university of groningen / faculty of economics and business

12006-EEF

Participation and Performance at the London 2012 Olympics

Gerard H. Kuper
Elmer Sterken

SOM is the the research institute of the Faculty of Economics & Business at the University of Groningen. SOM has six programmes:

- Economics, Econometrics and Finance
- Global Economics & Management
- Human Resource Management & Organizational Behaviour
- Innovation & Organization
- Marketing
- Operations Management & Operations Research

Research Institute SOM
Faculty of Economics & Business
University of Groningen

Visiting address:
Nettelbosje 2
9747 AE Groningen
The Netherlands

Postal address:
P.O. Box 800
9700 AV Groningen
The Netherlands

T +31 50 363 7068/3815

www.rug.nl/som

Participation and Performance at the London 2012 Olympics

Gerard H. Kuper
University of Groningen
g.h.kuper@rug.nl

Elmer Sterken
University of Groningen

Participation and Performance at the London 2012 Olympics

Gerard H. Kuper and Elmer Sterken

Institute of Economics, Econometrics, and Finance,

University of Groningen, PO Box 800, 9700 AV Groningen,

The Netherlands

Corresponding author: Gerard H. Kuper

email: g.h.kuper@rug.nl

telephone: +31 50 3633756

fax: +31 50 3637337

July 2012

Acknowledgements: We would like to thank Rense Deen for excellent research assistance.

Participation and Performance at the London 2012 Olympics

Abstract

The current paper predicts the medal tally for the London 2012 Olympic Games. The forecast procedure consists of analyzing participation and success at the country level of the three most recent editions of the Olympic Summer Games. Potential explanatory variables for medal winnings are income per capita, population, geographical distance to the Games, success in terms of medals won at World Championships, and the home advantage. Our forecasts show that the China takes first place in the medal tally with 44 gold medals, followed by the United States of America winning 33 gold medals. We expect Great Britain to take fourth place winning 23 gold medals.

Keywords: Olympic Summer Games, Medal predictions, Panel data models.
JEL-code: R0, O1, C23, Z10.

1. Introduction

Ever since the first Ancient Olympic Games in 776 BC the ultimate aim of competing, especially in athletics, was to be the best. Winning an Olympic event was the highest honor people could achieve (Lämmer, 1992, p109 refers to Homer, Book VIII, pages 147-148). What started as a competition to strengthen the bond between Greeks became an international affair in the 2nd century AD, when competitors from outside Greece competed in the Olympic Games.

In 393 AD the ancient Olympic Games were abolished because they were considered unchristian. After about 15 centuries the Greek government reinstated the Olympic Games as an international competition for the best amateur athletes. At the start of the modern Olympic Games in 1896, the Olympic Games were an elitist event, mostly for men (Wallechinsky and Loucky, 2012). Similar to the ancient Olympic Games, the Games were held every four years (a period called the Olympiad). The main purpose was to foster the ideal of "…a sound mind in a sound body…," and to promote friendship among nations. Initially, a second series of Games, the so-called intercalated Games, were supposed to be organized every four years in Athens in between the Olympic Games. These series was not successful, but neither were the 1900 and the 1904 Olympic Games. To revitalize the Games Athens organized the first, and last, intercalated Games in 1906, which were successful. However, the medals awarded in 1906 are not acknowledged by the International Olympic Committee (IOC).

This paper presents forecasts for medal winnings at the 2012 Olympic Games in London. There is a huge literature by sociologists and economists analyzing the impact of social and economic conditions on the outcomes of the Olympic Games competition. We review this literature in Section 3. This is the fourth time that we apply the methodology we have developed for forecasting the medal tallies for the Olympic Summer Games. The forecasts we present are to be interpreted as expectations based on past performance. After the Games we can identify which countries underperformed, and which countries performed better than expected. In what follows we model participation and success at the most recent editions of the Olympic Summer Games. Our goal is to investigate the role of key determinants such as population size, income per head, distance and home advantage in determining participation and success. Before we discuss the methodology, the data and the econometric model, we first give an overview of related work in Section 2. Section 3 presents some facts of the London 2012 Games. In Section 4, the determinants for success are discussed, and Sections 5 and 6 present our model and the forecasts. We summarize our findings in the last section.

