# Explaining International Soccer Rankings


Peter Macmillan and Ian Smith



*Journal of Sports Economics* 2007; 8; 202

DOI: 10.1177/1527002505279344

# Explaining International Soccer Rankings

PETER MACMILLAN
IAN SMITH
*University of St. Andrews*

*Existing research on the determinants of FIFA's international soccer rankings suffers from serious statistical problems, particularly sample selection bias and nonnormal errors. The authors correct for this by extending the data set by an additional 100 countries. Furthermore, they find important roles for new variables in the form of the size of population and a long history of international soccer in explaining world football rankings. The authors also investigate the determinants of an alternative ranking measure to that constructed by FIFA.*

**Keywords:** *international football rankings; history*

## INTRODUCTION

There can be little doubt that the team sport of soccer boasts both the greatest number of participants and spectators worldwide. But despite the evident global popularity of soccer and its significance as a leisure pursuit and business, studies of national team performance are scarce. There are only two empirical contributions of which we are aware. First, Torgler (2004) recently investigated the factors that affected the outcomes of matches in the FIFA World Cup 2002 hosted by Japan and South Korea.[1] Second, Hoffman, Lee, and Ramasamy (2002b; henceforth HLR) estimated the determinants of the variation in FIFA soccer rankings dated January 2001 for a cross-section of 76 countries.[2]

These rankings receive considerable public attention primarily for reasons of national prestige and esteem, and also as a visible indicator of relative performance and progress. As they measure the average level of success over a number of years, the FIFA rankings are not intended for use as a device for forecasting the results of either individual matches or major tournaments. Indeed, unless the difference in rankings is sufficiently large, they would not be expected to be particularly useful for predicting the results of individual games given the significant role in outcomes

JOURNAL OF SPORTS ECONOMICS, Vol. 8 No. 2, April 2007 202–213
DOI: 10.1177/1527002505279344

played by essentially random factors such as player injuries, motivation, fitness, and team spirit. The focus of our research, therefore, is on what determines longer term international soccer strength, as measured by the FIFA ranking scores, rather than the predictive accuracy of those rankings in international competition.

The same objective of explaining the variation in international soccer rankings was adopted in the study conducted by HLR. However, the method they use to select the sample of 76 countries is nonrandom. Because an overwhelming majority of their chosen countries is located above the median of the ranking distribution, the authors unintentionally sample on the dependent variable. As a result, their estimates exhibit both sample selection bias and also nonnormal errors. Our contribution remedies these problems by augmenting their sample with data on 100 additional countries. We also include new variables to capture the importance of population and country-specific football history for international performance. Furthermore, given doubts about the utility of FIFA's ranking algorithm, the robustness of the results is evaluated using an alternative ranking system.

## THE SAMPLE, VARIABLES, AND DATA

The 76 countries chosen by HLR are the medal winners at the Sydney 2000 Olympic games used in a companion study of Olympic success (Hoffman, Lee, & Ramasamy, 2002a). This simple sampling device was intended to avoid bias. However, HLR earlier state that

> these studies have found that the number of medals won by a country at the Olympic Games is partially explained by factors such as its per-capita GNP, population size as well as certain geographical, political and cultural influences. It would seem reasonable to suspect that variables explaining performance over a range of sports should partially explain the success of countries in international football. (pp. 256-257)

In other words, HLR expect that both Olympic medal winning and international soccer success are correlated with the same or similar set of explanatory variables. If this is so, then confining the sample of countries used to investigate football performance to Olympic winners is unlikely to generate a random selection. Rather, the likelihood is that a disproportionate number of countries will be chosen from the upper end of the international soccer ranking distribution. The descriptive statistics listed in the upper part of Table 1 document this outcome. Of the 76 countries selected by HLR, 66 are above the median FIFA ranking of 412 points and only 10 are below.

