Forecasts and evaluation of the 2011 Pan American Games

GERARD H. KUPER, ELMER STERKEN[1]

- **Abstract**: The current paper predicts the medal tally for the 2011 Pan American Games. The forecast procedure consists of analyzing success at the four latest editions of the Pan American Games at the country level. Potential explanatory variables for medal winning are GDP, population, geographical distance to the Games and home advantage. Our forecasts show that the US takes first place in the medal tally. We expect Mexico to take fourth place wining 37 gold medals. The final results as they were published after the Games prove us right in this respect. However, the forecasts are less precise than similar predictions for the Olympic Games.

- **Resumen**: Este ensayo pronostica el número de medallas para los Juegos Panamericanos del 2011. El procedimiento del pronóstico consiste en analizar a los ganadores de las últimas cuatro ediciones de los Juegos Panamericanos por país. Las posibles variables explicativas de las medallas ganadoras son PIB, población, distancia geográfica con respecto al sitio donde se celebran los Juegos y la ventaja del país sede. Nuestro pronóstico muestra que Estados Unidos obtendrá el primer lugar en el total de medallas ganadas. Esperamos que México ocupe el cuarto lugar obteniendo 37 medallas de oro. Los resultados finales publicados después de los Juegos prueban nuestras predicciones. Sin embargo, los pronósticos son menos precisos que otras predicciones similares hechas para los Juegos Olímpicos.

- **Keywords:** Pan American Games, Medal predictions, Panel data models.

- **JEL Classification:** R0, O1, C23, Z10.

- Recepción: 02/12/2011 Aceptación: 07/03/2012

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

[1] Department of Economics, Econometrics, and Finance, University of Groningen.
Corresponding author: Gerard H. Kuper email: g.h.kuper@rug.nl
Acknowledgements: We would like to thank the two anonymous referees for constructive comments and suggestions. We thank Solmaria Halleck Vega and Javier López Arnaut for translating the abstract and one of the referee reports.

honour 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. The ancient Olympic Games were abolished in 393 AD 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, 2008). Similar to the ancient Olympic Games, the Games were held every four years (the Olympiad). The main purpose was to foster the ideal of "…a sound mind in a sound body…" and to promote friendship among nations.

In 1925, the International Olympic Committee (IOC) encouraged the establishment of regional Games (Olderr, 2003). One year later, the Central American Games were organized and, in 1936, the concept of Pan American Games was approved. The official start of the Pan American Games, however, was delayed due to the Japanese attack on Pearl Harbor and the Second World War. The first Pan American Games were held in 1951 in Buenos Aires (Argentina). In 2011, the sixteenth Pan American Games have been organized in Guadalajara (Mexico).

This paper presents forecasts for medal winnings at the 2011 Pan American Games. 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 first time that we apply the methodology we have developed for forecasting the medal tallies for the Olympic Winter and Summer Games since 2002, for the Olympic Games to the Pan American 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 success at the most recent editions of the Pan American Games. Our goal is to investigate the role of key determinants such as population size, income per head, distance and home advantage in determining success. Before we discuss the methodology, the data and the econometric model, we first present a brief history of the Pan American Games. In Section 3 we give an overview of related work. In Section 4, the determinants for success are discussed, and Sections 5 and 6 present the model and the forecasts. Section 7 evaluates our forecasts for the 2011 Pan American Games. We summarize our findings in the last section.

■ *A brief history of the Pan American Games*

There are similar developments between the Olympic Games and the Pan American Games, but there are also differences. Similarities relate to the size of the event, and the issues like politics and drugs. A difference is the scope of the event and the fact that organizing Winter Pan American Games failed because of the lack of American support (Olderr, 2003). Over the years, the Olympic Games grew in size, in terms of number of sports, athletes and countries participating. A similar development is seen for the Pan

American Games as is illustrated in Table 1. The number of countries and athletes participating doubled since the first Pan American Games in 1951. The number of sports increased from 18 in 1951 to 46 in 2011.