2. Literature

There is a huge literature on the Olympic Games, and its interaction with economic and political developments. Firstly, in the early editions of the Games economic conditions determined participation probably more than athletic qualities. At the end of the 19th century sports were the exclusive right of the wealthier people in mainly developed countries. Secondly, the Games have been used to stimulate nationalistic sentiments. Some examples are mentioned in the previous section. Thirdly, it may be argued that organizing large scale sporting events, like the Olympic Games, lead to significant economic benefits. National success at the Games may even lead to higher rates of economic growth by raising consumer and producer confidence (see Sterken, 2006).

For the post-World War II Games sociologists and economists have analyzed the impact of social and economic conditions on the outcomes of the Olympic Games competition. Earlier examples relating success to social conditions are Ball (1972), Levine (1972), and Grimes et al. (1974). They show that socialist and host countries systematically outperform other countries. Shughart and Tollison (1993) focus on the consequences of the end of Soviet socialism for Olympic performances. Another strand of literature analyses recent editions of the Olympic Games with a focus on predicting Olympic success. Examples are Johnson and Ali (2004) and Bernard and Busse (2004). This literature shows that for the post-World War II editions of the modern Games factors like income, the home advantage, and the fact that a country has a socialist/communist tradition have a major impact on position of countries in the final medal tally (see also Kuper and Sterken, 2011). According to these studies a higher income allows for labor specialization, gives possibilities to train athletes better, to send a larger group of athletes to the Games, etc. The home advantage helps to send more athletes by regulation (the home country participates in a large majority of all events) and to get more crowd support during the Games. The post-war studies estimate the home country advantage to be about two percentage points of the share in medals earned (see Courneya and Carron, 1992, and Nevill and Holder, 1999). After World War II both professionalization of sports in the Western world and the communist tradition helped to create a professional sports environment and to increase labor division even further. The impact of being a communist country is even estimated to be higher leading to about a three percentage points increase in the medal share.

3. Some facts about the London 2012 Olympics

The Games of the XXX Olympiad in London start on July 25, 2012 with the football competition two days before the actual opening of the Games. The Games close on Sunday August 12.

Athletes from over 200 countries are expected to compete in 26 sports. Compared to the Beijing 2008 Games, in London 2012 baseball and softball are dropped from the program, and women's boxing makes its Olympic debut in three weight classes: flyweight, middleweight, and lightweight. One weight class in boxing for men is cancelled. There are more changes: in sailing the Tornado class catamaran is dropped, and the women's fleet race in the Yngling class is replaced by the Elliot match race. In tennis mix doubles are introduced. London initially made a bid for 28 sports with golf and rugby sevens replacing softball and baseball. But the IOC voted against golf and rugby sevens. These sports will be introduced in 2016 in Rio de Janeiro. In total there are 302 medal events (the same number as in Beijing 2008), with two bronze medals awarded for 53 events (35 events for men and 18 for women) in boxing, judo, wrestling, and taekwondo.

Table 1 shows how the modern Olympic Summer Games have evolved. In 1896, 246 athletes of 12 nations competed in 43 events. All athletes in the first Modern Olympic Games were men. It is noteworthy that women did compete in the Ancient Olympics (see Wallechinsky and Loucky, 2012). Over time the Games grew in scale and scope. In the Games of the XXIX Olympiad in Beijing the number of athletes is 45 times higher, the number of participating counties increased to over 200, the number of events increased sevenfold, and 42.3% of all athletes are female. This trend was not continuous due to boycotts for political reasons, especially during the Cold War.

Table 1 – An overview of all editions of the Olympic Summer Games.
Source: W.J. Mallon, co-founder and previous chairman of the International Society of Olympic Historians.