As the selection process is related to the value of the dependent variable (FIFA ranking points, denoted $Y_i$), it can introduce correlation between the error term and the regressors, leading to bias and inconsistency in the ordinary least squares (OLS) estimators. This will arise insofar as there are omitted variables that are correlated with the regressors and that jointly determine Olympic success and interna-

TABLE 1: Descriptive Statistics

| | *Mean* | *Median* | *Maximum* | *Minimum* | *Standard Deviation* |
|---|---|---|---|---|---|
| FIFA ranking points, $Y_i$ | | | | | |
| FIFA 203 countries | 385.1 | 412 | 821 | 8 | 201.1 |
| HLR 76 countries | 543.7 | 572.5 | 821 | 117 | 139.4 |
| HLR + 100 countries | 405.7 | 434 | 821 | 15 | 195.0 |
| $ELE_i$ | 55.3 | 54.4 | 87.1 | 28.7 | 13.2 |
| $GNP_i$ | 5884 | 1940 | 41860 | 90 | 8792 |
| $(TEMP - 14)_i^2$ | 81.2 | 70.6 | 237.2 | .0 | 68.3 |
| $HOST_i$ | .07 | .0 | 1.0 | .0 | .26 |
| $LATIN_i \times POP_i$ | .05 | .0 | 2.90 | .0 | .27 |
| $POP_i$ | 33.3 | 6.3 | 1262.5 | .04 | 125.6 |
| $HISTORY_i$ | 1946.1 | 1953 | 1999 | 1882 | 27.3 |
| $REPUBLICS_i$ | .08 | .0 | 1.0 | .0 | .27 |

NOTE: Unless otherwise indicated, the summary statistics refer to the extended sample of 176 countries. The variable labels are as defined in the text.

tional football performance. The magnitude of the bias is, of course, an empirical matter.

To address the sample selection bias issue, we collected data on 100 excluded countries. This omits 27 teams in the full FIFA list for which GNP per capita data were unavailable for 2001 or a nearby year in the World Bank database. Typically, these nonreporting countries are small island states (such as Montserrat and Anguilla) or political outliers (such as Iraq, Libya, and North Korea). Because their ranking is generally below the FIFA median, this still imposes some residual selection on the value of the dependent variable. Table 1 indicates that the extended sample of 176 countries has a ranking points mean of 405.7, marginally higher than that for the 203 FIFA teams of 385.1. The $z$ statistic for the null hypothesis that the mean $\mu$ of the extended sample is drawn from a population with $\mu = 385.1$ is 1.36, indicating nonrejection at conventional levels of statistical significance. However, this conclusion does not hold for the 76 country sample of HLR where the $z$ statistic is 6.87, easily rejecting the null hypothesis of an unbiased sample.

The algorithm adopted by FIFA for calculating its world ranking points by country has been criticized for generating apparently counterintuitive orderings.[3] For example, just prior to the 1998 World Cup, FIFA ranked Egypt in equal 17th place with the same points as France, despite the fact that the French team was clearly superior in the eyes of most experts at the time. One possible explanation for such anomalies is the fact that the FIFA rankings include the results of friendly matches. This is a cause for concern because these games lack the performance incentives of competitive matches. Their outcomes, therefore, are more likely to be misleading as a measure of footballing strength. Indeed, Torgler (2004) reports that FIFA world ranking did not play an important role in predicting the outcome of matches in the 2002 World Cup. But because he did not test alternative ranking measures, it is not possible to assert a priori whether this reflects a failure of the ranking system itself, a problem with his equation specification, or simply the vagaries of one-off competitive matches at the highest level.

