IMPACT ON COMPETITIVE BALANCE FROM ALLOWING FOREIGN PLAYERS IN A SPORTS LEAGUE: EVIDENCE FROM EUROPEAN SOCCER

by

Ramón Flores*,

David Forrest**

and

J. D. Tena***

December, 2009

* Universidad Carlos III, Spain
** University of Salford, UK
*** Universidad Carlos III, Spain and Universita di Sassari, Italy

The authors acknowledge funding from the Consejo Superior de Deportes, Spain. J.D. Tena also acknowledges financial support from the British Academy under the Visiting Scholars Programme while working on this paper. Helpful comments were received from The Editor and from Ian McHale.
Abstract

Different frameworks yield different predictions concerning impacts on competitive balance from admitting foreign players to a sports league. We identify one reason to be that clubs may be represented as hiring simultaneously or in a leader-follower fashion. With the latter assumption, we show ambiguity in sign of effect, depending on how much extra talent becomes available. Whether balance improves or worsens when foreign players are introduced is therefore an empirical matter. We test for effects from liberalisation of player labour markets associated with the Bosman Ruling. Across seventeen European football leagues, results indicate improvement in within-season (but not cross-season) balance.

JEL: L1, J2, F2.
IMPACT ON COMPETITIVE BALANCE FROM ALLOWING FOREIGN PLAYERS IN A SPORTS LEAGUE: EVIDENCE FROM EUROPEAN SOCCER

I. BACKGROUND

The globalisation of the player labour market is one of the most striking developments in contemporary sport. Its effects are observed in America as clubs draw on players from, for example, Mexico and the Caribbean. But its most visible manifestation is in European soccer where lifting of restrictions on imported players by the governing body, UEFA, following the judgement in the Bosman case at the European Court of Justice in 1995, has lead to several leagues being staffed by a clear majority of foreign players. This has troubled UEFA sufficiently for it to lobby the European Commission for adjustment of the legal framework that would permit it, as it proposes to do, to introduce minimum numbers of domestic players who would have to be fielded by clubs. In other words, partial reversal is proposed of measures that effectively created a single market in player services.

Motives for objecting to the presence of foreign players are varied and may include simple national pride. But economic analysis can explore ways in which introducing foreign players impacts in tangible ways on the welfare of supporters. A number of potential effects could be examined, for example that on the performance of national teams. However, we focus here on the dimension of supporter welfare that has been most central in sports economics, the degree of competitive balance. We ask whether permitting international trade in player services will have any effect on balance in domestic leagues. So far, a consensus has failed to emerge on this question as different models yield different predictions.

Both Haan et al. (2002) and Késenne (2007) develop a framework derived from the pioneering El-Hodiri and Fort (1971) model of a two team league. They each predict that moving from autarky to free trade in player services will not result in any change in competitive balance. Some preliminary empirical support for this prediction is offered by Haan et al. (2007) who present standard summary indicators of balance that show little variation in European football as between the pre- and post- Bosman eras, at least in the seven ‘big’ leagues to which their analysis was restricted.

On the other hand, Schmidt and Berri (2003) and Berri et al. (2005) adopt a different theoretical perspective and it yields contrary predictions. They posit that a fixed number of places on professional teams is to be filled from an eligible playing population. Recruitment will be from the right tail of the

---

1 Effects here are evaluated in Baur and Lehmann (2007) for football and Alvarez et al. (2009) for basketball.
distribution of talent. Strong clubs will hire the most talented players on the extreme right and weaker clubs the next best. If the eligible population is small, there is liable to be a significant difference in ability between the best players and the remainder recruited. This is what generates competitive imbalance. However, if the pool from which players are selected is increased, the difference between best and next best becomes less pronounced. The gap in level of talent between clubs is then smaller and balance improves. Schmidt and Berri demonstrate that there were indeed historically significant improvements in Major League Baseball when the sport was racially integrated and, later, when foreign players were introduced. On this view of the world, a change from autarky to free trade in player services should change competitive balance, and favourably (at least in countries which import players).

