Modelling Olympic performance

Economic briefing paper

113
US

87
China

68
Russia

54
Great Britain

Estimate of medal total at the 2012 London Olympic Games

pwc

Executive summary

As we are getting closer to the start of the 2012 Olympic Games in London, so interest is rising in the likely medal tallies of different countries. As a contribution to this debate, this paper presents analysis on the determinants of past Olympic performance and uses this to produce some benchmarks against which performance at the 2012 Olympics can be judged. This updates similar analysis we produced around the time of the 2000, 2004 and 2008 Olympics.

The following economic and political factors were found to be statistically significant in explaining the number of medals won by each country at previous Olympic Games:

- population;
- average income levels (measured by GDP per capita at PPP exchange rates);
- whether the country was previously part of the former Soviet/communist bloc (including Cuba and China); and
- whether the country is the host nation.

In general, the number of medals won increases with the population and economic wealth of the country, but less than proportionately: David can sometimes beat Goliath in the Olympic arena, although superpowers like the US, China and Russia continue to dominate at the top of the medal table.

Many countries from the former Soviet bloc continued to outperform relative to the size of their economies at the Beijing Olympics, despite it being held nearly two decades years after the fall of the Berlin Wall. This effect is fading, however, and is no longer statistically significant if recent Olympic performance is also included in the model. We can, however, see a similar effect at work in China more recently, where state support contributed greatly to their Olympic success in Beijing: sport it seems is one area where a planned economy can succeed!

Now it is no longer the host country, however, China may find it more difficult to stay ahead of the US (as it did in Beijing on gold medals although not total medals). Indeed we find that host nations generally ‘punch above their weight’ at the Olympics, which bodes well for the British team in London.

Our model suggests that the British team could win around 54 medals this time around, beating an already exceptionally good performance of 47 medals in Beijing. This would still leave Britain in fourth place in the total medal table, behind the US (113 medals), China (87) and Russia (68), but ahead of old rivals Australia (42) and Germany (41) according to the model projections. But all models are subject to margins of error and they can never take full account of the human factor of exceptional individual performances – so we will be only too pleased if the British team can beat our model projection in London this summer!

Economic briefing paper: modelling Olympic performance

With the 2012 Olympic Games in London fast approaching, there will inevitably be much speculation about how many medals each country will win. In this paper we consider, as a light-hearted (but nonetheless reasonably rigorous) contribution to the debate, how far statistical models can help to explain the number of medals won by each country in past Olympics. We published the results of a similar modelling exercise around the time of the Beijing 2008, Athens 2004 and Sydney 2000 Olympics and have now updated this analysis, taking into account also the results of other past studies in this area[1].

Key features of our model

The updated version of our model includes data on medal performance from the Olympic Games since 2000[2]. We find that, in explaining the share of the total medals awarded to each country, the following economic and political factors are statistically significant in a model that does not include past Olympic performance (see Annex for further technical details):

- population;
- average income levels (measured by GDP per capita at PPP exchange rates);
- whether the country was previously part of the former Soviet bloc (including Cuba in this case); and
- whether the country is the host nation.

For predictive purposes, we also took account of performance in the previous two Olympics, as discussed further below and in the Annex.

David vs Goliath

In the case of both population and average income levels, we found that the best fit[3] was obtained by using the logarithm of these variables as the explanatory factor, which implies that the number

[1] In particular, A.B. Bernard and M.R. Busse, 'Who Wins the Olympic Games: Economic Resources and Medal Totals', *Review of Economics and Statistics*, 2002; and D.K.N. Johnson and A. Ali, 'A Tale of Two Seasons: Participation and Medal Counts at the Summer and Winter Olympic Games', Wellesley College Working Paper 2002-02, January 2002.

[2] In earlier versions of this paper we looked back to 1988, but on reviewing the data we felt that the most information was contained in the more recent Games since 2000 when it came to setting medal targets for 2012. We have therefore focused on these last three Olympic Games in our statistical analysis. Our results are, however, broadly comparable to those of the studies quoted in the previous footnote, which do cover a longer time span.