Alternative unofficial indices to FIFA's measure are calculated by enthusiasts and published on the Internet. One of the most carefully constructed is the Elephant ranking, which is based on results in the World Cup and continental championships, including qualifiers, and specifically excludes friendly games.[4] In what follows, results are presented using both FIFA's ranking points and the Elephant alternative as the dependent variable. Note that all of the regressions are estimated using (FIFA or Elephant) calculated ranking points rather than the actual ranks themselves as the dependent variable. This is because the use of pure ranks would result in the loss of information with regard to the distance between countries in terms of relative strength. Notwithstanding some significant discrepancies with respect to the position of individual countries, the FIFA and Elephant measures are highly correlated for the 76 countries in HLR's sample.[5]

Given the paucity of readily available soccer-specific inputs by country such as expenditure on football or numbers of players or teams or leagues, HLR's study was constrained to use explanatory variables unrelated to football. There are four explanatory variables in their preferred equation. The first is gross national product per capita, $GNP_i$, collected from the World Bank database. The variable is a proxy for the effect of private and public football funding. Higher levels of GNP per capita are anticipated to have a positive but decreasing effect on international football success across countries below a threshold income level, above which the relationship becomes negative due to the relatively low income elasticity of demand for participation in football. A quadratic specification in GNP per capita is employed to capture this postulated inverted U-shaped relationship with ranking points.[6] This is consistent with the approach of Johnson and Ali (2000, 2002) who adopted the same nonlinear specification in per capita income for their Olympic medal count equation.[7]

A temperature variable $(TEMP - 14)_i^2$ is introduced to capture the effect of climate on football performance. The authors argue that temperate countries with an average temperature around 14°C (57.2°F) are conducive to outdoor activities and this promotes sporting success, whereas deviations in either direction are detrimental. The climate variable is measured simply using the squared deviation of average annual temperatures from 14°C for the capital city of each country.

Noting the footballing success of some Luso-Hispanic cultures, HLR specify a dummy variable, $LATIN_i$, set to one for all Spanish and Portuguese speaking countries. This is designed to capture the special cultural factors that promote football in such nations, although the authors are unable to identify precisely what such elements might be. In the preferred specification, $LATIN_i$ is interacted with $POP_i$, country *i*'s share of the world population, reflecting the fact that the marginal effect of a larger population on football performance depends on the cultural predisposition to participate in the sport. Neither $LATIN_i$ nor $POP_i$ were statistically significant when included individually as explanatory variables.

Finally, for countries that have acted as a host nation for the World Cup (finals) competition, a dummy variable is set to one. In studies of international tournament outcomes, host status is typically considered to confer enormous benefits arising from the strength of spectator support, familiarity with the local conditions, and the intensity of public expectations. Torgler (2004), for example, reports a strong favorable effect of host status on the relative success of Japan and South Korea in the 2002 FIFA World Cup.[8]

In the context of this study, however, because interest is not in the outcome of a particular event or tournament, the host variable is interpreted somewhat differently by HLR. The condition of having hosted the World Cup is taken to measure cultural affinity toward soccer because hosting typically reflects a strong footballing tradition and widespread public support for the event. Naturally, cultural affinity is expected to have a positive influence on international performance. However, interpreting hosting status as a proxy for football tradition is not altogether satisfactory. The set of hosts (prior to 2001) includes, for example, the United States, which has a lower cultural affinity with soccer than many countries that have never acted as World Cup hosts such as Russia, the Netherlands, Turkey, Hungary, and the Czech Republic.

More generally, the $HOST_i$ dummy regressor may be jointly determined with international soccer performance. In other words, the choice of host itself could be partly endogenous to a country's football world ranking, resulting in biased and inconsistent estimates. Although the $HOST_i$ variable relates to tournaments prior to the FIFA world rankings for 2001 that define the dependent variable, it cannot be considered independent of these rankings due to the high degree of serial correlation in the rankings over time. Given the potential endogeneity of $HOST_i$, the natural approach is to search for valid instruments, that is, exogenous sources of variation in hosting that are unrelated to (omitted) factors that affect international football performance. As it is difficult to think of any such candidate instruments, an alternative strategy is either to reestimate the equation with the host variable omitted or to find a better variable to model football tradition. We adopt the latter approach.