Edition	**Year**	**City**	**Nations**	**Events**	**Athletes**	**Women**	**% Women**
I	1896	Athens	12	43	246	0	0%
II	1900	Paris	29	94	1613	22	1.4%
III	1904	St. Louis	14	94	649	6	0.9%
-	1906[1]	Athens	21	74	840	6	0.7%
IV	1908	London	22	106	2002	37	1.8%
V	1912	Stockholm	27	102	2377	53	2.2%
VII	1920	Antwerp	29	152	2576	65	2.5%
VIII	1924	Paris	44	126	3066	135	4.4%
IX	1928	Amsterdam	46	109	2871	274	9.5%
X	1932	Los Angeles	38	117	1329	126	9.5%
XI	1936	Berlin	49	129	3955	329	8.3%
XIV	1948	London	59	136	4070	393	9.7%
XV	1952	Helsinki	69	149	4931	521	10.6%
XVI	1956[2]	Melbourne	72	151	3345	383	11.4%
XVII	1960	Rome	83	150	5348	612	11.4%
XVIII	1964	Tokyo	93	163	5136	680	13.2%
XIX	1968	Mexico City	112	172	5555	783	14.1%
XX	1972	Munich	121	195	7122	1059	14.9%
XXI	1976	Montreal	92	198	6071	1261	20.8%
XXII	1980	Moscow	80	203	5253	1120	21.3%
XXIII	1984	Los Angeles	140	221	6793	1569	23.1%
XXIV	1988	Seoul	159	237	8423	2201	26.1%
XXV	1992	Barcelona	169	257	9385	2723	29.0%
XXVI	1996	Atlanta	197	271	10329	3512	34.0%
XXVII	2000	Sydney	200	300	10647	4068	38.2%
XXVIII	2004	Athens	201	301	10558	4301	40.7%
XXIX	2008	Beijing	203	302	10906	4611	42.3%

The most medals until and including the Beijing 2008 Games are won by the United States of America as is shown in Table 2, about twice as much as the former Soviet Union. China ranks seventh, but we expect China to move to fifth place in the all-time medal count in 2012.

[1] The Games of 1906 are *Intercalated* Games.
[2] Including the equestrian events which were held in Stockholm, Sweden due to quarantine regulations in Australia.

Table 2 – All-time medal count of the Olympic Summer Games before London 2012 (Top 25). Source: Wallechinsky, D.; Loucky, J. (2012).

Country	**Gold**	**Silver**	**Bronze**	**Total**
United States of America	929	729	638	2296
Soviet Union (until 1988)	395	319	296	1010
Germany	247	284	320	851
Great Britain	207	255	253	715
France	191	212	233	636
Italy	190	157	174	521
China	163	117	105	385
Hungary	159	141	159	459
East Germany (until 1988)	153	129	127	409
Sweden	142	160	173	475
Australia	131	137	164	432
Japan	123	112	126	361
Russia (since 1996)	108	97	112	317
Finland	101	83	115	299
Romania	86	89	117	292
The Netherlands	71	79	96	246
South Korea	68	74	73	215
Cuba	67	64	63	194
Poland	62	80	119	261
Canada	58	94	108	260
Norway	54	48	42	144
Bulgaria	51	84	77	212
Czechoslovakia (until 1992)	49	49	45	143
Unified Team (1992)	45	38	29	112
Switzerland	45	70	66	181

4. Determinants of participation and success

In our earlier forecasts of success for the Olympic Games – since the Winter Games of Salt Lake City in 2002 – we have modeled success conditional on participation, and we use the results of World Championships in the years prior to the Games as an additional, and powerful, explanatory variable because many of the athletes who participate at the World Championship also enter the Olympic Games. For an analysis of our forecast performance we refer to Appendix A.

We apply econometric models to quantify and identify determinants of participation and success at the Olympic Games. These determinants are based on the literature and our experience in predicting participation and success at the Olympic Games. We estimate the model in a combined time-series cross-section form, and we use the fixed-effects estimator to account for unobserved differences between countries and/or time periods. We present simple models that explain participation and success at the national level. There are various reasons to model at the national level instead of individual or event cases. First, the impact of income cannot be measured on the individual level. Second, modeling at the individual or event level is more sensitive to measurement errors. Thirdly, success is mostly discussed at the country level.

The determinants for participation are demographic (population), economic (income), and geographic (distance to the host country) in nature. Also home advantage may determine participation. These determinants are predetermined. So, there is no endogeneity bias. The distance to the Games translates into travelling costs, which could also be considered as an economic component. We measure the distance to the Games as the shortest distance between two points on a sphere (Sinnott, 1984, see Appendix B for details). The main argument why economic welfare is important in explaining Olympic participation is division of labor. If a country becomes wealthier, specialization of labor input is allowed and individuals can make a living out of their special sports competitive advantages: we assume that income will determine the training, access to training facilities, and health conditions of the potential athletes. The home advantage is a dummy variable (1 if a country hosts the Games, and 0 in other cases). Home countries are allowed to send more athletes.

