Available online at www.sciencedirect.com
ScienceDirect
International Journal of Forecasting 26 (2010) 576–588
www.elsevier.com/locate/ijforecast

Forecasting national team medal totals at the Summer Olympic Games☆

David Forrest[a,*], Ismael Sanz[b,c], J.D. Tena[d,e]

[a] University of Salford, UK
[b] Universidad Rey Juan Carlos, Spain
[c] GRIPICO, Spain
[d] Universita di Sassari, Italy
[e] Universidad Carlos III, Spain

Abstract

The paper reports the results of an exercise to forecast national team medal totals at the Beijing Olympic Games, 2008. Forecasts were released to the media before the competitions commenced. The starting point was an established statistical model based on a regression analysis of medal totals in earlier Games, with past performance and GDP among the principal covariates. However, we based our own forecasts on a model with additional regressors, including a measure of public spending on recreation. This adaptation is shown to have improved the forecasting performance. We also made subjective, judgemental adjustments before releasing our final public forecasts, and we demonstrate that this led to a further increase in accuracy. These final forecasts were successful in predicting the principal changes in medal shares relative to the 2004 Games, namely the surge in medals for China and Great Britain and the substantial fall in medals for Russia.

Keywords: Sports forecasting; Olympic Games; Medals

1. Background

Whether one considers the number of athletes and countries participating, the range of disciplines represented, or the size of television audience, the Summer Olympic Games, held every four years, is unquestionably the world's leading sports event. The Beijing Games of 2008 is generally considered to have been one of the best ever in terms of organisation, spectacle and record performances. In each country considerable attention is paid to the performance of its national team, and there was universal interest in whether the host nation, China, would continue its apparent progress, evident in recent championships, towards a super-power status in sport.

This paper reports an exercise of forecasting in advance the medal total of each national team in

☆ An earlier version of this paper was presented at the Workshop on the Economics of the Olympic Games, held at the University of Groningen, Netherlands, in May, 2008.

* Corresponding author.
E-mail address: d.k.forrest@salford.ac.uk (D. Forrest).

0169-2070/$ - see front matter © 2010 International Institute of Forecasters. Published by Elsevier B.V. All rights reserved.
doi:10.1016/j.ijforecast.2009.12.007

Beijing. The forecasts generated were placed in the public domain, by press release, on the day of the Opening Ceremony, before the competition commenced, and, while no newspaper published the forecasts in full, reference was made to the projected medal totals of various individual countries on the BBC website and during interviews on BBC (domestic and World Service) radio and French Radio 1. One of the reports is still available (May, 2009) at http://news.bbc.co.uk/today/hi/today/newsid_7549000/7549969.stm.

Most of the models of sports that are reported in the forecasting literature relate to a single sport, for example association (Goddard, 2005) or American (Sauer, Brajer, Ferris, & Marr, 1988) football. Such models are based on collecting a sample of matches and regressing the outcome on specific sporting data from previous fixtures, such as teams' goal scoring records in soccer or yards gained in American football. Sometimes models are based on power scores (Boulier & Stekler, 2003), but power scores themselves are only aggregated measures of similar team data from previous matches.

However, forecasting medal totals in the Olympic Games cannot rely on using past sporting data as the primary input to the statistical model. Too many disparate events are involved for it to be practical to derive aggregate medal predictions based on a micro model for each of the several hundred individual competitions that make up the Games. When analysing Olympic performances by national teams, as initiated by Ball (1972), the tradition has therefore been not to use sporting data but rather to use more fundamental variables that capture the economic, demographic and political factors which are relevant in determining the shape of the final medals table.

Bernard and Busse (2004) made what is probably the best known contribution to the field, on account of both the econometric sophistication of their model and the fact that it was presented in a very high profile journal. Bernard and Busse estimated a tobit regression equation (which allows for the large number of countries with no medals in a Games), employing data from 1960 to 1996 (excluding, in their preferred result, those years affected by large-scale political boycotts). Their dependent variable was the medal share of country i in year t. The key explanatory variables were population and GDP per capita (the very similar values of the coefficients estimated on these two variables led Bernard and Busse to conclude that "total GDP is the best predictor of national performance"). In addition, categorical variables were also included to represent: the host country of a Games; membership or former membership of the Soviet bloc; and 'other planned economies'. The coefficient estimates on all variables were positive and the levels of statistical significance suggested that the model had successfully identified the key structural determinants of Olympic performance.

However, the goodness-of-fit of the Bernard and Busse specification, when it included only population, GDP and categorical variables to represent hosts, Soviet bloc countries and planned economies, was relatively poor (the same was true of a very similar work that fitted a Poisson model in preference to a tobit; see Lui & Suen, 2008). The problem was that, for unknown reasons, perhaps related to the sporting culture, some countries consistently performed much better or worse than would be expected from the size of their economies. For forecasting purposes, a lagged dependent variable was therefore added to the specification. With a country's previous medal performance taken into account, the potential of the model to be used for forecasting was considerably enhanced. Bernard and Busse demonstrated this by releasing to the media forecasts of medal totals for the Sydney Games in 2000. They commented that the predictions turned out to perform "quite well" according to formal distance metrics.

