Sweet FA: Why foreign player crackdowns hurt English football

Ben Southwood
Abstract

It is a very common view that “importing” foreign football players into the UK to play in the Premier League leads to less opportunity for English players to play for these teams. This means that English players get less high-level experience, and consequently aren’t as good as the players of Spain, France, Italy or German, who make up a larger fraction of the players playing in their home leagues. This, the argument runs, is an important factor in explaining the English national team’s perceived underperformance in international competitions. I review the literature and present novel data establishing a negative relationship between current performance (as measured by FIFA ranking) and the current amount of football played in a league by native players (across Spain, England, Germany and Italy). Further, I find no relationship between minutes played by English players in the Premier League five or ten years ago and current performance. Finally, I find strong evidence that a league’s overall strength (as measured by its UEFA coefficient) is predicted by the current amount of foreigners playing in it. To restrict foreign players would not directly benefit the English national team, but it would risk substantially curtailing the overall quality of the world’s most popular football league.
Literature Review

Though this is a question of considerable academic interest, as well as practical import, it has barely been studied directly. Baur & Lehmann (2007) found that greater outward and inward mobility was associated with national success. Berlinschi, Schokkaert & Swinnen (2010) found that greater outward mobility led to learning which McGovern (2002) and Maguire & Pearson (2000) mention the issue but provide little evidence either way. Milanovic (2005) finds that greater international footballing mobility increases worldwide inequality among clubs, but decreases inequality among national teams, as players from countries with weaker national leagues can get high quality experience abroad.

Elliott & Weedon (2010) argues that an influx of foreign youths to English football academies is more “foot exchange” than “foot drain”—more foreigners means more exchange and circulation of knowledge and skills. According to Flores, Forrest & Tena (2010) more openness to foreigners in a league enhances its competitiveness. Alvareza, Forrest, Sanz & Tena (2011) look at the general case for amateur sports and find that extra foreigners competing in a given league means greater likelihood of qualification for world, European and Olympic championships.
Introduction

The intention in this paper is to provide a very simple and basic challenge to the most simple and basic (but widely repeated) claims of the football migration restrictionists. It is certainly possible that mechanisms outside the purview of this introductory paper are working, such that restricting football migration (e.g. along the lines of FA Commission (2014)), would improve the national team in other ways.

For example, the FA Commission argues that the key mechanism is 18-21 year old English players playing in some sort of competitive league with older players. For those purposes, the Championship, League One, League Two and others would suffice, not to mention foreign leagues. This paper cannot tackle that question—it looks only at countries’ top leagues and players of any age. Having said that, while the story the FA tells is reasonably plausible, the evidence they present is mainly anecdotal, sparse and lacking in rigour. Since the mean and median English professional footballer do not typically play for the national team, we need independent evidence that the mean or median minutes matter for the top of the distribution. And the evidence they present cannot determine whether English players get less experience than Spanish or German because fewer talented English players enter the sport or because of systemic/structural issues.

This paper is, as mentioned, a mainly “negative” paper, challenging a handful
of simple (but perenially popular) explanations for perceived underperformance. Thus, it does not advance an alternative explanation of why England’s men typically do worse at international football competitions than might be expected for a country with our population and GDP. But I will, fleetingly, raise one possibility that ought to be examined in future research: the idea of spreading out a fixed amount of national talent among a varying number of different sports.

England is competitive at the top level of cricket, rugby league, rugby union, golf, cycling, motorsport, recently the Olympics and even does relatively well at football, for all its fans’ lamentations. By contrast Germany, Europe’s most successful footballing nation, is particularly good at only football, ice hockey, handball and to some extent basketball, with a vast gulf between football and all the others. Assuming only that sporting talent is reasonably general (i.e. skill in one sport correlates well with skill in another sport), which it seems to be, countries will tend to do better at sports when they focus their talent on fewer areas. As I say, this is a very tentative hypothesis, and may only be seriously advanced if supported by future research.

