first publication online on 25 February 2010: Equation (7), \( \frac{\partial C_i}{\partial w_i} = \int_{t'_i}^{\infty} \frac{\partial D_i}{\partial w_i} dt_i \) was corrected to \( \frac{\partial C_i}{\partial w_i} = \int_{t'_i}^{\infty} \frac{\partial D_i}{\partial w_i} dt_i \) and \( \frac{\partial R_i}{\partial w_i} = t'_i \frac{\partial D'_i}{\partial w_i} = \text{MRP}'_i \), where \( t'_i = t_i(w'_i) \).] [Correction added on 20 May 2010 after first publication online on 25 February 2010: Equation (8), \( \frac{\partial R_i}{\partial w_i} = t'_i \frac{\partial D'_i}{\partial w_i} = \text{MRP}'_i \) +]
\[ t'_i \frac{\partial D_j}{\partial w_i} = \frac{\partial t'_i}{\partial w_i} D'_i + \text{MRP}'_i, \]
was corrected to \[ t'_i \frac{\partial D'_i}{\partial w_i} = \text{MRP}'_i. \]

In Equations (7) and (8), no terms containing \( \partial t'_i / \partial w_i \) appear since ticket revenues adjust optimally. Further, in Equation (7), regardless of which demand function is determined by a particular \( w_i \), the quantity demanded at \( t_i = \infty \) must be zero, that is, \( D_i|_{t_i=\infty} = 0 \), for any level of team quality, if the owner charges an infinite price then zero tickets are sold. Substituting Equations (7) and (8) into Equation (6), the first-order conditions become:

\[ \frac{\partial L_i}{\partial w_i} = \int_{t'_i}^{\infty} \frac{\partial D_i}{\partial w_i} dt_i - \text{MRP}'_i + \kappa = 0, \]
\[ i = 1, \ldots, n \]
and \[ \frac{\partial L}{\partial \kappa} = \sum_{i=1}^{n} w_i - \frac{n}{2} = 0. \]

[Correction added on 20 May 2010 after first publication online on 25 February 2010: the preceding paragraph has been amended in line with the post-publication corrections for Equations (7), (8) and (9), and Equation (9) was corrected.]

Finally, Equation (9) implies the following for the planner’s equilibrium:

\[ \text{MRP}'_i - \text{MRP}'_j = \int_{t'_j}^{\infty} \frac{\partial D'_j}{\partial w_j} dt_j - \int_{t'_i}^{\infty} \frac{\partial D'_i}{\partial w_i} dt_i, \]
\[ i, j = 1, \ldots, n. \]

As our earlier observation on the league profit-maximizing outcome, the first thing we observe here is that it is possible for welfare to be maximized, consistent with Equation (9) and the Ticket Price Assumption, at a constant vector \( w^* = 0.500 \), that is, a perfectly balanced league. Once again, identical demands allow this to happen since both the right-hand side of Equation (10) vanishes and \( \text{MRP}'_i = \text{MRP}'_j \) for all \( i, j = 1, \ldots, n \) and for all values of \( w \). However, as long as there is variation in talent demand itself, the league cannot be perfectly balanced since the right-hand side of Equation (10) cannot be zero.

However, unlike the league profit-maximizing outcome, adding the GIDP Assumption to the Talent Demand Assumption does not settle anything. Under these two assumptions (the latter has declining MRP with respect to \( w_i \)), we can only find that \( i > j \) implies \( (\partial D'_i/\partial w_i) < (\partial D'_j/\partial w_j) \) in equilibrium with \( w^*_i > w^*_j \); the marginal effect of an increase in talent on ticket demand is larger in the smaller-revenue market \( j \). But this is not enough to settle the issue since the sign of the right-hand side of Equation (10) also rests on the range of integration of these demand effects, that is, \( t_j > t'_j \) and \( t_i > t'_i \). Reasonably, \( t'_i > t'_j \), by the GIDP Assumption, but the sign of Equation (10) rests on the magnitude of this difference.

So, the difference between the league revenue-maximizing equilibrium result in Equation (4) and the planner’s optimum rests on the right-hand side of Equation (10). On the right-hand side of Equation (10), attendance for each team changes directly with a change in that team’s own talent level and surpluses follow suit. But it is the size of the marginal fan surpluses with respect to talent, in the smaller-revenue market compared to the larger-revenue market, which determines whether an increase in balance is welfare enhancing. Without assuming the problem away, and with the virtue of highlighting what ultimately will be an empirical issue, the following can be shown.