In this paper, we propose to adapt the (moderately) successful Bernard-Busse forecasting framework which was used to predict medal totals at the 2000 Games (an exercise reported in their 2004 paper) to include two new covariates, one of which attempts to capture the influence of the level of public expenditure on sport in each country and one of which tests whether future hosts of the Games have such a great incentive to raise their performance standards that this is already reflected in their achievements in the current Olympiad. Thus, we introduce the effects of national policy into the model in an attempt to improve the forecasting performance. In Section 2 we describe and report estimation results for this augmented model, using data from the four Olympics between 1992 and 2004. In Section 3, we describe how the results from this statistical model (which relates to the *share* of medals by country) were used to predict the *number* of medals by country at the Games in Beijing, 2008. We also explain the limited judgemental adjustments made in an attempt to further refine our predictions. In Section 4, we compare the forecasting performances of: a baseline model similar to that of Bernard and Busse (2004); our own statistical model; and our own statistical model following the incorporation of judgemental adjustments. The final two sections contain an overview of the success of the exercise and discuss its practical purpose.

2. The underlying statistical model

We obtained the data on medal totals by country at each Olympic Games from the website of the International Olympic Committee (www.ioc.com). Because our analysis was to extend over a number of Olympiads, and because the number of medals varies over time as both sports and the events within them are added to or removed from the programme, we followed Bernard and Busse (2004) and modelled each country's share of medals rather than the actual number. Like them, we expected a country's level of GDP to be a predictor of the share of medals, and, for consistency across the equation, we represented the relevant covariate as the share of GDP (relative to all other countries in the sample). Again like Bernard and Busse, we included a lagged dependent variable, and also categorical variables to represent countries which had been members of the Soviet bloc, other countries which currently operate as 'planned' economies (China), and the host country of a Games (Spain, USA, Australia and Greece in 1992, 1996, 2000 and 2004 respectively).

Bernard and Busse (2004) themselves described their specification as sparse. As was mentioned above, we decided to add two covariates which we suspected might be relevant.

The first 'new' variable (*antehost*) is a categorical variable which is set equal to one in observations relating to the country which is due to host the next Games, four years after the current event. Countries are awarded the right to stage the Games seven years before they take place. We reasoned that a country will wish to do particularly well in terms of medals when it stages the Games, and the government and governing bodies will begin preparing straight away, acting to improve standards by, for example, hiring eminent foreign coaches or allocating more of the sports budget to elite, as opposed to recreational, athletics programmes. Such measures should be reflected to some extent in the country's performance at the Games held between the time when a country is awarded and when it actually hosts the Games.

The second variable we added seeks to capture the influence of public subsidies to sport on national team performance. It was not possible to obtain the size of the sports budget for each country and each year in the sample. However, we identified a potentially adequate proxy. The United Nations data on public expenditure by country breaks down spending into nine categories defined by the Classification of Functions of Government (United Nations, 2000), and one of these, "recreational, cultural and religious affairs", includes sport. Of course, the data here include a large volume of spending that is unrelated to sport. On the other hand, EUROSTAT does disaggregate the expenditure under this heading for five European Union countries from 1990 onwards, and there proves to be remarkable level of consistency across space and time in the proportion (close to 35%) of this budget which is devoted to sport. This encouraged us to suspect that the correlation between expenditure on the broad 'recreation' category and expenditure on sport itself might be high enough, at least across western, developed countries, that the figures in the UN data would allow our statistical model to control for different levels of expenditure on sport. Unfortunately, there is no evidence on how highly correlated total 'recreation' and sports spending are across all countries, so that ultimately the adequacy of using the UN data can only be judged based on whether their inclusion in the model improves the forecasting performance.

The data were collected from various issues of *Government Financial Statistics*, published by the International Monetary Fund. They were not available until 1990, which explains why our model of the generation of medals was estimated across the Games from 1992 onward (an incidental advantage of this was that there were no distortions arising from political boycotts in this period). The variable (*recspendshare*) that we constructed from this data, expressed in such terms that it was consistent with the way in which the other variables in the model were defined, was the share of aggregate public expenditure on recreation across the sample (in the four years up to and including the Olympic year — three for the 1992 Games) accounted for by the subject country.

Our final model, to be estimated as a tobit regression, was therefore:

$$medalshare_{i,t} = f(medalshare_{i,t-4}, gdpshare_{i,t}, host_{i,t}, sovietbloc_{i,t}, planned_{i,t}, antehost_{i,t}, recspendshare_{i,t}), \quad (1)$$

where i is an index of countries and t signifies the year of the Games.

Earlier on we estimated a version of this model which also included the population size (as a share of the population of all countries in the sample), but the coefficient estimate proved to be insignificant even at a generous level of significance. It was therefore omitted from the final model as shown in Eq. (1). Bernard and Busse (2004) included the population size and GDP per unit of population together, but found that they attracted very similar coefficient estimates, implying that the level of GDP alone would provide equivalent results. The redundancy of population in the presence of GDP, as here, is therefore consistent with their findings.