In the following sections I will prevent evidence that:

(a) for none of the top four European Leagues (Italy’s Serie A, Spain’s La Liga, Germany’s Bundesliga, or England’s Premier League) does a lower total number of minutes played by natives lead to lower international performance during that year; in fact the opposite is more likely to be true

(b) neither Italy, Spain, Germany or England seems to do any better when their citizens are playing more total minutes, across the top four leagues as a whole (there seems to be more or less no relationship)

(c) the proportion, or total amount, of minutes played five, or ten, years ago by English players in the Premier League, is unrelated to international performance
Basic glossary and sources

**Minutes:** Total minutes played by a given nation’s native players in their top national league (e.g. total minutes played by English players in the English Premier League) over a year, given by its end-point (i.e. the 2013-14 season is recorded as 2014) [provided by Opta Sports]

**Proportion:** The minutes played by a given country’s natives in their national league, as a proportion of the total minutes played [author’s own calculations]

**UEFA coefficient:** A measure of the success of a given country’s teams in European competitions for a given football season, particularly the UEFA Champions’ League and the UEFA Europa League [retrieved from http://kassiesa.home.xs4all.nl/]

**FIFA ranking:** Calculated by FIFA on the basis of results, weighted by importance, location and the strength of the two teams—an accurate general measure of team performance, with some issues, but frequent enough for some statistical power. For a given season, the December of the year that season ended is used for comparison [retrieved from en.fifaranking.net]

**Total minutes across top leagues:** The total minutes played by a given nation’s footballers across all of the top four leagues (Italy, Spain, Germany, and England). For most years in my sample no English players played any minutes in most
Success in major championships: I give a nation a ranking for their performance in the European Championship and World Cup (first, second, third and fourth are easy to award; for further positions they get the average of those in their position class; if they fail to qualify I give them the total number of teams in the tournament plus one) [author’s own calculations]
Data

I have data on minutes for all seasons of the Premiership from 1992-3 until 2013-14 (the 2013-14 figure is proxied from the number I had 85% of the way through the season since most of the work was done when the 2013-14 season was unfinished). I have data on La Liga, the Bundesliga and the Serie A from the 2009-10 season up until the 2013-14 season. This means there are 22 English observations of minutes, and 15 observations from abroad, 37 in total.

There are an equal number of observations for FIFA Ranking and UEFA coefficient. For total minutes across top leagues there are only 20 observations—each of the four countries for five years each. There were major championships every two years from 1994, within my sample, so I have 20 observations across the four countries.

This is a fairly small sample size, especially in the lagged regressions, and so must be seen as a step toward a fuller picture of how national success is determined, rather than the endpoint. Still, if the effects were as strong as typically suggested in the media and by the so-called man on the street, we might expect to nevertheless see significant results.
Results: Same year
Figure 1. FIFA Ranking (one is best) vs. total minutes played by native players
Figure 2. FIFA Ranking (one is best) vs. proportion of home league minutes played by natives
Regression output 1: \texttt{allfifr} refers to the FIFA ranking a country had at the end of a year after the given season was through; \texttt{allmin} is the total minutes played by native players in their home league over a year; \texttt{allprop} is the proportion of minutes of a league that were played by native players; \texttt{alluef} is the home league’s UEFA coefficient for the season in question.

The first result is that the amount of football native players play in their home league appears to have a significant, and negative, relationship with the country’s FIFA ranking. When you put both proportion and minutes in a multivariate least squares regression you get a small negative coefficient on proportion (implying a country’s rank improves by 0.065 places for each extra percentage point of proportion) but at the same time you get a small positive coefficient on total minutes (implying that an extra 1,000 minutes of native players would worsen a country’s rank 0.067 places). Both are significant at the 1% confidence level. UEFA coefficient points (a decent proxy for the strength of the top teams in the league) appears to have no relationship.
Figure 3. FIFA points vs. total minutes played by native players
To test for robustness, we can use raw FIFA points instead of the FIFA rank. We might want to do this because a country’s rank is affected not just by its own endeavours but also those of others. If Spain, Germany and Italy do very well, England’s rank could fall even if its own quality hasn’t changed at all. Thus we can look at the raw points put out by FIFA’s algorithm, which track team quality more directly. However, we can only do this for the reduced sample (2009-2014) because FIFA’s algorithm changed twice between the start of my full sample (1992) and today, in ways that are hard to correct for (I try to do so in two experimental
ways later, but they are ultimately unsatisfactory). As you can see, even given this confound, rank and points are quite closely linked in our sample.