[3] This was true for the variant of the model excluding performance in previous Olympics. When including these additional factors, we found that simply including the level of GDP in the model gave the best fit (see Annex for details). However, while such a model has greater predictive power, it is less interesting in terms of explaining how economic and political factors influenced past Olympic performance.

of Olympic medals won rises less than proportionately as population and/or income levels increase. The coefficients on the population and income variables were similar, suggesting that it is total GDP that matters most in predicting Olympic performance rather than how this splits down between population size and average income levels. The less than proportionate relationship between Olympic medal success and GDP implies that there are diminishing returns from economic size in terms of increased sporting success. There are a number of possible reasons for this:

- as discussed further below, we find clear statistical evidence that the former Soviet bloc countries significantly outperformed expectations based on their relatively low GDP levels;

- outstanding athletes from smaller countries may be able to train in wealthier countries (e.g. by attending US universities) but may continue to represent their own countries in the Olympics; for a small country, one or two gold medal performances from such athletes can make a large percentage difference to their overall points scores; this may well, for example, be a factor in the success of some Caribbean sprinters;

- while outstanding athletes in large countries like the US may be spread across a very wide range of Olympic and non-Olympic sports, there may be more of a tendency for athletes in smaller countries to specialise in a narrower range of disciplines where there is a local track record of success (e.g. long distance running for Kenya or sprinting for Jamaica); this strategy of specialisation can prove proportionately very successful in producing Olympic medals (and is perhaps analogous to the development of specialised industry 'clusters' in particular countries/regions, where a virtuous circle can then develop to generate world class performance[4]); and

- although this is difficult to prove, it could be that there is more focus on sport in some poorer countries where other life opportunities are more limited; if true, this greater motivation to participate and excel in sport may make up to some degree for inferior training facilities, at least in the early stages of a career.

Whatever the explanation, the bottom line is that size matters, but it is not everything. David can sometimes slay Goliath in the Olympic arena.

Long live the USSR – though the memory is fading

As noted above we found, in common with previous studies and our own earlier analysis, that whether a country was formerly in the Soviet bloc (or is in a communist regime like China or Cuba) was statistically significant[5], given the high political importance of sporting success in many of these countries. This shows that sport is one area where state planning and intervention can produce results[6], which still persisted in Beijing almost two decades after the fall of the Berlin wall. However, successive estimations of our model over the last four Olympics show that these effects are gradually fading for the ex-Soviet countries, except for China which still has a strong

[4] This cluster theory was first developed in detail by Michael Porter in his book, *Competitive Advantage of Nations* (1990). Perhaps he could include a review of sporting excellence clusters in the next edition?

[5] Except where past Olympic performance was included in the model, in which case this already capture this 'Soviet bloc' effect (see Annex for details).

[6] Many would argue unfair results due to the well-documented use of performance-enhancing drugs by some Soviet bloc countries before the 1990s, but this is unlikely to be such a factor in relative performance now.

state policy of promoting Olympic sport. On the other hand the strong sporting traditions created in these countries could last for some time yet, so there is still likely to be some outperformance in London by former Soviet bloc countries relative to what might have been expected based on their GDP alone.

We also found that it was worth distinguishing here between the group of ex-Soviet bloc or communist countries where a particularly high priority was given to sport (in particular, Russia, Ukraine, Poland, Bulgaria and Romania, as well as China and Cuba) and other ex-Soviet or planned economies where this was less of a priority. For unified Germany, we included a dummy variable value of 0.5 here to allow for the influence of the former East Germany.

Home country advantage: good news for Britain in London

We also found the home country effect to be significant. In practice, however, this effect will vary across countries depending on their size and the strength of their sporting traditions. It was particularly strong for China in Beijing (where its medal total rose to 100 from 63 in Athens) and was also evident in the Sydney 2000 Olympics, where Australia performed very well to win 58 medals (compared to 49 in Athens and 46 in Beijing as this effect faded). But it was somewhat less evident in Athens, where Greece only increased its medal total to 16, as compared to 13 in Sydney (and then fell back to just 4 medals in Beijing).

Overall, our model estimates suggest that home country advantage should on average boost medal share by around 2 percentage points, which might translate to around 19 extra medals for Britain in 2012. However, this needs to be tempered by the fact that Beijing was already an exceptional performance for Britain that may have seen preparations for London 2012 already starting to bear fruit in areas like cycling, rowing and sailing. So our overall model projection, as discussed further below, suggests a solid but more modest increase in Britain's medal total to 54 in London.

History matters

Finally, we found that the explanatory power[7] of the model was increased significantly by including medal shares at the previous two Games, which can be interpreted as reflecting the fact that sources of comparative advantage in sport tend to persist over time. Once this past performance is allowed for some other factors (particularly the ex-Soviet bloc effect) are no longer statistically significant as they are already captured in past performance.