We did not include a constant term in the model, as this would unreasonably imply that, for example, a microstate with a negligible share of the world GDP and recreational spending would still be expected to claim some share of the medals available. Furthermore, when we experimented with the inclusion of an intercept term, its value was effectively zero (-0.002, t-statistic $= 0.33$). Bernard and Busse (2004) also excluded the constant term from their preferred specification (with a lagged dependent variable).

Eq. (1) was estimated using all observations from all Games between 1992 and 2004 for all countries for which complete data on the regressors were available. Note that although public expenditure on recreation is now recorded for most countries (and therefore most of the countries which took part in the 2008 Games could be included in the forecasting exercise), data from the earlier years were scarcer, which accounts for the relatively few observations included in the modelling of outcomes between 1992 and 2004. Of course, there is a risk here of bias in the estimation of the model, since the countries which submitted data in the early years probably formed a non-random sample of countries to which the model was to be applied for forecasting purposes. It was not possible to assess the degree or direction of any bias, but, if serious, presumably it would subsequently be reflected in low quality forecasts.

Table 1
Summary descriptive statistics ($n = 196$).

	Mean	Median	Standard deviation
Medalshare	0.0110	0.0027	0.2199
Gdpshare	0.0076	0.0010	0.0234
Host	0.0069	0.0000	0.0828
Sovietbloc	0.0759	0.0000	0.2650
Planned	0.0069	0.0000	0.0828
Recspendshare	0.0121	0.0012	0.0718

Table 1 presents summary descriptive statistics for the covariates included in the model. The results from estimating Eq. (1), our full model in which we augment the Bernard and Busse approach by including the two additional covariates, are displayed in column (1) of Table 2.

The results are dominated by the lagged dependent variable, confirming the tendency of individual countries to persist in a good (or bad) performance from Games to Games. As in previous studies, GDP is a positive and highly significant predictor of success. Again as in earlier work, both host nations and ex-Soviet bloc countries outperform what might be expected given their income levels. Of the two new covariates we introduced, *antehost* attracts a positive and strongly significant coefficient: hosts of an Olympic Games appear to raise their level of performance as much at the preceding Games as when they stage the festival itself. The coefficient estimate on *recspendshare* is also positive, as anticipated, and is significant at a level of 0.057. It is, of course, encouraging that the influence of sports spending on international success can be detected even when it has to be proxied by data for a much more widely defined category of government expenditure.

We used the results reported in column (1) of Table 2 as the basis for our initial forecasts of medals won at Beijing. Column (2) of the same table reports the results for a specification which includes only

Table 2
Tobit regression results for the share of medals, estimated over the period 1992–2004.

	(1) Full (augmented) model	(2) Benchmark model (full sample)	(3) Benchmark model (restricted sample)
$Medalshare_{,t-4}$	0.737	0.776	0.778
	(21.95)	(26.43)	(25.43)
Gdpshare	0.082	0.095	0.095
	(3.17)	(3.91)	(3.75)
Host	0.011	0.009	0.009
	(2.47)	(2.20)	(2.10)
Sovietbloc	0.003	0.003	0.003
	(2.24)	(1.90)	(1.80)
Planned	0.007	0.005	0.004
	(1.48)	(1.02)	(0.96)
Antehost	0.015		
	(3.14)		
Recspendshare	0.047		
	(1.92)		
Number of observations	196	263	196

Note: t-statistics are in parentheses.

covariates which were consistent with the Bernard-Busse model. As a comparison, we employed this to produce a benchmark set of forecasts for the Beijing Games. Note that the sample size was larger for the column (2) exercise than for the column (1) exercise because of the missing data on government spending in the latter case.

Column (3) of the table reports the results of estimating the Bernard-Busse model using only the observations available when estimating the model represented in column (1). The coefficient estimates prove to be remarkably similar in columns (2) and (3), which suggests that the differences in coefficient estimates following the introduction of the extra variables *antehost* and *recspendshare* to the Bernard-Busse model can be attributed to changes in the specification rather than to the unavoidable employment of a more restricted sample.

3. Forecasting medal totals using the statistical model

Once estimated, the model in Eq. (1) was used to generate fitted out-of-sample values that were then applied to the question of how many medals a country would be expected to win in 2008. For some covariates, 2008 data were of course not yet available; in these cases, latest published estimates were used instead.

One complication was that the model predicts the proportion of medals a country will win relative to other countries included in the exercise. However, predictions could not be made for some of the countries, because they did not declare recreational spending, for example; but that of course does not stop them winning medals.

At Beijing, 929 medals were planned to be awarded (the actual number can vary slightly, for instance if a tie occurs for third place in some events). We assumed that the countries for which we were able to make forecasts (Cuba was an important omission) would collectively win the same proportion of all medals available as they had in the preceding Games at Athens. This meant that our forecasts were based on the expectation that the 127 countries included in the forecasting exercise would secure 893 of the 929 medals available in Beijing (in the event, they actually won 900). To obtain forecast numbers of medals by country, we therefore multiplied the fitted values from the statistical model by 893. The same procedure was used to generate forecasts from the benchmark model based on the more restricted set of covariates suggested by Bernard and Busse (2004). Below, we compare the performances of the forecasts from the two models.