Regression output 2: **justnowmin** is the total number of minutes played by a country’s native players in their native league between the 2009-10 season and the 2013-14 season (inclusive), **allallfifp** is the FIFA points their national team earned.

What we see is a small positive relationship (an extra 1,000 minutes played by native players in their native league gets you approximately an extra FIFA point) but the relationship is insignificant.
Figure 5. FIFA points vs. total minutes played by nation’s players across top leagues

Regression output 3: allallmin is the total number of minutes played by a country’s players across the top four European leagues between the 2009-10 season and the 2013-14 season (inclusive), allallfifp is [again] the FIFA points their national team earned.

If we do the same thing for minutes played by a country's players across all the top four leagues (presumably an even better test of the total "good experience" players of a country are getting) then we get a very similar result—a small, but insignificant (closer to significance this time) benefit to the national team of having extra players play across all the top leagues. It fails to provide any extra support for the popular thesis.
Figure 6. Total minutes played by natives vs. position in major championship
Regression Output 4: minutesforevens is the minutes played by native players in their home leagues on even years (i.e. the years ending in European Championship and World Cups years), sameyearchamps is their position in those championships that year (as an example, if they went out in the Euro group stages, they’d be placed 12.5th, since the top eight teams go through; if they went out in the World Cup group stages, they’d be placed 24.th, since the top sixteen go through)

Here we find even smaller, and even less statistically significant results. Presumably some of this is coming from the impreciseness of the measure, but even so this does add to the weight of the evidence suggesting little or no relationship between current amount of minutes being played by native players in their top home league and national team success.
Results: Five- and ten-year lags
Figure 7. Minutes played by English players in Premier League vs. FIFA ranking five years later.
Figure 8. Proportion of Premier League minutes played by English players vs. FIFA ranking five years later

Proportion of total home league minutes played by natives 5 years prior
Regression Output 5: fif5 is the English FIFA ranking at the end of the year, min5 is the total number of minutes played by English players in the Premier League season ending five years earlier, prop5 is the proportion of total Premier League minutes five years ago that were played by English players, and uef5 is the UEFA coefficient earned for English teams’ success in Europe in that season five years earlier.

Overall what we see here is that there’s practically no relationship whatsoever. We can give numbers for the observed associations between UEFA ranking, proportion of minutes played by English players, and total minutes played by the English, but p-values are so high it’s pointless (i.e. this is very close to what we might expect to see if they varied independently). On top of that, the adjusted R-squared—a measure of how much of the variation in FIFA ranking five years later these five years earlier numbers explain—is just over 0.09. 91% of the relationship is left unexplained. This suggests that the total amount of minutes English players play at a given date has, on its own, little to no relationship with how well England does (five years) down the line.
Figure 9. Minutes played by English players in Premier League vs. FIFA ranking ten years later
Figure 10. Proportion of PL minutes played by English players vs. FIFA ranking ten years later
Regression Output 6: \texttt{fif10} is the English FIFA ranking at the end of the year, \texttt{min10} is the total number of minutes played by English players in the Premier League season ending ten years earlier, \texttt{prop10} is the proportion of total Premier League minutes ten years ago that were played by English players, and \texttt{uef10} is the UEFA coefficient earned for English teams' success in Europe in that season ten years earlier.

Taking the lag up to ten years reduces our observations, but we actually get closer to significance. However, what we see are results roughly in line with what we saw in the current time comparisons. Each 10 percentage point higher proportion of Englishmen playing the Premier League is associated with a 0.13 rank improvement on FIFA's scale. But each 1,000 extra minutes to Englishmen is associated with a 0.08 rank worsening. The adjusted R-squared is about 0.24 here, an improvement on the five year lag if still very low.
Results: The link between foreigners and club success in Europe

One good measure of the strength of the top teams in a league, and by extension the strength of the league (especially if those top teams rotate) is that league’s UEFA coefficient. This records how well (or poorly) their teams have done in European club competitions (such as the UEFA Champions League and the UEFA Europa League). Countries whose clubs do well are rewarded by more places (and more guaranteed places) in future competitions. The top three teams get four spots in the Champions League, and so on. It doesn’t accurately measure the strength of the whole league, since it only looks at the top seven or so clubs. Future research could probably look at the points schedule down the league to correct for the inaccuracy of the measure—for our purposes the measure will suffice, since relative to other European leagues, the English Premier League has a relatively shallow points curve.