Table 1
An overview of all editions of the Pan American Games

Location				Number		
Edition	City	Country	Dates	Countries	Athletes	Sports
I	Buenos Aires	Argentina	25/02 - 09/03 1951	21	2513	18
II	Mexico City	Mexico	12/03 - 26/03 1955	22	2583	17
III	Chicago	United States	27/08 - 07/09 1959	25	2263	18
IV	São Paulo	Brazil	20/04 - 05/05 1963	22	1665	19
V	Winnipeg	Canada	23/07 - 06/08 1967	29	2361	19
VI	Cali	Colombia	30/07 - 13/08 1971	32	2935	17
VII	Mexico City	Mexico	13/10 - 26/10 1975	33	3146	19
VIII	San Juan	Puerto Rico	01/07 - 15/07 1979	34	3700	22
IX	Caracas	Venezuela	14/08 - 29/08 1983	36	3426	22
X	Indianapolis	United States	08/08 - 23/08 1987	38	4453	27
XI	Havana	Cuba	02/08 - 28/08 1991	39	4519	34
XII	Mar del Plata	Argentina	11/03 - 26/03 1995	42	5144	33
XIII	Winnipeg	Canada	24/07 - 08/08 1999	42	5275	34
XIV	Santo Domingo	Dominican Republic	01/08 - 17/08 2003	42	5325	40
XV	Rio de Janeiro	Brazil	12/07 - 29/07 2007	42	6035	41
XVI	Guadalajara	Mexico	14/10 – 30/10 2011	42	5932	46

Source : http://en.wikipedia.org/wiki/2011_Pan_American_Games

Olderr (2003) presents a chronological overview of the Pan American Games until and including the Games of 1999. Here we focus on some of the issues that plagued the Pan American Games. The first issue relates to politics. In 1951, Juan Perón used the Games for propaganda purposes in the same way Hitler did during the 1936 Olympics. In 1971, Cuba emerged as a sporting power, but four of its athletes defected while another one died in suspicious circumstances. Again, in 1999, 13 Cuban athletes defected.

Table 2
All-time medal count of the Pan American Games before 2011 Guadalajara

Rank	Nation	Gold	Silver	Bronze	Total	Rank	Nation	Gold	Silver	Bronze	Total
1	United States	1747	1295	873	3915	4	Argentina	258	279	363	900
2	Cuba	781	531	481	1793	5	Brazil	239	283	401	923
3	Canada	348	547	682	1577	6	Mexico	157	217	409	783
						7	Venezuela	73	156	224	453

Rank	Nation	Gold	Silver	Bronze	Total
8	Colombia	57	109	162	328
9	Chile	37	70	108	215
10	Puerto Rico	21	72	113	206
11	Jamaica	21	33	59	113
12	Dominican Republic	19	43	85	147
13	Ecuador	14	13	36	63
14	Uruguay	11	22	42	75
15	Trinidad and Tobago	8	17	25	50
16	Guatemala	7	12	29	48
17	Bahamas	6	11	9	26
18	Peru	5	28	58	91
19	Netherlands Antilles	4	9	16	29
20	Costa Rica	4	6	10	20
21	Suriname	4	2	5	11
22	Panama	3	20	24	47
23	Guyana	2	4	11	17
24	El Salvador	1	6	12	19
25	Bermuda	1	4	3	8
26	Antigua and Barbuda	1	0	3	4
27	United States Virgin Islands	0	4	5	9
28	Barbados	0	3	7	10
29	Nicaragua	0	3	7	10
30	Cayman Islands	0	3	0	3
31	Haiti	0	2	5	7
32	Paraguay	0	1	6	7
33	Honduras	0	1	4	5
34	Bolivia	0	1	2	3
35	Grenada	0	1	2	3
36	Dominica	0	1	1	2
37	Belize	0	0	2	2
38	Saint Lucia	0	0	2	2
39	Aruba	0	0	1	1
40	Saint Vincent and the Grenadines	0	0	1	1
41	British Virgin Islands	0	0	0	0
42	Saint Kitts and Nevis	0	0	0	0
Totals		3496	3477	3875	10848