Two different theoretical approaches therefore yield conflicting implications. The present paper aims partially to reconcile the two. We suggest in Section 2 that a key difference between them is that Haan et al. (2002) and Késenne (2007) both assumed that the two clubs act simultaneously in hiring players whereas in Schmidt and Berri (2003) the implicit assumption is of a leader-follower relationship where the very best talent talks to the stronger club first and the weaker club has to select from what remains. We view this as potentially more realistic. If the Haan et al. analysis is adapted accordingly, and if account is also taken of a practical upper limit on the number of players that can be absorbed by one club, predictions change such that there can now be an impact from introducing foreign players. But the sign of the impact is ambiguous and depends on the number of foreign players.

Given we then have ambiguity in the implications of opening up player labour markets, we go on to treat the issue as one that can be settled only empirically. In Section 3 we test for structural breaks in measures of competitive balance in seventeen European leagues at around the time of Bosman. The results suggest consistently that within-season competitive balance was affected by the Bosman Ruling and the impact was favourable. However, there is little evidence of change in the correlation of rankings across seasons.

The paper, then, essentially presents results from event studies. The event on which we focus we call ‘Bosman’ but by this we refer not just to the judicial ruling, as delivered at the European Court of Justice in December, 1995, but to the measures easing nationality restrictions that were introduced in the following year as football responded. In fact, the Bosman case, which was about whether clubs could require the payment of a transfer fee if an out of contract player wanted to join a new team, itself had rather limited applicability to the question of restrictions on foreign players (the issue of nationality was incidental to the case, but relevant to the extent that the footballer Bosman wished to make a cross border move between Belgium and France). At the time, UEFA enforced a ‘3+2’ rule on its members, meaning that no club could play a domestic first division match or a European competition match with more than three foreigners on the field (two more foreign born players were permitted if they met a test based on length of service and early career association with the host country). The Court found that this was in fact
inconsistent with the freedom of trade provisions of the Treaty of Rome in so far as it was applied to players moving between states in the European Economic Area (EEA). But UEFA’s response went further than just making changes required by the judgement. Its jurisdiction extended to many more countries than that of the European Court of Justice. Plainly, anomalies could arise if some countries’ clubs could field a full team of foreigners (so long as they were from within the European Economic Area) but clubs in other countries, outside the EEA, were still bound by 3+2. In due course, therefore, the response of UEFA to the Bosman case was to remove all its restrictions on the participation of foreign players in national leagues and in its own pan-European club competitions.

The liberalisation of the player labour market after Bosman saw substantial increases in the proportion of expatriate players across Europe. For example, Andreff (2006, p. 327) notes the growth in the proportion of foreigners in each the ‘big five’ European leagues as between 1995-6 and 1998-9. The largest changes were in Germany (27% to 39%), Italy (17% to 33%) and Spain (29% to 40%). Further, the effect was not confined to major leagues: in Belgium, more than 40% of First Division players were foreign. Though there was migration internal to the EEA, large numbers of players also reached big and small country leagues in Western Europe from Eastern Europe, South America and Africa. Gerrard (2004, p. 50) illustrates this by reference to the high proportions of players in 2002 World Cup squads who represented Western European clubs in domestic competition. So not only were the migrants numerous, they were also, as evidenced by taking part in the World Cup, disproportionately ‘talented’ players.

This sharp increase in foreign players had not been sought by the football authorities, which had indeed argued in the Bosman case that sport was a special case where nationality restrictions had specific and justifiable purposes (Breillat and Lagarde, 2006, p. 740). To this extent, the changes in rules that followed the judgement may fairly be regarded as exogenous. ‘Bosman’ therefore constitutes an experiment in which the event to be studied effectively allowed unlimited recruitment of extra territorial talent to a sports league. Our focus is on what this did to competitive balance.

