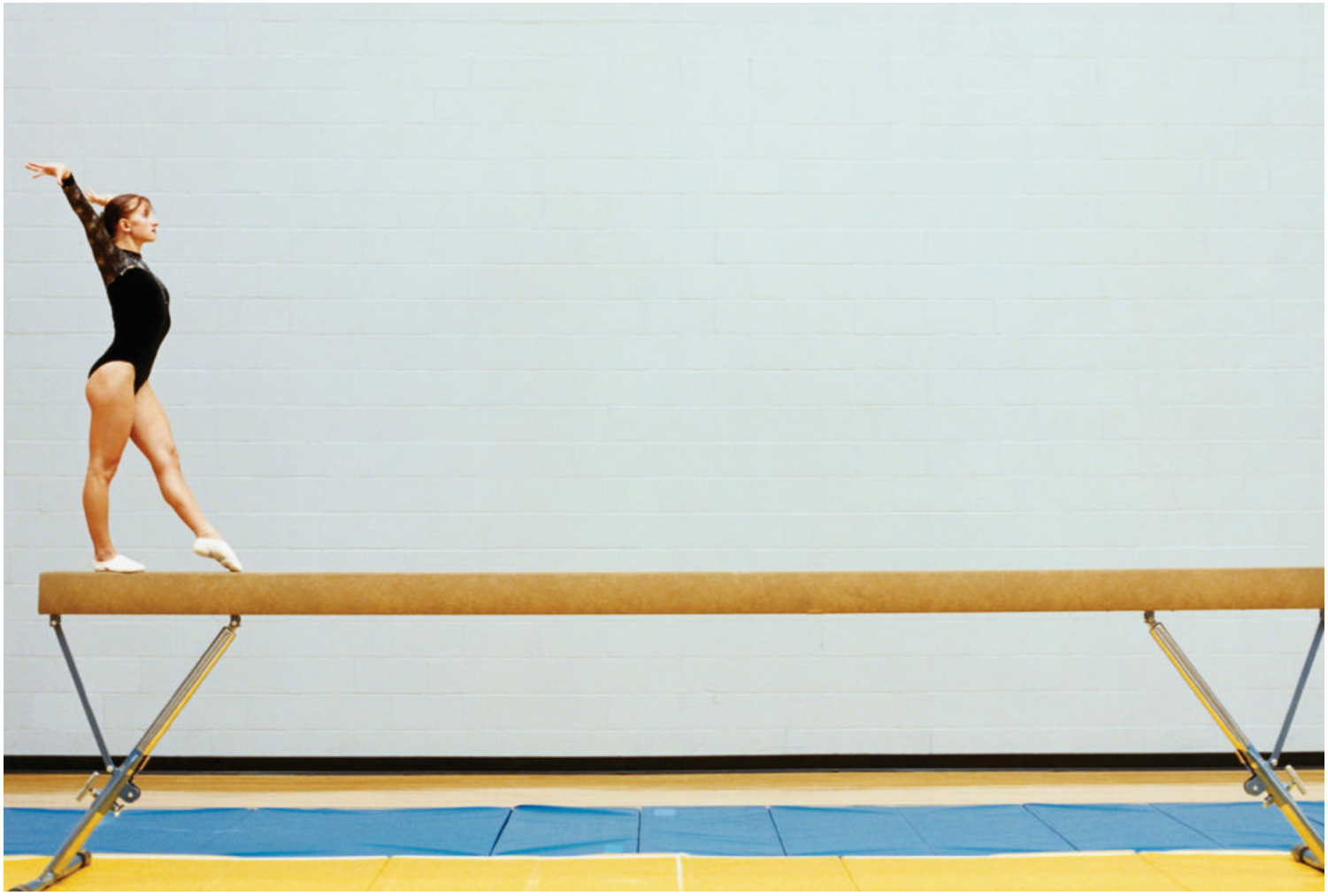
Economics

Economic Briefing Paper: Modelling Olympic Performance*



*connectedthinking

PRICEWATERHOUSECOOPERS

Economic Briefing Paper: Modelling Olympic performance

Executive Summary

As a light-hearted contribution to the debate on likely medal tallies in Beijing, this paper presents some analysis on the determinants of past Olympic performance and uses this to produce some benchmarks against which performance at the Beijing 2008 Olympics can be judged.