Table 1 provides descriptive statistics for the variables used in the regression analysis for the extended data set of 176 countries.

## REGRESSION RESULTS

HLR's results are reported in column 1 of Table 2. Note that this and subsequent equations are estimated by OLS. Column 2 replicates the results using the alternative Elephant ranking index $ELE_i$. Given that the scale of this ranking measure differs from FIFA's, the coefficient estimates are not directly comparable across the

TABLE 2: Regression Results

| *Dependent Variable* | $Y_i$ *(1)* | $ELE_i$ *(2)* | $Y_i$ *(3)* | $Y_i$ *(4)* | $Y_i$ *(5)* | $ELE_i$ *(6)* |
|---|---|---|---|---|---|---|
| Constant | 492.59 | 61.41 | 449.4 | 432.92 | 7077.0 | 567.14 |
| | (19.3) | (33.2) | (23.0) | (18.2) | (7.2) | (8.9) |
| $GNP_i$ | .01 | .001 | .01 | .01 | .008 | .0005 |
| | (2.4) | (3.3) | (2.6) | (3.0) | (2.1) | (2.0) |
| $GNP_i^2$ | $-2.45 \times 10^{-7}$ | $-2.46 \times 10^{-8}$ | $-2.58 \times 10^{-7}$ | $-3.07 \times 10^{-7}$ | $-2.80 \times 10^{-7}$ | $-1.88 \times 10^{-8}$ |
| | (1.7) | (2.3) | (1.9) | (2.3) | (2.3) | (2.4) |
| $(TEMP - 14)_i^2$ | –.49 | –.07 | –1.16 | –1.17 | –.80 | –.05 |
| | (2.0) | (3.9) | (6.5) | (6.6) | (4.6) | (4.33) |
| $HOST_i$ | 81.05 | 5.61 | 127.03 | 94.14 | 36.11 | 5.25 |
| | (1.8) | (1.7) | (2.3) | (1.7) | (.7) | (1.6) |
| $LATIN_i \times POP_i$ | 8587.46 | 370.37 | 109.64 | 85.66 | 71.93 | 2.86 |
| | (2.2) | (1.3) | (2.2) | (1.7) | (1.6) | (1.0) |
| $POP_i$ (millions) | | | | .91 | .57 | .01 |
| | | | | (2.5) | (1.7) | (0.5) |
| $POP_i^2$ | | | | $-7.54 \times 10^{-4}$ | $-5.31 \times 10^{-4}$ | $-1.36 \times 10^{-5}$ |
| | | | | (2.4) | (1.9) | (0.7) |
| $HISTORY_i$ | | | | | –3.41 | –.26 |
| | | | | | (6.8) | (8.0) |
| $REPUBLICS_i$ | | | | | –104.1 | –8.40 |
| | | | | | (2.5) | (3.0) |
| Sample size | 76 | 76 | 176 | 176 | 176 | 176 |
| $\chi_H^2$ | 3.56 | 9.72 | 8.79 | 8.34 | 12.63 | 14.5 |
| ($p$-value) | (.9) | (.3) | (.4) | (.7) | (.6) | (.4) |
| $\chi_N^2$ | 30.94 | .02 | 2.20 | 2.44 | 1.26 | 2.25 |
| ($p$-value) | (.0) | (1.0) | (.3) | (.3) | (.5) | (.3) |
| $\overline{R}^2$ | .318 | .425 | .362 | .378 | .508 | .54 |

NOTE: $\chi_H^2$ and $\chi_N^2$ are the chi-squared statistics for White's (1980) test for heteroskedasticity (no cross terms) and the Jarque and Bera (1980) test for nonnormal errors, respectively. $\overline{R}^2$ is the $R^2$ statistic adjusted for degrees of freedom. All equations are estimated by ordinary least squares.

two equations.[9] However, in terms of statistical significance, it is clear that only the coefficient of the $LATIN_i \times POP_i$ variable deteriorates somewhat, whereas the climatic and GNP variables have larger $t$-ratios. In other words, HLR's results do not appear to be too sensitive to the choice of ranking measure, although their specification explains the variation in the Elephant points better than that in FIFA's scale.