There is also a third set of forecasts. This was the set of forecasts we issued to the media in a press release before events began in Beijing. It was based on our statistical model but also incorporated judgemental adjustments that were made post modelling, and represented the final and public output of our forecasting experiment.

The first adjustment we made to reach this final set of forecasts related to the expected performances of the countries in our exercise which had been a part of the former Soviet bloc (Belarus, Bulgaria, Croatia, the Czech Republic, Estonia, Georgia, Hungary, Kazakhstan, the Kyrgyz Republic, Latvia, Lithuania, Moldova, Poland, Romania, Russia, the Slovak Republic, Slovenia, Tajikistan). It is well known that the Soviet Union and its then satellites overperformed considerably at the Olympic Games throughout the Cold War period. The results from our model, which are derived entirely from data from Games postdating the end of the Soviet era, confirmed that this continued even after the fall of communism. This is not likely to be *only* because more resources were allocated to securing international prestige: our model controls for public spending on sport, but *sovietbloc* still retains statistical significance. It might be speculated that the reason for this is that communist societies were able to be more ruthless in their pursuit of medals than western societies, for example in matters such as singling out future athletes at an early age and subjecting them to rigorous specialist training. However, whatever the reason for their past international success, the maturing of democracies in Eastern Europe is likely to lead to convergence with Western Europe in sport as in other areas of society. It might therefore be expected that the extra medals attributable to the old way of doing things will fade away over time. However, the period over which our statistical model was estimated is too short for it to have been practical to detect, with statistical precision, any long-term decline in the influence of the old political system. We therefore decided to make a judgemental adjustment by imposing a coefficient of 0.002 instead of 0.003 on the *sovietblock* variable, thus reducing by a third its expected impact on numbers of medals. We could of course have imposed a more substantial reduction, say to 0.001, but this would have implicitly suggested that we expected the *sovietblock* effects to disappear entirely by the time of the 2012 London Games. This we were reluctant to believe because of the substantial degree of inertia that appears to be present in the process of generating medals. For example, some features of the structure of sport under a previous political system might be expected to survive for a time, and thus, even now, some athletes still in active competition will have trained in a structure influenced by the mores of the Soviet era.

The second adjustment we made related to China, for which the predicted number of medals had already been raised through the influence of two categorical variables, *planned* and *host*. However, it is of course unusual for a host nation also to have a planned economy. It is true that the Soviet Union hosted in 1980, but the political boycott by the USA and others so distorted the final medal totals that the Games of that year provided no reliable guidance on whether the two effects, those from *planned* and *host*, might be compounded. In any case, we anticipated that there was indeed likely to be an interaction between the two variables. Perhaps one reason why planned economies have overperformed at Olympic Games in the past is that their leaders can direct the country's resources in ways which are unavailable in democracies, but which enable them to satisfy their own personal demand for international prestige. Prestige is even more at stake if a country is the host, so one might expect even more effort to ensure home success than from a western host. We decided to build in an interaction effect by doubling the independent effects from the *host* (0.011) and *planned* (0.007) variables shown in Table 2. This meant that, compared with the statistical model, and using exact rather than rounded coefficient estimates, we added 1.735% to the forecast medal share for China, increasing its forecast number of medals by fifteen.

4. Results from the Beijing forecasting exercise

In this section we compare the performances of our three forecasting models. We refer to our application of the Bernard-Busse (2004) model as the *baseline model*, to our own extension of it (including the two additional covariates) as the *augmented model*, and to the final adaptation, also including the two subjective adjustments, as the *adjusted model*.

Although their regression model was estimated over all Olympic countries, Bernard and Busse (2004) issued forecasts for the 2000 Games for only 36 countries, those which had won the most medals in 1996. Our exercise yielded forecasts for a much wider group of countries. However, to enable like-for-like comparisons with the degree of success they reported, we initially refer to a sub-set of our list of countries, namely the top 35 medal winners of the preceding 2004 Athens Games. We have only 35 cases rather than 36 because the information on Cuba was incomplete, which prevented our forecasting its performance in 2008.

For their 2000 forecasts, Bernard and Busse reported a mean absolute error of 4.3 medals, and noted that 35 of the 36 countries recorded a number of medals which was within two standard errors of the number predicted. This they viewed as a moderate success.

Our *baseline model* includes similar covariates to theirs, but was estimated over a later (though overlapping) data set. The mean absolute error in the resulting 2008 forecasts (for 35 countries) was 5.62, substantially larger than the equivalent measure reported for 2000. Only 24 of the 35 individual medal totals were within two standard errors of the corresponding predicted value.