The following factors were found to be statistically significant in explaining the number of medals won by each country at previous Olympic Games since 1988:

- 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);
- whether the country is the host nation; and
- medal shares in the previous Olympic Games.

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 Athens 2004 Olympics, despite it being held 15 years after the fall of the Berlin Wall. This effect is expected to continue to be evident at Beijing, although it could fade gradually over time. We can see a similar effect at work in China, where state support has boosted Olympic performance: sport it seems is one area where a planned economy can succeed!

There is a marked contrast between the success of China and the performance of India, which won only one medal in the Athens Olympics (the same as at Sydney), despite having a population of over a billion. The explanation seems to lie in India's passionate love of cricket, which is not an Olympic sport but may attract potential stars away from athletics and other Olympic events.

We find that host nations generally 'punch above their weight' at the Olympics, which bodes well for the Chinese team in Beijing. In fact, our model projects a very close race between the US and China for top place in the medals rankings in Beijing.

The benchmarking feature of the model can be illustrated with reference to the UK. In Athens, the UK won 30 medals, which is more than the 25 medals our model would have predicted. Even taking into account this strong past performance, our analysis suggests that a 'par performance' for the UK would be to win 28 medals in Beijing. So matching their Athens medal haul of 30 would be a significant achievement by the British team.

Economic Briefing Paper: Modelling Olympic Performance

With the 2008 Olympic Games fast approaching, there will inevitably be much speculation about how many medals each country will win in Beijing. In this paper we consider, as a light-hearted 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 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

Our model includes data on medal performance from the Olympic Games since 1988[2]. We find that, in explaining the share of the total medals awarded won by each country, the following factors are statistically significant (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);
- whether the country is the host nation; and
- medal shares in the previous Olympic Games.

David vs Goliath

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

---

[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] We have not taken the analysis back before this since the 1980 and 1984 Olympics were affected by major boycotts and analysing data for Games before 1980 is beyond the scope of this study. Our results are, however, broadly comparable to those of the studies quoted in the previous footnote, which do cover a longer time span. Because we have a lagged dependent variable in our preferred model, the 1988 medal shares are used only for this lagged variable in explaining 1992 medal performance.

- 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 and Ethiopia); 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[3]); and

- although this is difficult to prove, there is a popular view that some athletes from poorer countries may be 'hungrier' for success; if true, this greater motivation 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

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 highly significant, 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[4], which still persisted in Athens more than a decade after the collapse of the old Soviet regime. These effects may gradually fade if the new regimes in ex-Soviet countries invest less in sport than before, and the athletes trained as youngsters under the old regime retire, but on the other hand the strong sporting traditions created in these countries could last much longer.

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 China in Beijing

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 strong in Sydney, where Australia performed very well to won 58 medals (compared to 49 in Athens), but somewhat less evident in Athens, where Greece only increased its medal total to 16, as compared to 13 in Sydney. In percentage terms, however, the Greek medal total increased by over 20%, so this was still statistically significant.

In Beijing, the combination of the home country effect and the state support for sport highlighted above is expected to lead to a particularly significant boost to Chinese medal performance, allowing them to challenge the US for the top position in the medal table at Beijing.

[3] 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?

[4] 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.

History matters

Finally, we found that the explanatory power[5] of the model was increased significantly by including medal shares at the previous Games, which can be interpreted as reflecting the fact that sources of comparative advantage in sport tend to persist over time. The coefficient on this lagged dependent variable was around 0.6.

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 was higher in Athens than our model would have predicted;

- **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 Beijing 2008

If we apply the model to the latest available data for each country, we get estimated medal targets for Beijing 2008 as shown in Table 1 (assuming that the total number of medals awarded is the same[6] as in Athens to allow direct comparison with results from 2004).

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

[5] The explanatory power of the preferred model was reasonably high, as indicated by an adjusted R-squared of 0.89 (i.e. the model explains just under 90% of the variance in medal shares between countries). If we exclude the lagged dependent variable, the explanatory power of the model drops to only around 50%.