There are several arguments why participation at the Games is not proportional to the absolute size of the population. The main argument is that participation at the Games is not proportional to population since the number of athletes that represent their country at the Games is restricted. Another argument – which is based on Reiss (1989) – states that the maximum performing individual of a population of size N will be of the order $(\log N)^{1/2}$. However, this argument is valid for standard normal series, and population is not normally distributed. Nevertheless, in this paper we use the square root of population (in logarithm) as explanatory variable because experimenting with other specifications in earlier forecasts yields similar estimation results and forecasts.

The main determinants for medals won are the results at the World Championships in Olympic events and participation. We also include interaction effects. This will be discussed in more detail below. Note that in our set-up income per capita, population and distance has an indirect effect on success through participation.

Just as in the participation equation we also include the home advantage in success. In both equations we include one-period lags of the home advantage dummy: we hypothesize that a country that has organized the Games may benefit also four years after the Games. There may also be a lead effect because cities are elected seven years prior to the Games they have bid on. The lead effect is not yet considered in this paper. In the evaluation of the London Games we will analyze the lead effect.

5. Modeling participation and success

We define medals s won by a country i (for colors gold, silver and bronze, indexed by c) at the Summer Games in year t as shares in total gold, silver and bronze medals awarded. In similar fashion we define medal shares w won by a country i at the World Championships at the year before the Olympics in year t. Also participation p for each country i is defined as shares of total participation at the Games in year t. Modeling in shares may reduce problems of nonstationarity. However, tests for unit roots in a sample with a very small time series dimension (four periods) are not very powerful. Another advantage of modeling shares is that we directly can compare the performance of countries if a different number of medals are awarded at subsequent Games. For instance, the Sydney 2000 Games include 300 medal events, the Athens

2004 Games includes 301 medal events, while the Beijing 2008 Games and the 2012 London Games each feature 302 medal events. Note that the number of bronze medals differs from the number of gold and silver medals, because in boxing, judo, wrestling, taekwondo and karate two bronze medals are awarded in each event class. Finally, in a case of a tie sometimes two gold or silver medals awarded.

Note that participation and medal shares at Summer Games and World Championships in year t, $p_{e,i,t}$, $s_{c,e,i,t}$, and $w_{c,e,i,t}$ are available for different events, indexed by e. We distinguish team events and individual events. The latter is split in female and male events (events for men include mixed events in equestrian, badminton and tennis).

In our sample we include 126 countries that cover all medal winning countries, 95-97% of total participation at the Olympic Games, 99% of real World GDP, and 91% of the world population. Below we present the definitions of the variables. For the sources and definitions of variables we refer to Appendix B.

Modeling participation

The model for participation shares $p_{e,i,t}$ for different events (index e=teams, and male and female individual events) is a fixed effects panel model:

$$p_{e,i,t} = \underbrace{b_{e,i} + b_1 \frac{Y_{i,t}}{N_{i,t}} + b_2 (\log N_{i,t})^{1/2} + b_3 d_{i,t} + b_4 h_{i,t} + b_5 h_{i,t-1}}_{\hat{p}_{e,i,t}} + e^{p}_{e,i,t} \quad (1)$$

Cross-section fixed effects – included as $b_{e,i}$ – measure unobserved differences between countries. An example is the difference in sports culture between countries). We also include a one period lag for the home dummy h: A country may profit from the bigger delegation sent to the home Games also four years after the home Games. One could argue that there is also a lead effect because a country may prepare itself by sending more athletes also four before the home Games. This effect is not considered in this paper.

Table 3 – Estimation results for participation at the London 2012 Olympic Games with fixed effects for countries (robust standard errors are in brackets). The fixed effects are not reported.

Dependent variable: participation share p for female and male individual events and team events.

Explanatory variables:

Y/N = income per capita;

N = population;

d = distance from the capital of the host country to the capital of the participating country;

h = 1 if a country hosts the Games, else 0.