Technical details of the model are discussed further in the Annex. It is not surprising that the model cannot explain all the variation in medal shares across countries as this will also be influenced by individual athletic performances, as well as by policy-related factors such as:

- **the relative level of state and corporate funding** of Olympic athletes in each country (as a % of GDP); unfortunately we do not have data to hand on this, but comparatively high levels of corporate sponsorship may help to explain why the US medal share remains so high;

[7] The explanatory power of the preferred model was reasonably high, as indicated by an adjusted R-squared of 0.96 (i.e. the model explains around 96% of the variance in medal shares between countries). If we exclude the lagged dependent variables, the explanatory power of the model drops to only around 50% (see Appendix for technical details).

- **the relative effectiveness of this funding**, which could reflect the extent to which it has been focused on building up successful sporting clusters of genuine world class, rather than being more widely dispersed across a range of different sports; it would also reflect the effectiveness of sports administrations in different countries; and

- **the relative importance given to athletics and other Olympic sports** where significant numbers of medals are at stake (e.g. swimming, cycling, sailing, shooting, amateur boxing and rowing), as opposed to other sports which are either not represented at the Olympics (e.g. golf, rugby, American football and cricket) or where relatively few medals are at stake (e.g. football and basketball). This is likely to be related to a complex mix of historical and cultural factors as, indeed, will be the importance given to sport per se in different societies.

It follows that, if a country's performance at the Olympics differs significantly from what our economic model would predict, this could have some policy implications in relation to the level and effectiveness of sports funding as compared to other countries.

Model estimates of medal targets for London 2012

If we apply the model to the latest available data for each country, we obtain estimated medal targets for London 2012 as shown in Table 1 (assuming that the total number of medals awarded is the same[8] as in Beijing to allow direct comparison with results from 2008).

These model estimates represent one possible benchmark or target against which to calibrate how well a country does at the London Olympics given its size, income levels, political history and past performance. We would note in particular that:

- as host country, all eyes will be on Britain, where as noted above our model indicates a target for London of 54 medals, better than the already very good performance in Beijing due to the significant role that home advantage seems to play in Olympic performance; this has not been true for all hosts, however, as the example of the US in 1996 showed, so this is quite a challenging target for the Great Britain team to meet - but its strategy of specialising in key sports like cycling[9], sailing and rowing should pay dividends again;

- China did well as the host nation in Beijing, topping the table on gold medals and not far behind the US on total medals won; however, despite its continued strong economic performance since 2008, there must be some question as to how far China can match its exceptionally good performance in Beijing now it no longer has the huge government push for success at its home Olympics; our model therefore suggests that China will remain in second place in the total medals table but a little further behind a resurgent US team than in 2008; but this is far from certain since there is still considerable state support for Chinese Olympic sport and many of the champions of 2008 will still be strong contenders in 2012;

[8] In practice, there could be some change in the total number of medals awarded in London, in which case the model estimates in Table 1 would need to be adjusted pro rata. But there is unlikely to be a large change.
[9] Although we understand that there will be fewer medals to win in track cycling this time.

- Russia is projected by the model to continue to perform strongly relative to the size of its economy in third place (68 medals), but it does continue to drift down the table relative to the heights of its performance in the old USSR era;

- the two countries with by far the largest populations in the world are China and India, but their past Olympic performances could be not be more different: China is very strong as noted above, while India won only 3 medals in Beijing (though this was an improvement on just one medal in both Athens and Sydney); our model can explain some of this divergence, but still suggests that India is a significant underperformer, with a model target of around 5-6 medals for London. The most plausible explanation is that, with the exception of hockey, Indian sport tends to be focused on events that are not included in the Olympics, most importantly cricket; China, by contrast, is an example as noted above of the effectiveness of state planning in sport, comparable to the former Soviet bloc countries[10];

- our model estimates suggest that larger Western European countries such as Germany, France, Italy, Spain and Netherlands might be expected to broadly match their Beijing 2008 performances in London, though they will no doubt hope to do better;

- countries where the model targets for London are below those in Beijing include Australia (still in gentle decline from the heights of Sydney in 2000) and some former Soviet bloc countries where the legacy advantages of strong state support from that pre-1991 era may be gradually fading such as Ukraine, Belarus and possibly also Cuba;

- as well as Great Britain, countries that the model suggests have the potential to do better than in Beijing include Japan, Brazil (in the run up to being the host country in 2016), Romania and Turkey; it will be interesting to see if they can improve their standings in London as the model suggests; and

- overall the model estimates suggest that the top 30 countries might be expected to win around 80% of all the medals awarded in London, which would actually be slightly below the 82-83% shares of the top 30 countries in Beijing, Athens and Sydney. This also broadly mirrors the shape of the global economy, in which the top 30 countries account for just over 80% of world GDP.