**PROPOSITION 2.** In an \( n \)-team season ticket league, with a competitive talent equilibrium, with GIDP among teams, with a planner that maximizes the sum of consumers’ and producers’ surpluses, if, for \( i > j \), \( \int_{t'_j}^{\infty} (\partial D'_j/\partial w_j) dt_j - \int_{t'_i}^{\infty} (\partial D'_i/\partial w_i) dt_i > 0 \) then welfare is maximized at \( (t(w^*),w^*) \) where (1) \( t'_j > t'_i \) and \( w^*_i > w^*_j \), for all \( j \neq i \) and (2) \( w^*_i < w^*_j \) and \( w^*_j > w^*_j \), for all \( j < i \).

Proposition 2 shows that the planner’s equilibrium is also characterized by imbalance, but less so than the decentralized league equilibrium, as long as the marginal fan surpluses with respect to talent are larger in the smaller-revenue market than the larger-revenue market [i.e., the right-hand side of Equation (10) is positive]. Otherwise, it could be the case that the league has chosen the welfare maximizing level of balance [the right-hand side of Equation (10) is zero], or that a decrease in balance could be welfare enhancing [the right-hand side of Equation (10) is negative].

To us, the intuition behind the condition in Proposition 2 concerning the relationship between \( \int_{t'_i}^{\infty} (\partial D'_i/\partial w_i) dt_i \) and \( \int_{t'_j}^{\infty} (\partial D'_j/\partial w_j) dt_j \) goes as follows. In the last five seasons (2003–2007), Forbes annual reports on NFL team finances show that the New England Patriots have been among the revenue leaders (second, four times; fifth, the other) and the most
successful team on the field among the revenue leaders (14-2, 14-2, 10-6, 12-4, 16-0) (data can be checked in Fort 2010). The Arizona Cardinals were at the other end of the revenue spectrum (among the bottom five revenue teams four times and dead last three times) and equally unsuccessful (4-12, 6-10, 5-11, 8-8). In the condition in Proposition 2, marginal surpluses would be lower in New England (the larger-revenue team) because the much higher level of success on the field would have their fans with a lower marginal willingness to pay for additional winning. We emphasize that this occurs at the margin—it is not that fans in smaller-revenue markets are willing to pay more in total to see winners than fans in larger-revenue markets. Of course, this is just an example and rigorous empirical work is required to pass judgment on this theoretical construct.

We also note that the foregoing allows for the possibility that the planner may choose balance in such a way that a dynasty team dominates. Technically, this is a postseason issue beyond our regular season model. However, there is a relationship since the winner of the regular season could be that dynasty team. The planner’s optimal level of balance may result in the same team winning year after year, albeit at a lower expected value if the planner would increase league balance. But from the planner’s perspective, this would occur because somehow the presence of a dynasty team was also an important element determining fan willingness to pay. Of course, the ultimate value of Proposition 2 is that it states the conditions under which an increase in balance will be welfare enhancing in the form of a testable hypothesis. A suitable statistical test is to determine whether the right-hand side of Equation (10) is positive. Marginal impacts of winning percent on attendance can be obtained from estimating attendance demand. Since ticket prices are also available, the data required to do the test are observable.

III. POLICY IMPLICATIONS

Starting from this basic welfare theory foundation, what are the implications for the policy issues detailed in Section I? For the season ticket league, whether or not improved balance also improves welfare (defined as the sum of fans’ and owners’ surpluses) depends on the marginal impact on consumers’ surpluses in larger-revenue and smaller-revenue markets due to talent rearrangements in the two markets. The verdict of empirical work on this ambiguity will determine whether welfare will be improved by enhancing competitive balance, reducing it, or simply leaving it alone. This finding suggests that settling the debate over competitive balance requires knowledge of the relative sizes of changes in fan surpluses and ticket prices, between larger- and smaller-revenue markets, that accompany alterations in team qualities.

In essence, estimates of the impact of changes in quality on attendance actually measure the empirical importance of Rottenberg’s (1956) uncertainty of outcome hypothesis; how does attendance respond to increases in a team’s own quality? Estimates of these impacts to date have been somewhat clumsy and not clearly directed to the issues raised in Proposition 2, but they are evolving [the review in Szymanski (2003) covers their findings while the review in Fort (2006a) covers the shortcomings of the approaches]. The analysis here dictates that these estimates are required before making any attempt to alter competitive balance.