	Women	**Men**	**Team**
Y/N ($\times 10^{-4}$)	0.209	-0.298	0.054
(se)	(0.047)	(0.072)	(0.033)
$\sqrt{(\log(N))}$ ($\times 10^{-3}$)	0.190	4.713	-0.113
(se)	(0.037)	(0.323)	(0.225)
d ($\times 10^{-7}$)	-0.031	0.178	0.002
(se)	(0.012)	(0.014)	(0.005)
$h(-1)$ ($\times 10^{-2}$)	-0.019	0.058	0.254
(se)	(0.039)	(0.063)	(0.139)
h ($\times 10^{-2}$)	0.463	0.652	1.219
(se)	(0.061)	(0.070)	(0.106)
R^2	0.997	0.997	0.991
Countries	126	126	126
Observations	378	378	378

From Table 3 we conclude that, using 5% significance levels, income per capita has the expected positive effect on participation for female athletes. For teams the effect is weaker, both in size and significance (p-value for a one-tailed test is 0.051). The population size has a significant positive effect on male and female participation. The effect is particularly strong for men. Distance only has the expected negative effect for women. The home advantage effect on participation is significant and about twice as strong for teams as it is for individual athletes. The lagged effect of the home advantage is small, and only significant for teams.

Modeling success

The model for medal shares $s_{c,e,i,t}$, is also a fixed effects panel model, with cross-section fixed effects denoted by $a_{c,e,i}$. Medal shares for each medal color (index *c*=gold, silver and bronze) and each event (index e=teams, and male and female individual events) are explained by participation and world championship results prior to year *t*:

$$s_{c,e,i,t} = a_{c,e,i} + a_1 p_{e,i,t} + a_2 w_{c,e,i,t} + a_3 p_{e,i,t} \times w_{c,e,i,t}$$
$$+a_4 h_{i,t} + a_5 h_{i,t-1} + e^{s}_{c,e,i,t} \quad (2)$$

Again, we include a one period lag for the home dummy. This specification implies that income per capita, population and distance have an indirect effect on success through participation. Results of world championships have a direct effect on success. The interaction term allows the effect of participation on success to depend on the world championship results. Also the effect of world championship results on success depends on participation. We expect the interaction effect to be positive. Obviously, the overall effect of participation and world championship results on success is evaluated in a simultaneous test of coefficient a_1 and a_2 respectively, and the coefficient of the interaction term a_3.

Table 4 leads us to conclude that the world championship results and participation are important determinants for success, especially for individual male and female medals. This conclusion is based on joint significance tests reported in the last rows of Table 4. The interaction term is significantly positive for gold medals and silver and bronze medals for men. The home effect on success is not significant for teams. For men the home effect has the expected positive sign for Silver and Gold. The home effect for women is negative for Silver, but positive for Gold. This implies that women seem to benefit in the finals. There is also a lagged effect of the home dummy. However, the sign is ambiguous.

Table 4 – Estimation results for success at the London 2012 Olympic Games with fixed effects for countries (robust standard errors are in brackets). The fixed effects are not reported.

Dependent variable: medal share *s* for female and male individual events and team events, and for Gold, Silver and Bronze.

Explanatory variables:

w = world championship results (as share of total medals);

p = *p*articipation share;

d = distance from the capital of the host country to the capital of the participating country;

h = 1 if a country hosts the Games, else 0.

	Gold			**Silver**			**Bronze**		
	Women	**Men**	**Team**	**Women**	**Men**	**Team**	**Women**	**Men**	**Team**
w	0.220	-0.027	-0.054	0.223	-0.053	0.094	0.143	-0.008	0.224
(se)	(0.096)	(0.017)	(0.123)	(0.111)	(0.023)	(0.171)	(0.108)	(0.020)	(0.107)
p	-0.201	0.007	0.124	0.830	-0.000	0.254	0.192	0.005	0.035
(se)	(0.311)	(0.022)	(0.054)	(0.380)	(0.005)	(0.078)	(0.346)	(0.026)	(0.063)
$p \times w$ ($\times 10^2$)	0.198	0.404	0.360	-0.164	0.197	0.017	0.015	0.074	-0.167
(se)	(0.069)	(0.048)	(0.132)	(0.010)	(0.039)	(0.223)	(0.094)	(0.042)	(0.099)
$h(-1)$	0.015	-0.011	0.001	-0.021	-0.001	0.002	0.010	-0.000	-0.004
(se)	(0.002)	(0.002)	(0.001)	(0.002)	(0.004)	(0.001)	(0.002)	(0.001)	(0.001)
h	0.014	-0.004	-0.000	-0.007	0.013	-0.001	0.008	0.009	-0.000
(se)	(0.002)	(0.002)	(0.009)	(0.003)	(0.003)	(0.001)	(0.003)	(0.001)	(0.001)
R^2	0.952	0.955	0.820	0.902	0.927	0.592	0.890	0.956	0.675
Countries	126	126	126	126	126	126	126	126	126
Observations	378	378	378	378	378	378	378	378	378
Hypotheses									
Effect of p									
F-stat	4.228	35.049	6.986	3.545	13.055	5.596	0.177	1.623	1.564
(p-value)	(0.016)	(<0.001)	(0.001)	(0.030)	(<0.001)	(0.004)	(0.838)	(0.199)	(0.212)
Effect of w									
F-stat	46.599	39.216	14.314	2.007	14.618	1.041	4.018	1.753	2.221
(p-value)	(<0.001)	(<0.001)	(<0.001)	(0.137)	(<0.001)	(0.355)	(0.019)	(0.175)	(0.111)