Here we see a highly significant (i.e. easily significant at the 1% threshold) negative relationship between the overall strength and quality of a league (proxied by success in European competitions) and the total amount of minutes played by native players in their home leagues. Division of labour and specialisation, as well as simple added quality inputs (from countries with less prestigious, successful and
monied leagues) appear to add to success for the Premier League and elsewhere.

Figure 10a. Minutes played by natives in their home leagues vs. UEFA coefficient

Quantified, an extra 1,000 minutes played by natives leads to 0.36 fewer UEFA points—about a sixth of a win in a Champions League group match. Going back to 1990s levels of native participation would involve around 300,000 extra Premier League Minutes being played by natives.
```
Call:
lm(formula = alluef ~ allmin + allprop)

Residuals:
  Min     1Q   Median     3Q    Max
-55.771 -17.663   -5.763  19.212  61.142

Coefficients:
                 Estimate Std. Error  t value Pr(>|t|)
(Intercept)  1.429e+02  2.288e+01  6.245   4.15e-07  ***
allmin     -3.608e-04  1.223e-04  2.951    0.00571   **
allprop     1.754e+02  1.164e+02  1.508    0.14091

---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 29.37 on 34 degrees of freedom
Multiple R-squared:  0.3705,   Adjusted R-squared:  0.3335
F-statistic: 10.01 on 2 and 34 DF,  p-value: 0.0003824
```
Conclusion

This paper finds little evidence of any connection between the total amount of minutes played by English players in the Premier League and the success of the English national team measured by its FIFA ranking. Nor does it find any similar link when looking more broadly, and including all of the four major European countries’ top club leagues (Germany’s Bundesliga, Italy’s Serie A, and Spain’s La Liga). Nor does it find any evidence that the number total number of minutes played five or ten years ago correspond with the success of the English national team.

The finding—no relationship between total minutes, or the proportion of minutes played and success as measured by FIFA ranking—is robust to a number of different specifications, including using raw FIFA points and counting total minutes played across extra leagues.

By contrast, this paper finds a highly significant link, across all the top four major leagues, between the number of minutes native players play and the success of major teams in continent-wide club competitions. Allowing more foreign talent into a league closely corresponds with how well that league’s top teams are expected to do on the continental level.
Altogether, this evidence suggests that “cracking down” on foreign players would drastically harm the success of the English Premier League—the world’s most successful, popular, and (arguably) interesting club competition, without any corresponding benefit to the English national team. One possible explanation is that quantity of minutes played does matter, but its (negative) effect is balanced out by an opposite (positive) effect from enhanced quality of experience.

The evidence is not conclusive, as there are many ways this exploratory research could be developed to achieve a more complete picture. For example, data on young players could establish if there is a more direct channel to do with quantity of competitive experience (e.g. along the lines of the FA’s claims) whether or not it occurs in the country’s top league. Similarly, further research might clarify whether the literature is correct in claiming that foreign players help young English academy players or hinder them. But for now, the state of the published research suggests foreign players enhance competitiveness and the quality of the league, and do not harm the national team. The FA must do this extra research, or keep the current state of the literature in mind as it sets football policy.
Appendix: Experimental measures

Included only for completeness and interest, are two experimental ways I tried to get around the problems of the FIFA ranking system. While FIFA points give a very accurate picture of the strength of a team, taking into account their results, weighted by the importance of the match, the strength of their opponent and the strength of their opponent’s confederation, FIFA rankings are confounded by also taking into account the strength of other teams. For example, if Spain is in first, and loses to a few poor opponents (evidencing genuine worsening in quality) they will lose points, and may drop into second. Germany might rise to first without any concomitant improvement. But I couldn’t use the pure FIFA points (the preferred option) because their calculation method changed three times over the period, and there are no available calculations of what points would be under prior (or later) systems.