[6] In practice, there is likely to be some change in the total number of medals awarded in Beijing, 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.

- as the host nation in Beijing and an economy which has grown very strongly since 2004, the medal 'target' of 88 for China according to our model is much higher than its actual medal totals in Athens (63) or Sydney (59); in fact, our model predicts that China may be very slightly ahead of the US (87) although this difference is well within the margin of error of the model so the race for top place is really too close to call based on this analysis[7];

- Russia is projected by the model to continue to perform strongly relative to the size of its economy in third place (79 medals), well ahead of Germany (43) and Australia (41) in fourth and fifth places; based on past performance, however, we believe there is a good chance that all three of these countries could out-perform their model medal targets given their strong Olympic traditions;

- 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 won 63 medals in Athens while India won only one (the same as in Sydney); our model can explain some of this divergence, but still suggests that India is a significant underperformer, with a model target of 6 medals for Beijing. 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[8];

- our model estimates suggest that European countries such as France, Italy and the Netherlands should be pleased if they can match their Athens medal totals, which were generally somewhat above the levels predicted based on the size of their economies and the lack of Soviet-style state support for sport in these countries;

- for Britain, our model indicates a target for Beijing of 28 medals, the same as in Sydney, but down slightly from the 'above par' result of 30 medals in Athens;

- countries where the model targets for Beijng are identical or nearly identical to actual medal totals in Athens include Ukraine, Romania, Spain, Hungary, Canada, Bulgaria, Turkey and the Czech Republic;

- countries that the model suggests have the economic potential to do significantly better than in Athens include Poland, Brazil, Mexico and Indonesia; it will be interesting to see if they can improve their standings in Beijing; and

- overall the model estimates suggest that the top 30 countries might be expected to win around 82% of all the medals awarded in Bejing, which would be the same as in Sydney but actually down slightly from 83% in Athens. This also broadly mirrors the shape of the world economy, in which the top 30 countries account for around 84% of world GDP.

[7] In addition, Olympic medal table rankings are traditionally based first on the number of gold medals, rather than total medals as in our model, and our model has tended to under-predict the performance of the US (and indeed Russia) in the past two Olympics.

[8] Chess (although not an Olympic sport) is another example of this phenomenon, as Chinese players (particularly women, where Xie Jun, Zu Chen and Xu Yuhua 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 current world chess champion) than to state support for chess.

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

John Hawksworth
Head of Macroeconomics
PricewaterhouseCoopers LLP (UK)
Tel: +44 207 213 1650
E-mail: john.c.hawksworth@uk.pwc.com

Table 1: Model estimates of Beijing 2008 Olympics medal totals as compared to Athens 2004 results

Country	Model estimate of medal total in Beijing 2008	Medal total in Athens 2004	Difference
1. China	88	63	+25
2. US	87	103	-16
3. Russia	79	92	-13
4. Germany	43	48	-5
5. Australia	41	49	-8
6. Japan	34	37	-3
7. France	30	33	-3
8. Italy	29	32	-3
9. Great Britain	28	30	-2
10. South Korea	27	30	-3
11. Cuba	24	27	-3
12. Ukraine	23	23	0
13. Netherlands	20	22	-2
14. Romania	19	19	0
15. Spain	19	19	0
16. Hungary	17	17	0
17. Belarus	15	15	0
18. Greece	15	16	-1
19. Poland	14	10	+4
20. Canada	13	12	+1
21. Bulgaria	13	12	+1
22. Brazil	12	9	+3
23. Turkey	11	10	+1
24. Thailand	10	8	+2
25. Czech Republic	9	8	+1
26. Kazakhstan	9	8	+1
27. Iran	8	6	+2
28. South Africa	8	6	+2
29. Indonesia	8	4	+4
30. Mexico	8	4	+4
Top 30 total medals	**762**	**776**	**-12**
Other countries	166	154	+12
Total medals	**928**	**928**	**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 Olympic Games as an additional independent variable and has much higher overall explanatory power (as indicated by the respective adjusted R-squared coefficients of 0.89 for the second model, as against 0.5 for the first model). This second model therefore forms the basis for the Beijing 2008 medal estimates quoted in Table 1 above. Since the unadjusted model estimates for medal shares in Beijing did not add up exactly to 100%, a scaling factor was applied to given the results shown in Table 1.