6. Forecasting success

The estimates presented above are used to forecast medal winning at the London 2012 Olympic Games. We apply a two-stage forecasting procedure (for teams, and male and female individual events). In the first step we forecast participation (for teams, and male and female individual events) for 2012. In the second step we calculate expected medal shares for 2012 by replacing the actual values for participation in Equation (2) with the fitted values for participation $\hat{p}_{e,i,t}$ from Equation (1). The only exception in this procedure is the forecasts for team events. For these events it is known well in advance which countries are qualified. This information is used in our forecasts for the team results.

Table 5 – Top-30 medal forecasts for the London 2012 Olympic Games.

Rank	Country	Gold	Silver	Bronze
1	China	44	11	22
2	United States of America	33	36	34
3	Russia	27	28	34
4	Great Britain	21	19	19
5	Australia	13	19	13
6	Japan	12	9	10
7	Germany	10	13	20
8	Italy	10	8	13
9	France	9	12	15
10	South Korea	8	11	12
11	Netherlands	8	8	7
12	Romania	8	4	7
13	Ukraine	7	7	13
14	Cuba	6	8	11
15	Hungary	6	5	3
16	Belarus	4	6	9
17	Greece	4	6	0
18	Brazil	4	4	6
19	Norway	4	3	1
20	Spain	3	7	6
21	Canada	3	6	7
22	Kenya	3	5	3
23	Poland	3	4	3
24	Jamaica	3	3	3
25	Ethiopia	3	2	3
26	New Zealand	3	1	3
27	Iran	3	1	1
28	Kazakhstan	2	4	3
29	Turkey	2	3	3
30	Czech Republic	2	3	2

We expect China to win the medal race, with the USA in second place. The USA wins more medals, but China wins more Gold medals. This is the same as in the Games of Beijing in 2008. Great Britain wins more medal than in 2008, but the home advantage is not big enough to pass Russia. The Netherlands again fail to enter the Top 10, but the difference with South Korea in tenth place is small.

7. Summary and conclusion

In this paper we present forecasts for medal winnings at the 2012 Olympic Games in London in a two-step procedure. We first forecast participation in the Games of 2012, and then we forecast success conditional on our forecasts for participation. We do this for male and female events. For teams participation is known well in advance of the Games, so in this case we use actual participation.

Our model includes key determinants such as population size, income per head, distance and home advantage in determining participation and success. Our sample includes 126 countries that cover all medal winning countries.

Income per capita has the expected positive effect on participation for female athletes. For teams the effect is weaker. The population size has a significant positive effect on male and female participation, and is particularly strong for men. Distance only has the expected negative effect for women. The home advantage effect on participation is significant and about twice as strong for teams as it is for individual athletes. The lagged effect of the home advantage is small and only significant for teams.

With respect to medal winning we conclude that the world championship results and participation are important determinants, especially for individual male and female medals. For men the home effect has the expected positive sign for Silver and Gold. The home effect for women is negative for Silver, but positive for Gold. There is also a lagged effect of the home dummy. However, the sign is ambiguous.

Our predictions show that China wins the medal race, with the USA in second place. Similarly to the Games of Beijing in 2008, the USA wins more medals. Great Britain wins more medal than in 2008, but the home advantage is not big enough to pass Russia. Finally, The Netherlands fail to enter the Top 10.

References

Ball, D.W. (1972) "Olympic Games Competition: Structural Correlates of National Success," International Journal of Comparative Sociology, 15, 186-200.

Bernard, A.B. and M.R. Busse (2004) "Who Wins the Olympic Games: Economic Development and Medal Totals," Review of Economics and Statistics, 86, 413-417.