II. GUIDANCE FROM THEORY

Models by Haan et al. (2002) and Késenne (2007) were developed explicitly to generate predictions of the effects there might be on competitive balance in European football given the deregulation of the labour market following Bosman. They differ in detail. In particular Haan et al. assumed profit maximising behaviour by clubs whereas Késenne, in the spirit of the ‘European model of sport’ (Sloane, 2006), took clubs as win maximisers. However, they came to the same conclusion (each predicted zero impact on competitive balance), perhaps unsurprisingly since each model is an adaptation of the same, very influential two team league model (El-Hodiri and Quirk, 1971).
In the two team model, two clubs play a series of matches against each other. The clubs differ in terms of ‘drawing power’. This power is likely to be associated with factors such as the size of the local population and of the stadium and the history of past success. Drawing power is exogenously given in the model. The two clubs hire from a labour market where there is an unlimited supply of ordinary players but a limited supply of talented players. It is the distribution of the latter group between the two clubs that determines the proportion of matches won by either side during the competition (and it is this difference in proportions of matches won that defines the degree of competitive imbalance in the league).

The distribution of talented players between the clubs is decided by the market, as if at an auction. In equilibrium, the wage of talented players is at a level that clears the market with the supply of talent equal to the sum of the demand from each club. The club with the greater drawing power has the higher demand for talent because, in a larger market, it can more readily convert playing success to revenue (through extra ticket sales, merchandising, etc). This results, in equilibrium, in the bigger club hiring a disproportionate share of the talented players and this is the source of competitive imbalance.

Now let the number of talented players present in the labour market increase (for example, through permitting immigration). The relative proportions of talented players hired by the two clubs will not change because their demand is in terms not of number of talented players but proportion of talented players (it is this that determines win percent and therefore revenue). To be sure, the wage rate for talented players will fall, permitting each club to have more talent on its staff. But in the new equilibrium the bigger club continues to have the higher share of talent because any given level of success has greater value in the larger market. Win proportions and therefore competitive balance are unaffected, hence the conclusion reached by Haan et al. (2002) and Késenne (2007).

However, the result depends on two assumptions that could be regarded as questionable. A formal model developed by Flores et al. (2008) shows that, if the two assumptions are relaxed, competitive balance is no longer necessarily invariant to the introduction of extra talented players to the labour market. The first assumption that is questionable is that all hiring of talent across clubs occurs simultaneously (as if at a single auction). In fact, talented players are likely to prefer to play for the club with the greater drawing power since big clubs offer superior opportunities for individual sponsorship and in terms of participation in international competition and selection for a national team. It may therefore be fairer to adopt a Stackelberg framework by assuming that the two clubs have a leader-follower relationship, where the bigger team leads by talking to the talented players first and hiring those that it wants. The smaller team then follows and hires from any talent remaining. Of course, in deciding how much talent to hire, the leader will take into account that engaging an extra talented player may not only make its own team
stronger but also make the other weaker (because that player will no longer be available for signing). Game theoretic strategic behaviour may therefore lead to more talented players being recruited to the big club than it needs to staff the team.

But this brings us to the second questionable assumption of previous models constructed to account for the effects of Bosman. This is that there is no constraint on hiring talented players in terms of absolute numbers. This assumption is crucial to the ‘no effect’ prediction because, when there is an exogenous increase in the number of talented players, each club continues, in equilibrium, to hire the same proportion of them. But a club might find it hard to absorb extra numbers. The bigger club, for example, might already have a full squad of talented players. In principle, it could hire more but some talented players would then find it difficult to maintain match fitness because they spend too much time on the bench, with implications for their future careers and selection for their national team. Under-used talented players are likely to demand a transfer and this places an effective constraint on how many talented players can be hired by a single club.

If the leader-follower framework is adopted, and if a constraint setting a maximum number of talented players at each club is introduced, then the predictions of the models change (Flores et al., 2008). An increase in the talent pool might now, after all, affect competitive balance. Three cases can be distinguished. An extreme case is where the initial size of the talent pool is small and the number of players added to it is small. In this case, all the extra talent is hired by the larger (leader) club and competitive balance will worsen. In another extreme case, the initial size of the talent pool is so large that the constraint on the maximum number of talented players is binding on both clubs. With each already satiated, increased availability of talented players has no effect on the competitiveness of the league. Yet another possibility is that the extra players are numerous enough, and the larger club already close enough to satiation, that more of the new talent goes to the small than to the big club. This will improve competitive balance.