It will be interesting to see how actual medal performance in London compares to the benchmarks represented by the model estimates. We will revisit this question after the Olympic Games.

John Hawksworth
Chief Economist, PwC (UK)
tel: +44 (0)207-213-1650
e-mail: john.c.hawksworth@uk.pwc.com

[10] Chess (although not an Olympic sport) is another example of this phenomenon, as Chinese players (particularly women, where Xie Jun, Zu Chen, Xu Yuhua and Hou Yifan have all won world championship titles since 1991) have increasingly been challenging the dominance of former USSR states now that state support for chess has been greatly reduced in Russia and other ex-Soviet countries. India has also been enjoying a chess boom over the past decade, but this has been more due to younger players emulating the example of Vishy Anand (the world chess champion since 2007) than to state support for chess.

Table 1: Model estimates of London 2012 Olympics medal totals as compared to Beijing 2008 results

Country	Model estimate of medal total in London 2012	Medal total in Beijing 2008	Difference
1. US	113	110	+3
2. China	87	100	-13
3. Russia	68	73	-5
4. Great Britain	54	47	+7
5. Australia	42	46	-4
6. Germany	41	41	0
7. France	37	41	-4
8. Japan	28	25	+3
9. Italy	27	28	-1
10. South Korea	27	31	-4
11. Ukraine	21	27	-6
12. Cuba	20	24	-4
13. Spain	18	18	0
14. Netherlands	16	16	0
15. Canada	15	18	-3
16. Belarus	14	18	-4
17. Brazil	14	13	+1
18. Kenya	13	14	-1
19. Romania	11	8	+3
20. Hungary	11	10	+1
21. Jamaica	11	11	0
22. Poland	10	10	0
23. Turkey	10	8	+2
24. Kazakhstan	9	13	-4
25. Greece	8	4	+4
26. Norway	7	10	-3
27. Bulgaria	7	5	+2
28. New Zealand	7	9	-2
29. Denmark	7	7	0
30. Argentina	7	6	+1
Top 30 total medals	**761**	**792**	**-31**
Other countries	197	166	+31
Total medals	**958**	**958**	**0**

Source: PricewaterhouseCoopers model estimates

Annex: technical details of regression models

Table 2 below shows two alternative regression equations that we have estimated. The first model variant excludes past Olympic performance from the set of explanatory variables and so provides a purer indicator of the ability to explain variations between countries purely on the basis of economic and political factors. The second model variant includes performance at the previous two Olympic Games as an additional independent variable and has much higher overall explanatory power (as indicated by the respective adjusted R-squared coefficients of 0.96 for the second model, as against 0.51 for the first model). This second model therefore forms the basis for the London 2012 medal estimates quoted in Table 1 above. Since the unadjusted model estimates for medal shares in London did not add up exactly to 100%[11], a small scaling factor was applied to given the results shown in Table 1.

Table 2: Alternative model specifications (dependent variable = % medal share)

Explanatory variables	Model without past Olympic performance variables	Model with past Olympic performance variables
Constant	-0.02	0.0006
Log (population: millions)	0.0057 (6.3)	-
Log (GDP per capita at PPPs : $000s)	0.0051 (4.2)	-
Level of GDP at PPPs ($ trn)	-	0.0022 (5.2)
Ex-Soviet bloc dummy	0.012 (3.2)	Not significant
Host country dummy	0.06 (4.1)	0.02 (5.0)
Medal share in previous Olympic Games	-	0.422 (3.5)
Medal share in previous but one Olympics Games	-	0.353 (3.2)
Explanatory power (adjusted R-squared)	*0.51*	*0.96*
Standard error of model	*0.013*	*0.004*
Number of countries covered	*101*	*101*

Note: t-statistics shown in brackets for explanatory variables

Source: PwC analysis using data from 101 medal-winning countries in 2000, 2004 and 2008 Olympics, plus IMF data on population and GDP per capita at PPP exchange rates. For the London projections in Table 1 above, the 2008 Beijing Olympics results were factored into this model together with the latest IMF GDP and population estimates for 2011.