As indicated by t-statistics greater than 2, all explanatory variables in both model variants were statistically significant at the 95% confidence level. When the lagged dependent variable is added, however, both the value and the statistical significance of the other explanatory variables falls considerably, but the explanatory power of the model as a whole rises significantly and the standard error of the model is reduced by more than half.

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

Explanatory variables	**Model without past Olympic performance variable**	**Model with past Olympic performance variable**
Constant	-0.0198	-0.0034
Log (population)	0.0056 (5.8)	0.00147 (3.3)
Log (GDP per capita at PPPs)	0.0066 (4.3)	0.00133 (2.0)
Ex-Soviet bloc dummy	0.026 (5.1)	0.003 (2.3)
Host country dummy	0.046 (3.2)	0.018 (5.8)
Medal share in previous Olympic Games	-	0.58 (19.0)
Explanatory power (adjusted R-squared)	*0.50*	*0.89*
Standard error of model	*0.014*	*0.006*
Number of countries covered	*80*	*80*

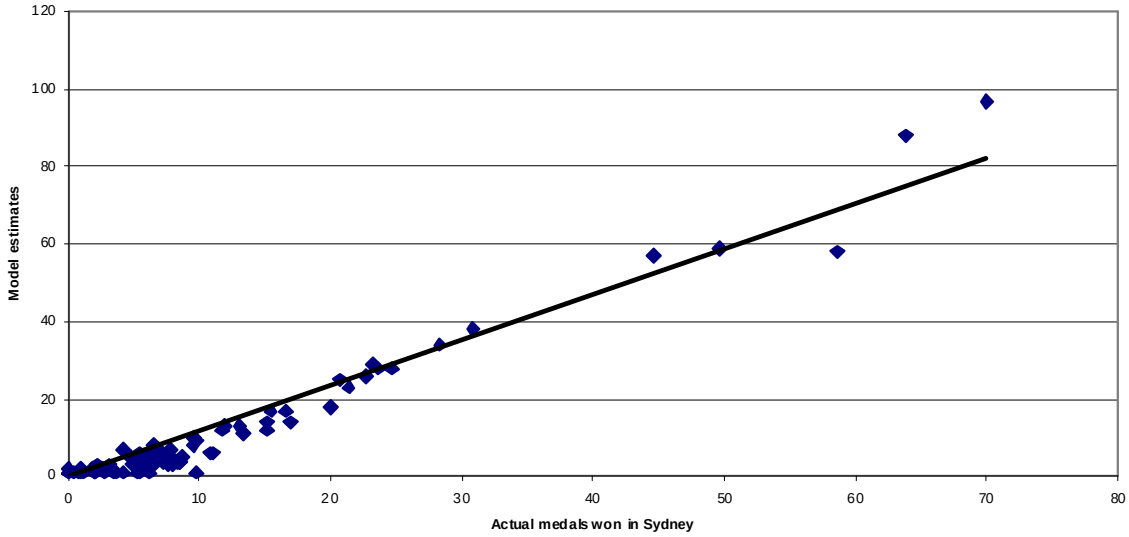
Note: t-statistics shown in brackets for explanatory variables

Source: PwC analysis using data from 80 medal-winning countries in 1988, 1992, 1996 and 2000 Olympics, plus World Bank data on population and GDP per capita at PPP exchange rates. For the Beijing projections in Table 1 above, the 2004 Athens Olympics results were factored into this model together with the latest GDP and population estimates.

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. Since the latter is likely to be of greater general interest, we have focused on the second model in the main text of the paper and used it to derive the results in Table 1 above.

The chart below gives a visual indication of how closely our preferred second model fitted the Sydney 2000 results. We can see from the chart that there was a long tail of relative 'underperformers' towards the bottom end of the medal table (i.e. where actual performance was below model projections), as well as some 'outperformers' (notably Russia and the US) at the top of the table. A somewhat similar pattern was seen for Athens 2004.

Model fit with past performance included



pwc.com/economics