With respect to the diagnostic statistics, the null hypothesis of the normal distribution of the error terms is easily rejected for the HLR specification reported in column 1. The worrying implication of rejecting normality is the possible invalidity of statistical inferences for this equation. However, the diagnostic tests for columns 2 and 3 indicate that substituting the Elephant rank index or expanding the sample size both ameliorate the nonnormal error problem evident in the HLR study.

Column 3 reestimates HLR's equation, including data for 100 additional nations, giving a total sample size of 176 country observations. Compared to column 1, the parameter estimates for $GNP_i$ are essentially unchanged but those for the other explanatory variables differ substantially. This provides evidence for the presence of serious sample selection bias in the original estimates. The effect of the temperature variable is more than twice as powerful, the estimated coefficient of the $LATIN_i \times POP_i$ interaction dummy is considerably smaller, and that of $HOST_i$ is more than 50% larger in magnitude.

### *Population*

HLR fail to find any significant independent effect of population size on international football performance and so only specify the variable as an interaction term with the $LATIN_i$ dummy. The authors attribute this result to the fact that some of the world's most highly populated countries such as China, India, and Indonesia have enjoyed only very limited success in international soccer. Because the same pattern is also true with respect to Olympic competition, Johnson and Ali (2000, 2002) proceed by employing a quadratic specification in population, using total population as a proxy for the number of people in the competitive age range, and find a negative coefficient on the squared term. This implies a diminishing marginal contribution of an additional person to sporting success as population increases.

Our estimated equation, including a quadratic in population ($POP_i$ and $POP_i^2$, measured in millions), is reported in column 4 and shows that both variables are statistically significant at the 5% level.[10] Compared to the results in column 3, the main effect on the estimates is a modest reduction in the magnitude and significance of the $LATIN_i \times POP_i$ and $HOST_i$ variables. Furthermore, in an unreported regression, it was found that excluding the highly populated countries of India and China from the sample appreciably increases the size of the estimated population coefficients but with very little effect on other parameters.

## Football History

The unavailability of football-specific input variables is a clear weakness in an equation explaining international soccer outcomes. One factor overlooked by HLR is the importance of football history. Given a particular level of GNP per capita, population, culture, and climate, international success would be expected to be increasing in the length of time that a team has participated in international competition. This can be justified theoretically in terms of the experience benefits of learning by doing. It takes time to build up a footballing infrastructure such as domestic leagues and coaching schools. Many of the African nations, for example, entered international competition comparatively late, beginning in the 1950s. It is only in the past 15 years that countries such as Nigeria, Senegal, and Cameroon have started to perform well on the world stage and challenge the traditional European and Latin American football elite.

To measure this history variable, we use the year of the first international football match by country and label it $HISTORY_i$.[11] Of course, a long history of play does not guarantee success. The Philippines were contesting the Far Eastern games as long ago as 1913 but are ranked by FIFA in 2001 just above the lowest decile of 203 countries in 179th place. Clearly, time is only one dimension of the development of football in a country, but it has the empirical advantage of ease of measurement and ready availability. A negative sign for the history variable is anticipated because a team with a more recent year for its first international match is expected to have a lower ranking ceteris paribus. A special case is that of the former Soviet republics.[12] Although these republics had national teams prior to the Soviet era, these were dissolved after the Russian revolution in the early 1920s and replaced by a single USSR team for purposes of international competition. The comparatively early start for these republics, then, is somewhat offset by later political developments. This handicap is modeled by setting a dummy variable, $REPUBLICS_i$, to one for these cases. A negative effect of the interrupted participation of their national teams at the international level on football rankings is expected.