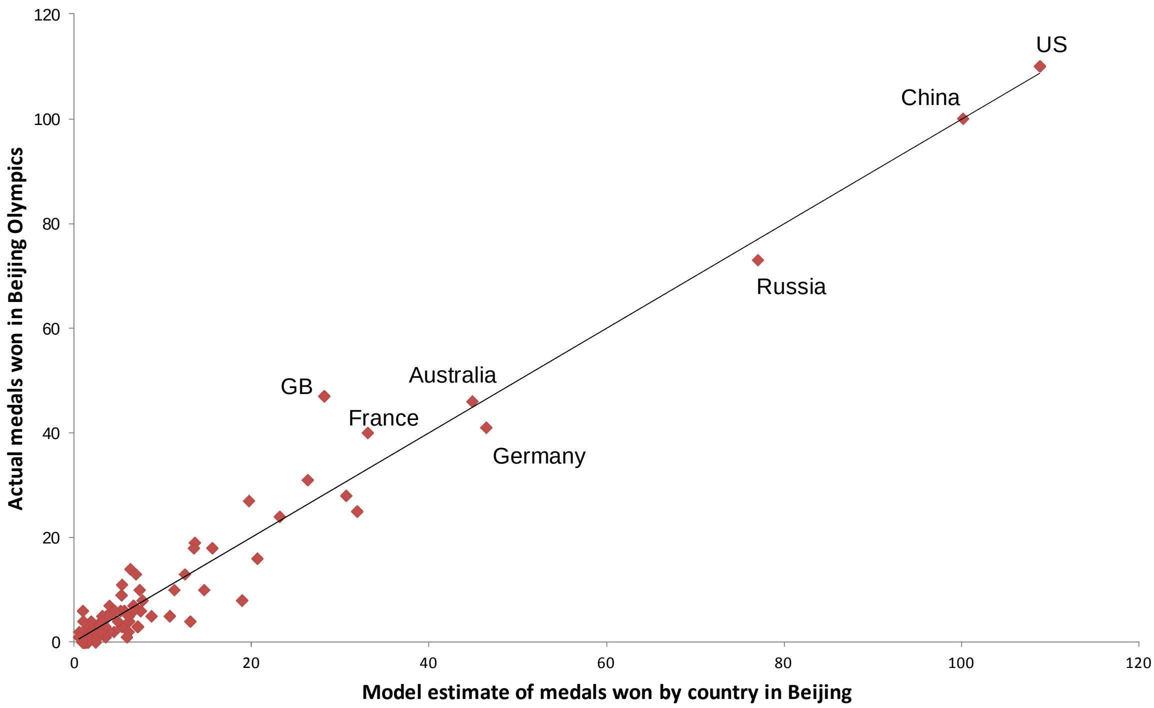
[11] The unadjusted medal shares added up to around 102%, so these were all scaled down by a factor of 1.02 to give the published results with the total number of medals set to 958, the same as were awarded in Beijing.

As indicated by t-statistics greater than 2, all explanatory variables in both model variants were statistically significant at the 99% confidence level. When the lagged dependent variable is added, however, we find that the level of GDP is the best economic variable to include while the ex-Soviet bloc dummy is no longer statistically significant because it is already captured in the past performance variables. The home country advantage variable remains highly significant statistically but smaller in magnitude than in the first model (where it is actually implausibly large). The explanatory power of the second model rises significantly and the standard error of the model is reduced by more than half.

In general, the first model is most interesting as a guide to the underlying economic and political drivers of past Olympic performance, while the second model is better for setting benchmarks against which to assess current and future Olympic performance. We have therefore used the second model to derive the medal projection results in Table 1 above, but the results of the first model are also discussed in some detail in the main text.

Figure 1 below gives a visual indication of how closely our preferred second model fits the actual Beijing 2008 results. We can see from this chart that there are a large number of countries clustered around the bottom left of the chart who won only a few medals and this is in line with the model estimates. The model also fits well the performance of China and the US at the top right of the chart, although Russia was a slight underperformer in Beijing relative to what the model would have suggested, as was Germany. The largest 'outperformer' in Beijing, however, was Great Britain, whose medal haul of 47 was well above the model estimate of around 30. France outperformed but by a smaller margin than Great Britain, while Australia delivered a par performance according to the model.

Figure 1: How well does our model fit the actual models won in Beijing?



Of course, the good model fit in Figure 1 is not so surprising given that we knew the actual results in Beijing and so could fit the data to these – the challenge of estimating likely medal hauls in London 2012 is, at this stage, much more difficult, which is why we regard our model projections in Table 1 above as benchmarks against which to judge performance, rather than precise forecasts. Whether Great Britain, with home advantage in London 2012, can maintain the model outperformance it showed in Beijing will be of particular interest.