To make these observations a bit more concrete, as mentioned earlier, some worry that the NFL may be too balanced. This view rejects Proposition 2 so that less balance would enhance fans’ welfare. But this type of rejection is intuitive; only a careful econometric analysis of the direct effects stated in Proposition 2 would actually decide the issue.

While identification of mechanisms to either enhance or reduce competitive balance is out of place given our findings, our examination of optimal competitive balance does offer the following observation. In the event that league choices are judged suboptimal, owners in NALs cannot be expected to violate their profit-maximizing choices of their own volition. Also, there is no external regulatory agency to enforce the talent redistribution in an NAL that would have to occur. To date, there has only been Congressional brow-beating of the variety cited in Section I (U.S. Senate 2001). Thus, there is a regulatory structure hurdle to leap in order to move decentralized league decision making toward the welfare improvements under the planner’s optimum.

From this perspective, one policy prescription—breaking up NALs into competitive separate leagues—has two attractive features to recommend it (see Horowitz 1976; Noll 1976; Quirk and Fort 1999; Ross 1989, 1991). First, the structure for this type of intervention already exists under the antitrust laws. Fort (2007) lists...
the references in the argument over practical, case-by-case antitrust intrusions into sports, but the first important observation is that the structure for antitrust intervention is in place and breaking up production units has been accomplished in the past (e.g., the AT&T break up into seven “baby Bells,” finalized in 1982).

The second attractive feature of an antitrust move to break up sports leagues involves the chances for a Pareto improvement in fans’ welfare. Proponents cited above have developed the argument that, if two competing leagues were created from an existing league, the result would unleash competitive forces so that a team would exist in every economically viable location without sacrificing the goodwill investment that the owners have already made in existing teams. If the result of such a break up approaches the competitive distribution of teams, Pareto optimality would reign over the distribution of talent among these teams with the sum of producers’ and consumers’ surpluses maximized. This would be the planner’s outcome with optimal competitive balance detailed in this article. Competition would distribute teams so that any remaining smaller- and larger-revenue potential among the franchises would approximate the optimal distribution of talent in the planner’s outcome in Equation (10) for the season ticket league.

Other impacts of the imposition of competition have been covered in detail in the references cited above. For example, well-known from basic welfare analysis, moving from an inefficient outcome to the planner’s surplus-maximizing optimum, it will be the case that the distribution of talent changes because there will be more teams in more cities. As a result, teams that had more talent may now have less. But if this redistribution truly is a welfare improvement, then gains in under-served markets will more than offset.

To conclude our policy discussion, we are not so naïve as to believe that the sum of fan and owner welfare matters in the policy process. Indeed, public choice analysis often reveals that social welfare may matter little in the policy process. Every move away from any league’s profit-maximizing choice will have distributional consequences on players, owners, and fans. Indeed, leagues have a variety of mechanisms for attaining their own ends and the ones chosen and favored by leagues must have favorable distributional consequences for them. But identifying the optimal level of competitive balance is important because doing so sets the stage for the analysis of the distributional consequences relative to welfare maximization; it is possible to know the cost in terms of fan welfare of violating the planner’s optimum.

IV. CONCLUSIONS AND SUGGESTIONS FOR FURTHER RESEARCH

We attempt to seek remedy for the absence of considerations of the optimal level of competitive balance in North American pro sports leagues. We devise a planner’s optimal talent distribution for regular season play that maximizes the sum of fans’ and teams’ surpluses. This outcome is compared to decentralized, profit-maximizing outcome for a closed league where season ticket sales dominate team revenues (like in the NFL).

As long as owners in different locations face variation in demand functions, the profit-maximizing outcome yields competitive imbalance. The balance that the season ticket league has depends on the marginal surpluses created in larger-revenue markets relative to smaller-revenue markets following a planner’s alteration in the distribution of talent. Theoretically, the possibility remains that the league’s profit-maximizing distribution of talent also maximizes fan welfare.