The results are reported in column 5. The coefficient estimates of the two new variables display the expected signs and are statistically highly significant. The penalty for starting late in international football is estimated to be 3.4 points per year on the FIFA ranking scale, all else equal. Note that the goodness of fit, controlling for degrees of freedom, has increased substantially from an adjusted $R^2$ of .318 in the original HLR equation to .508 when including the history and population variables for the full sample of 176 countries. The coefficient of $HOST_i$ is no longer significant with a $t$-ratio of .7. Because the dummy for World Cup hosts is only a weak proxy for cultural affinity, its effect loses statistical significance when the history variable, arguably a better measure of football tradition, is included in the specification.

Column 6 presents the results of reestimating the equation using Elephant ranking points as the dependent variable. Whereas the history variable and Soviet

TABLE 3: The 30 Countries With the Largest Underpredicted Values

| *Country* | *Actual FIFA Points* | *Predicted Points* | *Underprediction %* |
|---|---|---|---|
| Burkina Faso | 498 | 244 | –51 |
| Trinidad and Tobago | 600 | 299.4 | –50.1 |
| South Africa | 635 | 325.6 | –48.7 |
| Angola | 539 | 282.8 | –47.5 |
| Ivory Coast | 550 | 290.8 | –47.1 |
| Thailand | 521 | 301.2 | –42.2 |
| Cameroon | 585 | 344.5 | –41.1 |
| Guinea | 470 | 279.3 | –40.6 |
| Mali | 424 | 255.3 | –39.8 |
| Togo | 475 | 291 | –38.7 |
| Ghana | 531 | 326 | –38.6 |
| Liberia | 448 | 276.9 | –38.2 |
| Saudi Arabia | 585 | 362.4 | –38 |
| Nigeria | 555 | 344 | –38 |
| Turkmenistan | 305 | 191.7 | –37.1 |
| Namibia | 459 | 291.7 | –36.4 |
| Oman | 395 | 254.1 | –35.7 |
| Zambia | 555 | 360 | –35.1 |
| DR Congo | 499 | 329.3 | –34 |
| Tajikistan | 281 | 186.8 | –33.5 |
| Barbados | 407 | 275.5 | –32.3 |
| Congo | 455 | 308.9 | –32.1 |
| Kuwait | 483 | 330.2 | –31.6 |
| Tunisia | 611 | 418.2 | –31.6 |
| Jamaica | 557 | 385 | –30.9 |
| Morocco | 610 | 424.2 | –30.5 |
| Gabon | 442 | 307.5 | –30.4 |
| Paraguay | 706 | 492.2 | –30.3 |
| Senegal | 454 | 318.7 | –29.8 |
| St. Lucia | 275 | 193.3 | –29.7 |

NOTE: FIFA = Fédération Internationale de Football Association.

Republics dummy are robust to this change, the quadratic specification in population loses statistical significance. However, this outcome is sensitive to the inclusion of the outliers of India and China. If these countries are excluded, the population variables are significant at the 5% level.

## *Actual and Fitted Values*

Despite the additional data and variables, the estimates using the FIFA scores as the regressand in column 5 explain slightly more than half of the cross-country variation in international soccer rankings. A comparison of actual and fitted values can