On the face of it, imitating the Bernard-Busse approach shows that their model is now working much less well for forecasting. However, a closer examination reveals that, in an arithmetic sense, the deterioration in performance illustrated by the increase in the mean absolute error is, to an important extent, accounted for by the failure of the model to predict the dramatic increase in the number of medals won by two countries, China and Great Britain. In 2000, the model underpredicted for these countries by eight and ten medals respectively, but in 2008 the underpredictions from the same model here were 29 and 20. Had these errors been of the same scale as those for 2000, the overall mean absolute error would have been significantly closer to that of 2000. Therefore, the covariates proposed by Bernard and Busse (2004) should not necessarily be regarded as having a reduced relevance now.

Our *augmented model* recorded a mean absolute error of 4.71, while our *adjusted model* (which incorporated the judgemental adjustments) had a mean absolute error of 4.33. In each case, 27 of the 35 country predictions were within two standard errors of the forecast.

Both adding the new covariates and making the judgemental changes to the forecasts therefore appear to have helped our final set of forecasts to achieve the same accuracy (as measured by the mean absolute error) as the forecasts for 2000. We note, on the other hand, that the model produces more large errors than in 2000 (though the converse of this, of course, is that more individual country forecasts were very close).

We now consider the three sets of forecasts for the full set of 127 countries for which we were able to make forecasts (these are provided in an Appendix at the end of the paper). Note that 127 is a much larger number of countries than has been included in the modelling of past medal performances, due to a recent sharp increase in the coverage of the UN data on 'recreational spending'.

The same pattern emerges here as when considering only the leading countries. The mean absolute error was 2.04 for the *baseline model*, 1.89 for the *augmented model*, and 1.72 for the *adjusted model*. Of the 127 countries considered, the numbers for which forecasts were within two standard errors of the prediction were 116, 117 and 117, respectively. Thus, including the new covariates improved our predictions, and making the judgemental adjustments improved them further.

To test more formally whether our changes to the Bernard-Busse forecasting framework improved the performance, we used the encompassing test proposed by Clark and McCracken (2001), on account of its superiority over other standard procedures when comparing the forecasting accuracies of *nested* models. This test was used to compare the procedures in the *augmented* and *adjusted models* with that of the *baseline model*. In each case, the test statistic exceeded the critical value reported by Clark and McCracken, which is consistent with gains in forecasting performance.

5. Overview and evaluation

As an aid to discussion, Table 3 presents data for the fifteen leading medal-winning countries at Beijing, 2008 (thirteen of these had also been in the top fifteen positions in 2004; Romania and Cuba failed to retain

Table 3
Leading medal winners at the 2008 Games.

	Medals in 2004	Medals in 2008	Adjusted model forecast	Augmented model forecast	Baseline model forecast
USA	103	110	102	102	101
China	63	100	90	75	71
Russia	92	72	74	75	78
Great Britain	30	47	44	44	27
Australia	49	46	39	39	40
Germany	48	41	43	43	43
France	33	40	32	32	29
South Korea	30	31	25	25	25
Italy	32	28	28	28	28
Ukraine	23	27	21	21	21
Japan	37	25	35	35	36
Belarus	15	19	13	14	14
Spain	19	18	18	18	17
Canada	12	18	13	13	11
Netherlands	22	16	19	19	18

this status, and were replaced in the Beijing final standings by Belarus and Canada). The table shows the countries' final medal tallies for both 2004 and 2008, as well as our final predictions (rounded to the nearest integer) and the predictions we would have made had we employed only the *baseline model*.

Our forecasts were successful in predicting thirteen of the countries that filled the fifteen leading positions in the medals tally. Within the fifteen, the forecasts were exactly correct for only Italy and Spain. However, the model was broadly successful in picking out the rank order (the Spearman correlation coefficient was 0.920), and the R^2 between actual and predicted medals was 0.966. According to both of these measures, our forecasts were more successful than the forecasts based on the *baseline model*, where the corresponding indicators were 0.861 and 0.881.

An inspection of Table 3 confirms that the greatest changes between the 2004 and 2008 Games were the substantial gains on the parts of China and Great Britain, at the cost of Russia. Our forecasts successfully predicted these shifts very closely (except for the case of China, whose improvement in performance was even greater than we expected). By contrast, the *baseline model* actually predicted a fall in medals for Great Britain; for China, while it did forecast a sharp (13%) rise in medals, associated of course with the country's rapid increase in *gdp share*, this was nothing like the gain that China actually made; it also understated the size of the setback for Russia.

Augmenting and adapting the Bernard-Busse approach therefore enabled us to be much more successful than we would otherwise have been in anticipating the headline shifts in power that emerged from the Games. Clearly, one element in this was our identification of a strong tendency in past Games for future hosts to perform well. When fed into the forecasting equation for 2008, this made our forecast for Great Britain very close. It may be noted that the contribution of *recspendshare* to the accuracy of this particular forecast was only marginal, because the share of GDP devoted to public spending on recreation increased only slowly between 2001–2004 (0.38%) and 2005–2008 (0.40%) (EUROSTAT: General Government by function). The share for the United Kingdom is well above that for Germany (0.23%) and Italy (0.30%), but similar to that recorded for Spain and Sweden, for example.