Central Intelligence Agency's World Factbook, https://www.cia.gov/library/publications/the-world-factbook/.

Courneya, K.S. and A.V. Carron (1992) "Home Advantage in Sport Competitions: A Literature Review," Journal of Sport and Exercise Psychology, 14, 13-27.

Grimes, A.R., W.J. Kelly, and P.H. Rubin (1974) "A Socioeconomic Model of National Olympic Success," Social Science Quarterly, 55, 777-782.

International Monetary Fund, "World Economic Outlook Database," http://www.imf.org/external/pubs/ft/weo/2010/02/weodata/weoselgr.aspx.

International Society of Olympic Historians, http://www.isoh.org/.

Johnson, D.K.N. and A. Ali (2004) "A Tale of Two Seasons: Participation and Medal Counts at the Summer and Winter Olympics," Social Science Quarterly, 85, 974-993.

Kuper, G.H. and E. Sterken (2011) "Determinants of Participation and Success at the Earlier Modern Olympic Games," Journal of Olympic History, 19, 20-29.

Lämmer, M. (1992) "Myth or Reality: The Classical Olympic Athlete," International Review for the Sociology of Sport, 27, 107-114.

Levine, N. (1974) "Why do Countries Win Olympic Medals? Some Structural Correlates of Olympic Games Success: 1972," Sociology and Social Research, 58, 353-360.

London 2012 Olympic Games, http://www.london2012.com/.

Map Crow, http://www.mapcrow.info/.

Nevill, A.M. and R.L. Holder (1999) "Home Advantage in Sport: An Overview of Studies on the Advantage of Playing at Home," Sports Medicine, 28, 221-236.

Reiss, R.-D. (1989) "Approximate Distributions of Order Statistics," Springer-Verlag, Berlin.

Sinnott, R.W. (1984) "Virtues of the Haversine," Sky and Telescope, 68, 159.

Shughart, W.F. and R. Tollison (1993) "Going for the Gold: Property Rights and Athletic Effort in Transitional Economies," Kyklos, 46, 263-272.

UN, United Nations Statistics Division, http://unstats.un.org/unsd/snaama/dnlList.as

Economic Research Service of the U.S. Department of Agriculture, http://www.ers.usda.gov/Data/Macroeconomics/.

Wallechinsky, D. and J. Loucky (2012) "The Complete Book of the Summer Olympics, London 2012 Edition," Aurum, London.

Appendix A – Evaluation of Olympic Games forecasts

Table A1 summarizes our forecasting performance for four Olympic Games. For Turin and Vancouver we include all 26 medal winning countries (also countries that win only one medal), while for Athens and Beijing we only take the Top-30 into account. We compare our forecasting performance with those of Sports Illustrated (SI). This US based sports journal publishes forecasts for each event and for each individual medal. From these predictions we compile the medal tally. SI bases its predictions on their huge expertise of sports and athletes. Moreover, they publish the forecast very close to the start of the Games, so they are able to include the most recent information about the athletes who are competing and about their current form. Our predictions are based on statistical techniques, and are made a couple of months before the start of the Games.

The table reports two measures of forecast performance. The mean absolute error indicates that on average our predictions are off by about 2 to 3.5 medals for the Summer Games. The mean squared error penalizes big deviations from the realizations more severely, but is not easy to interpret. Despite these differences most of the time we outperform SI, as Table A1 illustrates.

Table A1 – Our forecasting performance (KS) at the most recent Olympic Games compared with those of Sports Illustrated (SI).

	Gold		**Silver**		**Bronze**	
Mean Absolute Error	**KS**	**SI**	**KS**	**SI**	**KS**	**SI**
Athens, 2004, Top-30	2.00	2.07	1.83	2.17	3.00	3.00
Turin 2006, all countries	1.54	2.31	1.50	1.46	1.85	1.46
Beijing 2008, Top-30	2.43	2.73	2.17	3.20	3.50	3.00
Vancouver 2010, all countries	1.50	0.89	1.81	1.50	1.65	1.69
Mean Squared Error						
Athens, 2004, Top-30	6.40	8.47	5.60	8.83	17.80	19.80
Turin 2006, all countries	7.00	12.62	4.04	2.62	4.38	6.23
Beijing 2008, Top-30	10.97	15.27	8.03	15.67	23.43	17.00
Vancouver 2010, all countries	4.81	1.65	4.50	3.73	4.81	5.23