TABLE 4: The 30 Countries With the Largest Overpredicted Values

| *Country* | *Actual FIFA Points* | *Predicted Points* | *Overprediction %* |
|---|---|---|---|
| Surinam | 173 | 283.1 | 63.7 |
| Samoa | 140 | 230.1 | 64.4 |
| Guinea-Bissau | 126 | 213.8 | 69.7 |
| Laos | 171 | 291.8 | 70.7 |
| Djibouti | 80 | 138.7 | 73.3 |
| Cambodia | 158 | 286.9 | 81.6 |
| Yemen | 190 | 366.6 | 92.9 |
| Nepal | 166 | 322.1 | 94.1 |
| Belize | 93 | 182.2 | 95.9 |
| Niger | 107 | 214.5 | 100.5 |
| Guyana | 105 | 221 | 110.5 |
| Vanuatu | 163 | 347.2 | 113 |
| Sao Tome | 115 | 245.3 | 113.3 |
| Central African Republic | 127 | 294.3 | 131.7 |
| Equatorial Guinea | 90 | 226.6 | 151.8 |
| Tonga | 98 | 274.8 | 180.4 |
| Bahamas | 117 | 333.7 | 185.2 |
| Macao | 116 | 338.2 | 191.6 |
| Palestine | 152 | 470 | 209.2 |
| Seychelles | 83 | 257.5 | 210.2 |
| Papua NG | 65 | 208.4 | 220.6 |
| Kyrgyzstan | 137 | 463.4 | 238.3 |
| Brunei | 64 | 252.5 | 294.6 |
| Philippines | 116 | 474.9 | 309.4 |
| Nicaragua | 71 | 334.7 | 371.4 |
| U.S. Virgin Islands | 33 | 185.7 | 462.8 |
| Pakistan | 73 | 466.1 | 538.5 |
| Mongolia | 35 | 224 | 540 |
| Puerto Rico | 52 | 373 | 617.3 |
| Bhutan | 15 | 321.8 | 2045.4 |

NOTE: FIFA = Fédération Internationale de Football Association.

provide a sense of what remains unexplained. Using the equation residuals, Table 3 lists those 30 countries with the greatest percentage underprediction of the actual FIFA ranking scores. It is notable that African countries are overrepresented, making up 19 out of the 30 countries with the largest positive residuals relative to the actual values of the dependent variable.[13] For example, Cameroon, one of the strongest emerging African teams, had a FIFA points score of 585 in January 2001 but the model predicts only 344.5 points.

By contrast, only 5 of the 30 largest overpredictions listed in Table 4 are for African countries. Instead, it is primarily countries ranked among the lowest FIFA scores, particularly Asian and Pacific nations, that dominate the more substantial negative residuals. This pattern reflects the difficulty, deriving from lack of suitable cross-country data, of modeling the comparatively low level of resources invested in producing soccer players in these nation states. It is for these countries that the absence of sufficient football-specific explanatory variables is particularly problematic.

## CONCLUSIONS

This article investigates the determinants of the relative strength of national soccer teams as measured by rankings of long-term performance. It demonstrates the existence of selection bias and nonnormal errors in the earlier study by Hoffman et al. (2002b), problems that arise from their sample selection procedure. Our contribution extends their sample by an additional 100 countries, adds new variables to capture the effect of history and population, and checks the robustness of the results relative to an alternative non-FIFA national football team ranking scale.

Inspection of equation residuals suggests that additional soccer-related variables are required to explain both the underpredicted strength in international competition of many African teams and the weakness of some Asian and Pacific countries.

Further research could usefully extend our findings through the collection of data on more cross-country football-related variables that are not readily available in standard sources, such as the number of professional players or teams or leagues. Such inputs will presumably account for much of the unexplained variation in long-run international soccer performance.

## NOTES

1. FIFA is the acronym of the Fédération Internationale de Football Association that governs world football.

2. The authors are grateful to Hoffman, Lee, and Ramasamy for kindly providing a copy of their data.

3. The algorithm takes account, on a monthly basis, of goals scored in each match, whether a team is home or away, importance of the game, venue, and regional strength over the previous 8 years.

4. Available together with details of the algorithm at http://www.elerankings.com/frameworldfootball .htm/ (retrieved September 2003).

5. The correlation in terms of pure rankings is smaller at .745. There are also some very large individual discrepancies. For example, Senegal is ranked 83rd in the world by FIFA but 18th by the Elephant rankings in January 2001. Eighteen months later, Senegal reached the quarter finals of the World Cup.