Source: http://en.wikipedia.org/wiki/Pan_American_Games

Another issue is the US dominance. Except for the first edition of the Pan American Games, the US dominates the Games; both in terms of the number of athletes as well as the number of medals won by US athletes (see Table 2). According to Avery Brundage - the former president of the IOC (1952 to 1972) and the first president of the Pan American Sports Organization (PASO) - this may have hurt the development of the Games and, in 1971, he called for the US to send second tier athletes. This proposal was met with criticism. Nevertheless, the US did and still does send lesser talented athletes to the Pan American Games, but probably for different reasons. A possible side effect of US dominance has been the hostility of the crowd toward US athletes in the 1975 (Mexico City) opening ceremonies and medal ceremonies. Obviously, the Pan American Games also have to deal with doping. Drug tests are used for the first time at the 1983 Pan American Games in Venezuela. Ten medalists were found to be taking illegal drugs.

Before we present the data and the econometric models, we review some of the literature on the Olympic Games that discusses the interaction with economic and political developments.

■ *Literature*

There is a huge literature on the Olympic Games and its interaction with economic and political developments. Literature on the Pan American Games is scarce, but what holds for the Olympic Games, to a lesser extent, is also probably true for the Pan American Games. We discuss some of these interactions of the Olympic Games 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 mainly in 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, 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. Also the home country may introduce new sports. For instance, Brazil introduced futsal in 2007 and Mexico introduces basque pelota and racquetball in 2011. 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.

■ *Determinants of 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. However, for the Pan American Games, we lack historical data on participation per country. Also using World Championship results is not meaningful, for various reasons. One reason is that the Pan American Games are regional Games, whereas the Olympic Games and the World Championships are not. A second and related reason is that many of the athletes who participate at the World Championship also enter the Olympic Games. This is not true for the Pan American Games for reasons discussed in Section 2. So, our model for success at the Pan American Games is a simplified version of our model for the Olympic Games. Hence, it remains to be seen whether our forecast performances carry over to the Pan American Games. This is discussed in Section 7.

We apply econometric models to quantify and identify determinants of success at the Pan American Games. These determinants are based on the literature and our experience in predicting success at the Olympic Games. Below we present the determinants. For definitions and sources, we refer to Appendix A. The forecasts are based on the last four editions of the Pan American Games.

The dependent variable is success measured by medals won (for gold, silver and bronze) as a fraction of the total number of gold, silver and bronze medals awarded. We calculate the shares of gold, silver and bronze medals won for the Games since the 1995 Pan American Games in Mar del Plata (Argentina). Because we are modeling medal shares, also the main determinants, like income and population, are included as shares in regional income and regional population.

The explanatory variables are:

I. *GDP share in regional GDP*. As an approximation for wealth, we use a four year average (three years before the event and the year of the event) of GDP for the period 1980-2011 as a share in regional GDP. Regional GDP is defined as the total GDP of the 42 countries that form the Pan American region.
II. *Population share in regional population*. The population size of participating countries also matters since bigger countries send more athletes. Again we calculate four year averages (three years before the event and year of the event) of the share of the population of the participating country in regional population (defined as the total population of 42 countries in the Pan American region) for the period 1980-2011.
III. *Distance*. We also analyze the impact of geographical distance of participating countries to the country that organizes the Games. This is an indication for travel costs. More specifically, we use data on the latitude and longitude of the cities that organized the Games. We use the so-called Haversine formula (Sinnott, 1984) to compute the distance between the capital city of the participating country and the host city as "the crow flies."
IV. *Home advantage*. Finally, the home dummy that measures the home advantage effect of hosting the Games is coded as follows: 1 for host country, 0 otherwise.