Appendix B – Definitions and data sources

Definitions

$p_{e,i,t}$ participation share of country i in year (Olympiad) t, by event e=women, men, team;

$s_{c,e,i,t}$ medal share of country i in year t, for c=gold, silver, and bronze and event e=women, men, team;

$N_{i,t}$ population of country i in year t (millions; averaged over four years: three years prior to the Games and the current year);

$Y_{i,t}$ income of country i in year t (real 2005 GDP in $ billions; averaged over four years: three years prior to the Games and the current year);

$d_{i,t}$ distance (kilometers) to the Games for country i in year t;

$h_{i,t}$ home advantage dummy (1 for host country, 0 else) for country i in year t;

$w_{c,e,i,t}$ medal share of country i in World Championships prior to year t for c=gold, silver, and bronze, by event e=women, men, team;

t time index, t = 2000, 2004, 2008, 2012, Averages of income and population are also available for 1996;

i country index, 126 countries that won at least one medal at the Olympics in year t = 2000, 2004, 2008;

c index for medal color, c = gold, silver and bronze;

e index for event, e = women, men, team.

Sources of data

Participation

The participation data for all modern editions of the Olympic Games are kindly provided by Bill Mallon (co-founder and later president of the International Society of Olympic Historians).

Medals

The main source of data on Olympic medals is Wallechinsky and Loucky (2012). Medal tallies for the world championships results are compiled from various internet sources.

Real Gross Domestic Product

Real Gross Domestic Product (GDP) in billions of 2005 dollars is published by The Economic Research Service of the U.S. Department of Agriculture. (http://www.ers.usda.gov/Data/Macroeconomics/)

Real GDP data for Montenegro, North Korea, Qatar, and Somalia are based on the Central Intelligence Agency's World Factbook. (https://www.cia.gov/library/publications/the-world-factbook/)

Population

The source for population data (in millions) is the International Monetary Fund's World Economic Outlook database (http://www.econstats.com/weo/V029.htm)

Population data for Afghanistan, Bermuda, Costa Rica, Cuba, Montenegro, Nicaragua, North Korea, Puerto Rico, Somalia, Serbia, Suriname, Trinidad and Tobago are based on various sources including the Central Intelligence Agency's World Factbook.

Distance

For any two points on a globe, identified by the latitude and longitude points, we have:

$$h = haversin\left(\frac{d}{R}\right) = haversin(lat_1 - lat_2) + cos(lat_1)cos(lat_2)haversin(lon_1 - lon_2), \quad \text{(A1)}$$

where $haversin(x) = sin^2(x/2)$ is the haversine function, d is the spherical distance, R is the radius of the sphere (for the earth we use R = 6367 km), lat_i is the latitude of point i=1,2, and lon_i is the longitude of point i=1,2. From this equality we can solve for the distance using the inverse sine (*arcsin*):

$$d = 2R \times arcsin\left(\sqrt{h}\right) \quad \text{(A2)}$$

This formula gives the shortest distance between two points on a sphere from their longitudes and latitudes. A source for the distance in kilometers to the host city for the Games is, for instance, Map Crow.

Home advantage

Finally the home dummy to measure the home advantage effect of hosting the Games is coded as follows: 1 for host country, 0 otherwise.

List of research reports

12001-HRM&OB: Veltrop, D.B., C.L.M. Hermes, T.J.B.M. Postma and J. de Haan, A Tale of Two Factions: Exploring the Relationship between Factional Faultlines and Conflict Management in Pension Fund Boards

12002-EEF: Angelini, V. and J.O. Mierau, Social and Economic Aspects of Childhood Health: Evidence from Western-Europe

12003-Other: Valkenhoef, G.H.M. van, T. Tervonen, E.O. de Brock and H. Hillege, Clinical trials information in drug development and regulation: existing systems and standards

12004-EEF: Toolsema, L.A. and M.A. Allers, Welfare financing: Grant allocation and efficiency

12005-EEF: Boonman, T.M., J.P.A.M. Jacobs and G.H. Kuper, The Global Financial Crisis and currency crises in Latin America

12006-EEF: Kuper, G.H. and E. Sterken, Participation and Performance at the London 2012 Olympics www.rug.nl/feb