Relaxing two critical but implausible assumptions therefore changes the predictions of models like those of Haan et al. (2002) and Késenne (2007). Essentially, their assumptions permitted them to focus on the distribution of a fixed supply of talented players. What mattered for competitive balance was not the number of such players but their proportionate distribution across clubs. This led to the conclusion that the degree of balance was invariant to the introduction of more talent. But, once the absolute number of talented players is accorded a role, it becomes clear that an increase in the talent pool can have a positive, zero or negative effect on competitive balance, depending on the size of the initial talent pool and the extent to which it grows. Whether Bosman had an effect in Europe, and if so in which direction, then becomes a matter that can be settled only empirically.
We noted above that Western European leagues generally appear to be net importers of talent. Since our focus is on Europe, we are therefore taking the perspective of the importer. The most important exporting regions are Africa and South America. The effects on countries in these regions could be analysed symmetrically with the approach described above. The sign of the impact on competitive balance in leagues in these regions will depend on the size of the initial talent pool and the number of those who migrate. For example, if the number of talented players is small and they are all initially at the big club(s), then their removal will promote greater balance. But, if the number of talented players is large and therefore dispersed across clubs, but so many of them migrate that only the big club(s) retain any talent at all, this will worsen competitive balance. Again, absolute numbers matter. Of course, it would be interesting to measure impacts for exporting countries empirically but we encountered practical difficulties in attempting to do so, in particular instability in league structures that made comparing measures of balance over time highly problematic. Our empirical tests below are therefore restricted to Europe.

III. IMPACT OF BOSMAN: EVIDENCE FROM SEVENTEEN EUROPEAN LEAGUES

We considered evidence from seventeen European leagues (the top divisions in Austria, Belgium, Denmark, England, Finland, France, Germany, Greece, Holland, Ireland, Italy, Norway, Portugal, Scotland, Spain, Sweden, and Switzerland). So far as we know, this constitutes, in terms of geographic coverage, the most comprehensive look so far at the consequences of Bosman.

The measures of competitive balance we considered were the same standard ones as those presented in Haan et al. (2002) and analysed again (this time with a non-parametric test instead of regression) in Haan et al. (2007). They were as follows.

(i) Concentration ratio is a measure, also employed by Koning (2000), to capture within-season competitive balance. For any season \( t \), it represents the number of points obtained by the top \( k \) teams in the final standings as a proportion of the maximum number of points those teams could have obtained. It is calculated as

\[
CR_{kt} = \left( \sum_{i=1}^{k} P_i \right) W \ast (R / 2) \ast k(2N - k - 1)
\]

where \( P_i \) is the number of points awarded during the season to the club that ended in position \( i \), \( W \) is the number of points awarded for a win, \( R \) is the number of times each team plays each other team (usually, but not always, two) and \( N \) is the number of clubs in the division. We computed this measure for \( k = 2 \), \( k = 4 \) and \( k = 8 \).
ii) *Fluidity* (labelled *DN* by Haan et al.) is intended to capture the dynamics in a national league as opposed to the dominance of a small number of clubs in a given season. It measures mobility in the final league standings between one season and the next and is calculated as $Fluidity_t = \frac{2}{N^2} \sum |r_i,t - r_i,t-1|$ where $r_i,t$ refers to the final ranking of club $i$ in season $t$ (with $i = 1$ signifying top place). Clubs that were not present in the division in both years in question are omitted from the calculation.

We had available for analysis a panel of seventeen countries observed over each season between 1986-7 and 2005-6. This includes more post-Bosman observations than was possible at the time of Haan et al. (2002). It is a shorter period than that included in the empirical analysis of Haan et al. (2007) whose ‘pre-Bosman’ era extended back to 1945. In our view this was too far as it leads to the measures of competitive balance before Bosman being influenced by evidence relating to pre-modern football, when many other things were different in the game.

For each of the measures in turn, we took a standard event study approach, testing for a structural break at the time the Bosman regime took effect. We count the Bosman era as starting from 1997-8 on the ground that the nationality restrictions were lifted in 1996 and, by 1997-8, there had been time for behaviour to adapt. We therefore compare 1986-7 to 1996-7 (pre-Bosman) with 1997-8 to 2005-6 (post-Bosman). It should be noted, however, that when we repeated our analysis with observations from 1996-7 omitted altogether (because it might be ambiguous whether Bosman had had an immediate impact), there were no material differences in findings compared with those presented below.