6. Following the suggestion of a referee, as an alternative to GNP, we also experimented with the United Nations's Human Development Index, which is available for 172 of the countries in the sample (the exceptions are Puerto Rico, U.S. Virgin Islands, Liberia, and Tahiti). In principle, as an overall measure of development, this index may represent a superior proxy for the ability of a nation to invest resources in soccer. However, this variable failed to achieve statistical significance in our regressions whether specified in linear or quadratic forms. We also tried including a dummy variable for less developed countries, but again this was not significant at conventional levels.

7. By contrast, in their study of Olympic medal outcomes, Bernard and Busse (2004) prefer to use the log of GDP per capita in addition to the log of population size.

8. Johnson and Ali (2000, 2002) also report strong hosting effects for home nations on Olympic success, controlling for the economic, geographic, and political attributes of participants.

9. The range of the Elephant points scale is 28.67 to 87.06, compared with 15 to 821 for the FIFA scale.

10. The estimated coefficient of $POP_i^2$ is negative as in Johnson and Ali (2000, 2002). For the original nonrandom HLR sample with 76 countries, the quadratic specification in population is only statistically significant if the highly populous China and India are omitted. However, the problems of nonnormal errors and sample selection bias still remain.

11. The date of first international match data are collected from The Rec. Sport. Soccer Statistics Foundation, available at http://www.rsssf.com/ (retrieved September 2003). A referee suggested that the age of the domestic league may be a better measure of tradition effects than the date of the first international game. Unfortunately, data on the length of time that national leagues have been in operation are not readily accessible for many countries. Instead, we collected data on a closely related variable, namely the date of foundation of the national football association by country, available at FIFA's Web site, www.fifa.com. This is expected to be highly correlated with timing of the formal organization of a football league in most cases. The foundation year variable performed well in the regressions and had the expected negative sign but with slightly lower significance than the date of the first international match variable reported in the text. Although other parameters were materially unaffected, the goodness of fit was a few percentage points lower in the case of equations using the foundation variable. Full results are available from the authors on request.

12. The former Soviet republics are Armenia, Azerbaijan, Belarus, Estonia, Georgia, Kazakhstan, Kyrgyzstan, Latvia, Lithuania, Moldova, Tajikistan, Turkmenistan, Ukraine, and Uzbekistan. Russia is excluded from the dummy variable on the grounds that the continuity between the national teams of the USSR and Russia was much greater than that for the smaller republics.

13. Notice that a positive (negative) residual in which the actual value exceeds (is less than) the fitted value represents an underprediction (overprediction).

## REFERENCES


Bernard, A. B., & Busse, M. R. (2004). Who wins the Olympic games: Economic resources and medal totals. *Review of Economics and Statistics*, *86*(1), 413-417.

Hoffman, R., Lee, C. G., & Ramasamy, B. (2002a). Public policy and Olympic success. *Applied Economics Letters*, *9*, 545-548.

Hoffman, R., Lee, C. G., & Ramasamy, B. (2002b). The socioeconomic determinants of international soccer performance. *Journal of Applied Economics*, *5*(2), 253-272.

Jarque, C. M., & Bera, A. K. (1980). Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Economics Letters*, *6*, 255-259.

Johnson, D. K. N., & Ali, A. (2000, September). *Coming to play or coming to win: Participation and success at the Olympic games* (Wellesley College Department of Economics Working Paper 2000-10).

Johnson, D. K. N., & Ali, A. (2002, January). *A tale of two seasons: participation and medal counts at the summer and winter Olympic games* (Wellesley College Department of Economics Working Paper 2002-02).

Torgler, B. (2004). The economics of the Football FIFA World Cup. *Kyklos*, *57*(2), 287-300.

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica*, *48*, 817-838.



*Peter Macmillan is a lecturer in economics in the School of Economics and Finance at the University of St. Andrews.*

*Ian Smith is a senior lecturer in economics in the School of Economics and Finance at the University of St. Andrews.*