By contrast with our success in the case of Great Britain, our set of forecasts, like those produced from the *baseline model*, underestimated the strength shown by Australia and Canada; and neither set of forecasts anticipated the sharp fall in the number of medals won by Japan.

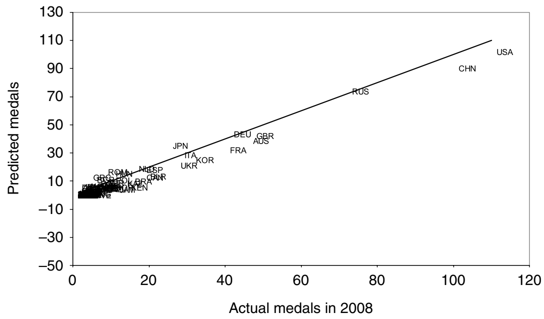

Fig. 1. Forecast and actual medal totals for our public forecasts (the adjusted model).

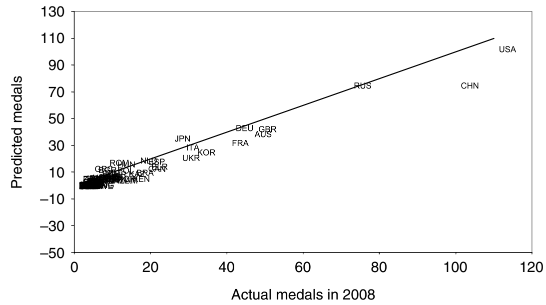

Fig. 2. Forecast and actual medal totals for the augmented model (Table 2, column 1).

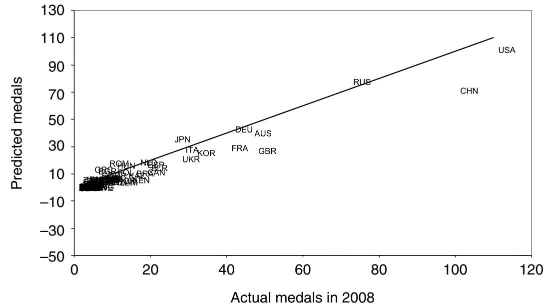

Fig. 3. Forecast and actual medal totals for the baseline model (Table 2, column 2).

At the other extreme from the elite nations, our *adjusted model* forecasts (rounded to the nearest integer) suggested that 60 countries would fail to win a medal. In fact, 47 of these duly recorded zero medals. However, one of them, Zimbabwe, achieved four medals, thanks to the efforts of a single swimmer, based, as it happens, in the USA. Of course, at the bottom end of the medals table, proportionately large errors are likely to occur because of the random availability of such individual talent.

Considering now all 127 countries for which forecasts were made, the goodness-of-fit between the actual and forecast medal numbers for our (final) *adjusted model* is illustrated by Fig. 1 (where countries represented by observations to the right of the 45° line performed better than predicted and those to the left performed worse than predicted). The model forecasts were very close for most countries. Of the 127 countries for which forecasts were made, 99 recorded medal totals within two of the forecast. The R^2 between the *adjusted model* forecasts and the actual medal totals was 0.970.

For comparison, Figs. 2 and 3 illustrate the goodness-of-fit of the alternative forecasts from the *augmented model* and the *baseline model*. The visual impression that the goodness-of-fit is lower in Fig. 2 than in Fig. 1 and lower in Fig. 3 than in Fig. 2 is confirmed by the R^2 values of 0.957 for the *augmented model* and 0.939 for the *baseline model*. Again, the pattern is consistent with gains in forecasting having been achieved *both* from modifying the Bernard-Busse framework to include policy related variables *and* from making limited judgemental adjustments to the new statistical model to take into account factors specific to China and the countries of the former Soviet bloc.

6. Closing remarks

Our goal in this exercise was to continue the tradition of attempting to forecast Olympic medal totals without employing any sporting data other than the medal count from the preceding Games. Earlier contributions to the literature had drawn attention to the explanatory power of the structural determinants of the numbers of medals, but Bernard and Busse (2004) were the first to demonstrate the forecasting capability of models based on covariates capturing countries' economic status and political systems. We have shown that, in the case of the 2008 Games, their style of model again predicted the pattern of medals awarded well, but also that when we added two new covariates, the forecasting power was increased further. One of these new covariates, a measure of public spending in the broad area of recreation, explains (in a general sense) the persistent tendency of countries such as Australia, New Zealand and The Netherlands, which are high spenders on recreation, to perform more strongly at the Olympic Games than the size of their economy suggests that they should. Judgemental adjustments to the output of statistical models are, of course, controversial. The statistical model assigns weights to information in an objective fashion, but the modeller, making changes on the basis of extra information, may over- or under-estimate the importance of this information, and could therefore worsen the forecast. However, in our case, the adjustments produced further improvements in the accuracy of the forecasts to which we committed ourselves before the Games.