For each measure of competitive balance (*BALANCE*) in turn, we estimated the fixed effects model

$$BALANCE_{i,t} = a_i + bBOSMAN + \varepsilon_{i,t} \quad (1)$$

where $BALANCE_{i,t}$ is a particular measure of competitive balance for country $i$ at season $t$; $BOSMAN$ takes the value one for observations from 1997-8 to 2005-6 and zero during 1986-7 to 1996-7; $a_i$ and $b$ are parameters to be estimated, and $\varepsilon_{i,t}$ is an error term. The parameter $b$ captures any impact on $BALANCE$ from $BOSMAN$ and the estimate of $b$ is reported below as model 1 in Table 1.

However, a test for the Bosman effect based on the specification shown as (1) is not correct if the observations for $BALANCE_{i,t}$ are not independent identically distributed (iid) but serially correlated. We therefore augmented model 1 with the inclusion of a lagged dependent variable:

$$BALANCE_{i,t} = a_i + pBALANCE_{i,t-1} + bBOSMAN + \varepsilon_{i,t} \quad (2)$$
Across the four measures of BALANCE that we considered, the parameter \( \rho \) was always significant at the 5% level (and significant at the 1% level in three of the four cases). Therefore model 2 is preferred to model 1. But, because of the presence of the lagged term, the (steady state) Bosman effect is now estimated not by \( b \) but by \( B = \frac{b}{(1 - \rho)} \). This is the Bosman effect shown for model 2 in Table 1.

Model 3 is a test for robustness that adds to specification (2), as in Haan et al. (2007), a dummy variable, CHAMPIONSLEAGUE, set equal to one from 1990-1 to represent any impact from the replacement of the European Cup by the European Champions League:

\[
BALANCE_{i,t} = a_i + \rho BALANCE_{i,t-1} + bBOSMAN + cCHAMPIONS LEAGUE + \epsilon_{i,t} \tag{3}
\]

The impacts of Bosman with this specification are shown in Table 1 as model 3.

Next we tested the restriction in model 2 that the parameter \( b \) was equal for all seventeen countries in the sample. \( F \)-tests rejected this restriction at the 1% level for three of the four measures of balance covered by our exercise. Accordingly, we adapted model 2 to permit different slope parameters for each country:

\[
BALANCE_{i,t} = a_i + \rho BALANCE_{i,t-1} + b_iBOSMAN + \epsilon_{i,t} \tag{4}
\]

Now, the impact of Bosman is captured as an average, 
\[
B = \frac{\sum_{i=1}^{17} b_i}{(1 - \rho)17}
\]
and it is this that is reported as the model (4) result in Table 2.

Results, shown in Table 2, on the impact of Bosman on the concentration ratio measures (defined over two, four and eight clubs) are consistent across the models to the extent that the sign is always negative. Further, at least for the four- and eight-team concentration ratios, the impact is always significant at at least the 5% level. This suggests that the dominance of top clubs in European leagues was typically lower in the years following Bosman than in the years before. Of course, event studies are blunt instruments. One cannot rule out that other factors than globalisation of the player labour market were responsible for the evident improvement in within-season competitive balance from 1997. However, one of the most obvious general influences, the increased importance of television rights income, which tends to be concentrated in the top clubs (in particular those that qualify for the European Champions League), might reasonably have been expected to generate less rather than greater equality. The result of the exercise

\[\text{\textsuperscript{2}}\text{ We ran a similar test allowing both autoregressive and slope parameters to change across countries with no material differences in the results.}\]
therefore has some claim to refute the validity of UEFA’s claims during the Bosman case that sport has specific features that justify restrictions on labour mobility. It claimed (Breillat and Lagard, 2006, p. 740) that nationality and other restrictions were needed to protect competitive balance within domestic leagues. But, in fact, permitting import of talent has coincided with some degree of improvement in respect of at least one type of competitive balance.