For policy implications, first, only careful empirical tests can determine whether enhancing or reducing balance is welfare improving for our season ticket league. Careful estimates of impacts of talent choice on attendance demand for all teams are required in order to choose intervention mechanisms that effectively hit the optimal level of competitive imbalance. To date, this type of careful assessment is missing in the debate over competitive balance. Second, unless it also maximizes profits, we do not expect that owners will choose the optimal level themselves. Currently, there is no external regulatory structure governing NALs that could impose the planner’s optimum using the variety of mechanisms that are capable of changing competitive balance. Third, the antitrust remedy of breaking up the leagues does already have the requisite legal structure and precedence. Also, if the forces of competition can drive a Pareto result, then the optimal level of balance that maximizes the sum of fans’ and teams’ surpluses, detailed in this article, would be achieved.
There are many avenues for future work suggested by this analysis. We utilize the competitive talent market model most applicable to NALs. But initial investigations by Szymanski (2004), Szymanski and Kesenne (2004), and Easton and Rockerbie (2005) suggest that other leagues around the world may be better treated with noncooperative models. Also, extensions beyond gate demand to include local TV will no doubt prove insightful, for example, beyond the NFL (a league that sells all games in a national contract). For that matter, different owner objectives would produce different decentralized league outcomes for comparison to the planner’s optimum [most recently, Fort and Quirk (2004) and Kesenne (2005)]. Further, we do not address the case where Fort and Quirk (2004) and Kesenne (2005).

APPENDIX: PROOFS OF PROPOSITIONS

PROPOSITION 1. By assumption, based on simple observation [item (iii) under our GDP Assumption], \( t'_i > t'_j \). We seek to show that in equilibrium, \( w^*_i > w^*_j \).

For \( w^*_i > w^*_j \), the GDP Assumption has \( MRP_i > MRP_j \) for any \( 0 \leq w \leq 1 \) and for all \( j \neq i \). Since our Talent Demand Assumption has \( \partial MRP_i / \partial w_i < 0 \), then \( MRP_i > MRP_j \) for any \( 0 \leq w \leq 1 \) and for all \( j \neq i \) implies \( w^*_i > w^*_j \).

PROPOSITION 2. Again, item (iii) under the GDP Assumption suggests \( t'_i > t'_j \). We seek to show that if \( \int_0^{w^*_i} (\partial D_i / \partial w_i) dt_i - \int_0^{w^*_j} (\partial D_j / \partial w_j) dt_j > 0 \) [the right-hand side of Equation (10) is positive] then (1) \( w^*_i > w^*_j \), for all \( j < i \) and (2) \( w^*_i < w^*_j \) and \( w^*_j > w^*_j \), for any \( j < i \).

For any \( j < i \), \( w^*_i > w^*_j \) follows lines similar to the proof of Proposition 1, using the additional assumption that \( \int_0^{w^*_i} (\partial D_i / \partial w_i) dt_i - \int_0^{w^*_j} (\partial D_j / \partial w_j) dt_j > 0 \). For \( w^*_i > w^*_j \), the GDP Assumption has \( MRP_i > MRP_j \) for any \( 0 \leq w \leq 1 \) and for all \( j < i \). Since our Talent Demand Assumption has \( \partial MRP_i / \partial w_i < 0 \), then \( MRP_i > MRP_j \) for any \( 0 \leq w \leq 1 \) and for all \( j \neq i \) implies \( w^*_i > w^*_j \).

For \( w^*_i > w^*_j \), it has been noted that first our Talent Demand Assumption has \( \partial D_i / \partial w_i > 0 \) and \( \partial D_j / \partial w_j > 0 \) so that and \( \int_0^{w^*_i} (\partial D_i / \partial w_i) dt_i > 0 \). If we evaluate the planner’s optimum depicted in Equation (10) at \( w^* \), the decentralized league equilibrium, then the planner’s optimum is the same as the league profit-maximizing equilibrium if and only if \( \int_0^{w^*_i} (\partial D_i / \partial w_i) dt_i - \int_0^{w^*_j} (\partial D_j / \partial w_j) dt_j = 0 \). But for the planner to rearrange talent so that \( \int_0^{w^*_i} (\partial D_i / \partial w_i) dt_i - \int_0^{w^*_j} (\partial D_j / \partial w_j) dt_j > 0 \), the planner would have to move talent toward the smaller-revenue market, \( j \), implying that (locally, at least) \( w^*_i < w^*_j \). Then \( \sum_{i=1}^{n} w_i = n/2 \) implies \( w^*_j > w^*_i \), for all \( j < i \). Moving from the league profit-maximizing equilibrium to a planner’s optimum would be in the direction of more competitive balance.

REFERENCES