Could a superior set of forecasts be derived for the 2012 Games? It would be desirable to consider alternatives to the tobit specification, for example by evaluating the out-of-sample performance of the Poisson count model employed by Lui and Suen (2008). However, the most striking failures in the forecasting exercise appear to be likely to be repeated unless further sporting information is taken into account. For example, the medal score for Jamaica was underpredicted, as was that for China; but both errors might have been avoided if the strong showings of those countries in track events at world championships held since the previous Olympics had been taken into account. Of course, judgemental adjustments on the basis of the myriad of sporting data available might actually worsen the forecasts, since it is unclear how much weight should be attached to each piece of information. However, given the large number of Olympic medals awarded for track and field events and for swimming, it might be appropriate to experiment with including covariates in the statistical model representing countries' performances at preceding world championships in these particular sports.

What purpose does all this serve? Of course, there is an immense media appetite for forecasts of high profile events. Efforts to satisfy that demand by statistical modelling can at least have the spin-off effect of aiding a more effective understanding of the structural determinants of national sporting success; and, by adding public spending to the model, we include something over which governments actually have influence. There are also some practical uses to which such forecasts can be put. They provide a framework for the objective benchmarking of national team performance, making sports governing bodies more accountable for any public funding they receive. Forecasts based on structural characteristics also have the advantage that, since the relevant covariates tend to change only slowly, they can reasonably be made a few years ahead. This means that they could be used to market the sponsorship of preparations by a national squad: potential partners could be given a realistic assessment of how successful a specific sporting enterprise on offer was.

There is every reason to expect that there will again be a strong demand for objective medal forecasts for the 2012 Games, perhaps especially in Great Britain: as host, it is likely to be particularly anxious to improve its showing even further than in 2008. Our long-range forecast for Britain is that the prospects for it doing so are good. Our modelling of past Games indicates that the future host would have been expected to win an extra 13 medals in Beijing relative to the number that would have been generated by reference to the other covariates alone. This accounted for most of the actual gain in medals made by Great Britain relative to the 2004 Games. Through the momentum effects captured by the lagged dependent variable, some of this gain in 2008 (amounting to ten medals) would be expected to be retained at the 2012 Games when it is the actual host, and there should then be a further gain (ten more medals) attributable to the hosting effect itself. Thus, we cautiously predict a successful 2012 for the British team, though of course, the forecasts will need to be reviewed over time to capture the effects of any variation in GDP and other covariates, which may potentially be very large in a new era of economic turmoil.

Appendix

See Table A.1.

Table A.1
Model forecasts for 127 countries at the 2008 summer Olympics.

	Medals in 2008	Adjusted model forecast	Augmented model forecast	Baseline model forecast
Afghanistan	1	0	0	0
Albania	0	0	0	0
Algeria	2	0	0	0
Argentina	6	5	5	5
Australia	46	39	39	40
Austria	3	6	6	6
Azerbaijan	7	6	7	6
Bahamas, The	2	2	2	2
Bahrain	1	0	0	0
Bangladesh	0	0	0	0
Barbados	0	0	0	0
Belarus	19	13	14	14
Belgium	2	3	3	3
Belize	0	0	0	0
Bhutan	0	0	0	0
Bolivia	0	0	0	0
Botswana	0	0	0	0
Brazil	15	10	10	11
Bulgaria	5	11	12	12
Burkina Faso	0	0	0	0
Burundi	0	0	0	0
Cameroon	1	1	1	1
Canada	18	13	13	11
Chile	1	3	3	3
China	100	90	75	71
Colombia	2	2	2	2
Congo, Dem. Rep.	0	0	0	0
Costa Rica	0	0	0	0
Croatia	5	4	4	4
Cyprus	0	0	0	0
Czech Republic	6	8	9	9
Denmark	7	7	7	7
Dominican Republic	2	1	1	1
Ecuador	1	0	0	0
Egypt, Arab Rep.	1	5	5	5
El Salvador	0	0	0	0
Estonia	2	4	5	5
Ethiopia	7	5	5	6
Fiji	0	0	0	0
Finland	4	2	2	2
France	40	32	32	29
Gambia, The	0	0	0	0
Georgia	6	5	6	5
Germany	41	43	43	43
Ghana	0	0	0	0
Great Britain	47	42	42	27
Greece	4	13	13	13
Grenada	0	0	0	0
Guatemala	0	0	0	0
Hungary	10	15	16	16
Iceland	1	0	0	0
Indonesia	5	4	4	4

Table A.1 (*continued*)