On the other hand, within-season balance, as measured by the distribution of points in the final rankings, represents only one possible conceptualisation of what might be understood by a competitive football environment. Whether clubs move easily between positions in the final rankings is another and, with this focus, the pattern in our results is rather different. The impact of Bosman on Fluidity is negative and significant in three of the four specifications reported in Table 1, indicating that deregulation of the labour market may have made for less movement in positions within the final standings as between one season and the next. This is a different result from Haan et al. (2007), who used the same measure as us, but on a smaller sample of countries, and Vrooman (2007), who offered a measure of season-to-season competitive balance based on the extent to which a club’s win percent in one year successfully predicted its win percent in the next. They found no significant difference between the pre- and post-Bosman years in the degree of fluidity across seasons. A caveat to our contrary finding is that it disappears in the specification that includes CHAMPIONSLEAGUE. In practice, of course, it is difficult to distinguish separate effects from Bosman and the Champions League because of the limited number of observations when one was in place but not the other. Certainly the Champions League may have contributed to less mobility in the rankings because teams which qualify one year are liable to do so again given the revenue they gain from participation. Hence, our attribution of an adverse impact from Bosman should be treated with caution.

Whether concentration ratios or Fluidity provided the focus, there were few cases in our estimation where the country-specific coefficient estimate for $b_i$ in (4) was individually statistically significant. At the level of the individual country, there were evidently insufficient observations to conclude confidently that concentration ratios had fallen. But for two-, four- and eight-club concentration ratios, the 17 slope coefficients included, respectively 10, 13 and 15 negative signs. This provided adequate ground for concluding from the pan-European evidence that big team dominance had fallen after Bosman. Similarly, the slope coefficient on Fluidity was significant only once but signed positive fourteen times and this collective evidence from the panel allowed us to conclude that dynamic balance deteriorated after Bosman.
IV. CONCLUSIONS

The existing literature produces conflicting predictions concerning the implications for competitive balance when a sports league is opened to foreign players. We identified a key source of differences between models: some represent strong and weak clubs as hiring personnel simultaneously but an alternative is to posit a leader-follower relationship such that strong clubs have first pick of talent. With the latter assumption, we pointed out the possibility that liberalisation will improve competitive balance, depending on numbers of foreign players joining the pool from which talent may be recruited.

We investigated the effect of the Bosman Ruling. Although it was not followed by an improvement in dynamic balance across seasons, we did find evidence of a ‘favourable’ impact having followed from the Ruling in seventeen leagues across Europe. Our results from an analysis of pan European data therefore appear to make the case of football similar to that reported for baseball following expansion of the racial and geographical scope of player recruitment in that sport. Further, they are consistent with the general prediction in economics that globalisation promotes greater competition in local markets.

References


Table 1. Bosman effect on four measures of competitive balance (seventeen European football leagues)

<table>
<thead>
<tr>
<th></th>
<th>two-club concentration ratio</th>
<th>four-club concentration ratio</th>
<th>eight-club concentration ratio</th>
<th>Fluidity</th>
</tr>
</thead>
<tbody>
<tr>
<td>model 1</td>
<td>basic model</td>
<td>-0.015</td>
<td>-0.021</td>
<td>-0.027</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(-2.63)</td>
<td>(-4.51)</td>
<td>(-6.32)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(***)</td>
<td>(***)</td>
<td>(***)</td>
</tr>
<tr>
<td>model 2</td>
<td>takes account of serial</td>
<td>-0.010</td>
<td>-0.016</td>
<td>-0.022</td>
</tr>
<tr>
<td></td>
<td>correlation</td>
<td>(-1.37)</td>
<td>(-2.58)</td>
<td>(-3.81)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(**)</td>
<td>(***</td>
<td>(**)</td>
</tr>
<tr>
<td>model 3</td>
<td>includes CHAMPIONS LEAGUE</td>
<td></td>
<td></td>
<td>-0.005</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(-0.68)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(*)</td>
</tr>
<tr>
<td>model 4</td>
<td>allows different effects</td>
<td>-0.012</td>
<td>-0.018</td>
<td>-0.023</td>
</tr>
<tr>
<td></td>
<td>across countries</td>
<td></td>
<td></td>
<td>(-2.02)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(***)</td>
</tr>
</tbody>
</table>

t- statistics are shown in parentheses; three, two and one asterisk indicate, respectively, significance at the 1%, 5% and 10% level