I tried to get around this in two ways. Firstly, what I called vs.top which took a country’s FIFA points (in December of a year) as a proportion of the top points in that December. Though this is also hopelessly confounded (your vs.top will rise if the top team’s FIFA points fall, even as yours remain the same) it allows one to make comparisons across periods and it is not obviously more confounded than FIFA rank (for one, they correlate pretty closely, see below).
Secondly, I calculated vs.max which took a country's FIFA points (again, in the December of a year) as a proportion of the total points scored during that era. For example, under FIFA’s third calculation system, running since July 2006, the top score ever earned is 1920pts, which Spain recorded in November 2010. A country’s vs.max is its December points as a proportion of 1920. This is a bit better, allowing us to compare periods but removing the most obvious confound from above. Essentially it allows us to combine the FIFA points series going back to 1992-93 (which we have English minutes data for) to the series for four countries 2010-2014 (regressions 2 and 3 above).

Figure 11. Correspondence between FIFA rank and FIFA points as a proportion of top scorer in that period (vs.top)
Figure 12. Correspondence between FIFA rank and FIFA points as a proportion of top scorer in that points system era (vs. max)
Figure 13. Correspondence between two experimental measures

For completeness, I also include here the outputs of regressions. The results we obtain are similar to the other regressions. For the second, and better measure (vs.max) we find no evidence of any relationship between our proxy for team quality and the amount of minutes played by natives in their home leagues (allmin), nor with the proportion of the total they play (allprop), nor with the European success of the league’s top teams (alluef).

Interestingly, for the first (and worse) measure vs.top we find some evidence of a relationship with the minutes played by natives in their top home leagues (but not the other regressors). Since all the better measures return a different result, it
seems likely that this comes as an artefact of the measure; perhaps because natives play more minutes when foreign players are weaker and hence when foreign national teams are weaker (and the denominator in vs.top is smaller). Further research may be able to further unpack this issue.

Call:  
`lm(formula = vs.top ~ allmin + alluef + allprop)`

<table>
<thead>
<tr>
<th>Residuals:</th>
<th>Min</th>
<th>1Q</th>
<th>Median</th>
<th>3Q</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-0.33225</td>
<td>-0.06535</td>
<td>-0.00569</td>
<td>0.09899</td>
<td>0.15914</td>
</tr>
</tbody>
</table>

| Coefficients:                      | Estimate | Std. Error | t value | Pr(>|t|) |
|------------------------------------|----------|------------|---------|---------|
| (Intercept)                        | 6.247e-01 | 1.352e-01 | 4.619   | 5.65e-05  *** |
| allmin                             | -8.548e-07 | 5.529e-07 | -1.546  | 0.1316 |
| alluef                             | -2.104e-04 | 6.918e-04 | -0.304  | 0.7629 |
| allprop                            | 1.046e+00 | 4.848e-01 | 2.158   | 0.0383 * |

---
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.1185 on 33 degrees of freedom  
Multiple R-squared: 0.1545, Adjusted R-squared: 0.07761  
F-statistic: 2.01 on 3 and 33 DF, p-value: 0.1317

Call:  
`lm(formula = vs.max ~ allmin + alluef + allprop)`

<table>
<thead>
<tr>
<th>Residuals:</th>
<th>Min</th>
<th>1Q</th>
<th>Median</th>
<th>3Q</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-0.22354</td>
<td>-0.07866</td>
<td>-0.03181</td>
<td>0.09270</td>
<td>0.26178</td>
</tr>
</tbody>
</table>

| Coefficients:                      | Estimate | Std. Error | t value | Pr(>|t|) |
|------------------------------------|----------|------------|---------|---------|
| (Intercept)                        | 6.914e-01 | 1.363e-01 | 5.074   | 1.48e-05  *** |
| allmin                             | -8.993e-07 | 5.570e-07 | -1.614  | 0.116 |
| alluef                             | -8.780e-04 | 6.970e-04 | -1.260  | 0.217 |
| allprop                            | 8.915e-01 | 4.885e-01 | 1.825   | 0.077 . |

---
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.1194 on 33 degrees of freedom  
Multiple R-squared: 0.1222, Adjusted R-squared: 0.04236  
F-statistic: 1.531 on 3 and 33 DF, p-value: 0.2248
References