	Medals in 2008	Adjusted model forecast	Augmented model forecast	Baseline model forecast
Iran, Islamic Rep.	2	6	6	6
Ireland	3	0	0	0
Israel	1	2	2	2
Italy	28	28	28	28
Jamaica	11	4	4	4
Japan	25	35	35	36
Jordan	0	0	0	0
Kazakhstan	13	8	9	9
Kenya	14	5	5	6
Korea, Rep.	31	25	25	25
Kuwait	0	0	0	0
Kyrgyz Republic	2	2	3	2
Latvia	3	5	6	5
Lebanon	0	0	0	0
Lesotho	0	0	0	0
Lithuania	5	4	5	5
Luxembourg	0	0	0	0
Madagascar	0	0	0	0
Malaysia	1	0	0	1
Maldives	0	0	0	0
Malta	0	0	0	0
Mauritius	1	0	0	0
Mexico	3	5	5	5
Moldova	1	2	3	2
Mongolia	4	1	1	1
Morocco	2	2	2	3
Myanmar	0	0	0	0
Namibia	0	0	0	0
Nepal	0	0	0	0
Netherlands	16	19	19	18
New Zealand	9	4	4	4
Nicaragua	0	0	0	0
Norway	10	5	5	5
Oman	0	0	0	0
Pakistan	0	0	0	1
Panama	1	0	0	0
Papua New Guinea	0	0	0	0
Paraguay	0	1	1	1
Philippines	0	0	0	0
Poland	10	10	11	11
Portugal	2	3	3	3
Qatar	0	0	0	0
Romania	8	16	17	18
Russian Federation	72	74	75	78
San Marino	0	0	0	0
Seychelles	0	0	0	0
Singapore	1	0	0	0
Slovak Republic	6	6	7	7
Slovenia	5	5	6	6
South Africa	1	5	5	5
Spain	18	18	18	17
Sri Lanka	0	0	0	0
St. Kitts and Nevis	0	0	0	0

(*continued on next page*)

Table A.1 (*continued*)

	Medals in 2008	Adjusted model forecast	Augmented model forecast	Baseline model forecast
St. Vincent and the Grenadines	0	0	0	0
Sudan	1	0	0	0
Swaziland	0	0	0	0
Sweden	5	6	6	6
Switzerland	6	5	5	4
Syrian Arab Republic	0	1	1	1
Tajikistan	2	2	3	2
Thailand	4	7	7	7
Trinidad and Tobago	2	1	1	1
Tunisia	1	0	0	0
Turkey	8	8	8	9
Uganda	0	0	0	0
Ukraine	27	21	21	21
United Arab Emirates	0	1	1	1
United States	110	102	102	101
Uruguay	0	0	0	0
Vanuatu	0	0	0	0
Venezuela, RB	1	2	2	2
West Bank and Gaza	0	0	0	0
Yemen, Rep.	0	0	0	0
Zambia	0	0	0	0
Zimbabwe	4	0	0	0

References

Ball, D. W. (1972). Olympic Games competition: Structural correlates of national success. *International Journal of Comparative Sociology*, *15*, 186–200.

Bernard, A. B., & Busse, M. R. (2004). Who wins the Olympic Games: Economic resources and medal totals. *Review of Economics and Statistics*, *86*, 414–417.

Boulier, B. L., & Stekler, H. O. (2003). Predicting the outcomes of National Football League games. *International Journal of Forecasting*, *19*, 257–270.

Clark, T. E., & McCracken, M. W. (2001). Tests of equal forecast accuracy and encompassing for nested models. *Journal of Econometrics*, *105*, 85–110.

Goddard, J. (2005). Regression models for forecasting goals and match results in association football. *International Journal of Forecasting*, *21*, 331–340.

Lui, H.-K., & Suen, W. (2008). Men, money, and medals: An econometric analysis of the Olympic Games. *Pacific Economic Review*, *13*, 1–16.

Sauer, R. D., Brajer, V., Ferris, S. P., & Marr, M. W. (1988). Hold your bets: Another look at the efficiency of the market for National Football League games. *Journal of Political Economy*, *96*, 206–213.

United Nations (2000). *Classification of the functions of government (COFOG)*. M. No. 84. Statistics Division, New York.

David Forrest (Ph.D., Economics, University of Western Ontario) is Professor of Economics, and a member of the Centre for the Study of Gambling, University of Salford, UK. He is also Honorary Professor at Macao Polytechnic Institute. His research interests range widely across the fields of sport and gambling. His recent outlets include *Economic Inquiry*, *Journal of the Royal Statistical Society (Series A)* and the *European Journal of Finance*. He has authored and co-authored a number of official reports, including, most recently, reports on match fixing in sport and on the prevalence of child gambling in Great Britain.

Ismael Sanz (Ph.D., Complutense University of Madrid) is currently Professor at the Applied Economics Department of Universidad Rey Juan Carlos, Madrid. He has been a Visiting Fellow at University of California Santa Barbara, Australian National University, University of Otago, Kennedy School of Government (Harvard University) and University of Nottingham. His areas of specialisation are public finance, globalisation and sports economics. He has published articles in journals such as the *European Journal of Political Economy*, *Public Choice*, *Contemporary Economic Policy*, and *European Political Economy Review*. He has co-authored reports for the World Bank, the European Commission and the New Zealand Treasury.

J.D. Tena (Ph.D., Economics, University of Newcastle-upon-Tyne) is a lecturer at Universita di Sassari, Italy, and visiting professor at Universidad Carlos III, Spain. He also serves as a forecasting consultant for Instituto Flores de Lemus, Spain. His research focuses on applied econometrics, with an emphasis on sports economics, macroeconomic forecasting and the identification of structural VAR models. His articles have appeared in journals such as the *European Economic Review*, *European Journal of Operational Research*, *International Regional Science Review* and *Economic Modelling*.