Table 3
Descriptive statistics

	Medal shares			Shares in		Distance
	Gold	Silver	Bronze	GDP	Population	
Mean	0.024	0.024	0.024	0.024	0.024	4392
Median	0.000	0.002	0.003	0.001	0.004	4730
Maximum	0.394	0.336	0.216	0.787	0.347	10219
Minimum	0.000	0.000	0.000	1.8E-05	2.4E-05	0.000
Std. Dev.	0.065	0.055	0.045	0.117	0.062	2240
Skewness	3.678	3.475	2.382	6.106	3.976	0.022
Kurtosis	16.905	16.582	8.296	38.862	19.014	2.569
Observations	168	168	168	168	168	168

Source: Own calculations.

Table 3 shows the descriptive statistics for the series in our model for all 42 countries and the last four editions of the Pan American Games. In total, there are 168 observations. It should be noted that by construction the shares add up to one. So the means for 42 countries are the same (1/42). The median country in the sample wins no gold medals; this means that fewer than 21 countries win at least one gold medal. The medal share series are skewed and exhibit excess kurtosis. The all-time medal count (in Table 2) shows that the Top-5 countries won 96% of all gold medals, and 84% and 72% of all silver and bronze medals, respectively. Also the GDP share and the population share series are skewed and show excess kurtosis. The median distance to the 2011 Pan American Games is 4,730 kilometers with a maximum of 10,219 kilometers (Argentinian athletes travelling to Winnipeg in Canada in 1999), and a minimum of 0 kilometers for home athletes.

■ *Modeling success*

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 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 cases is more sensitive to measurement errors. Thirdly, success is mostly discussed at the country level.

The determinants for national success are demographic (population), economic (income), and geographic (distance to the host country) in nature. Also home advantage may determine success. These determinants are predetermined, and there is no endogeneity bias. The distance to the Games translates into travelling costs, which could also be considered as an economic component. The main argument why eco-

nomic welfare is important in explaining Olympic welfare 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 model the national shares in medal totals, GDP and population. 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. An 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, at the Mar del Plata Games of 1995, a total of 432 gold medals were awarded, whereas in Guadalajara 2011, the number of gold medals was 361. Also 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 are awarded.

The main determinants are income and population as shares in regional income and regional population. 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_{it} will be of the order $(\log N_{it})^{1/2}$. In our model, we use population shares n_{it} and we approximate $\log(1+n_{it})$ with n_{it} (first-order Taylor series approximation). However, this argument is valid for standard normal series, and population shares are not normally distributed. Nevertheless, in this paper we use the square root of population shares as explanatory variable because experimenting with other specifications in earlier forecasts yields similar estimation results and forecasts.

Next, we assume that income Y_t will determine the training, access to training facilities, and health conditions of the potential athletes. We expect the population share and income share to have a positive effect on success. The potential share of athletes p_{it} is also affected by the home advantage and the geographical distance to the Games. The home advantage H_{it} is a dummy variable (1 if country i host the Games t, and 0 in other cases). Home countries are allowed to send more athletes. We measure the distance to the Games D_{it} as "the crow flies" (see Appendix A for details).

We model the share of medals (gold, silver, and bronze) at the Pan American Games o_m, with m = Gold, Silver, and Bronze, as a function of population share n, income share Y, the home advantage H, and distance D.

$$o_{mit} = c_1 n_{it}^{1/2} + c_2 Y_{it} + c_3 H_{it} + c_4 D_{it} + c_i + e_{it}, \qquad m=G, S, B \tag{1}$$

where e_{it} is a white noise residual. The country specific effects c_i represent factors that are constant over time but differ across countries. Countries that were successful at the previous Games probably are successful at the next edition as well.

Table 4 presents the results for the model with country-specific fixed effects for the Pan American Games. The sample is a balanced panel including the four most recent

Table 4
Medal counts for the Pan American Games with fixed effects for countries (robust standard errors are in brackets). The fixed effects are not reported

Dependent variable: percentage medal share o_m (m=G, S, B,).
Explanatory variables:

H = 1 if a country hosts the Pan American Games, else 0;
Y = GDP share of a country in total regional GDP;
n = country population share of the total regional population;
D = distance from the capital of the host country to the capital of the participating country.

	Gold	Silver	Bronze		Gold	Silver	Bronze
$n^{½}$	0.391	1.157	1.004	R^2	0.953	0.949	0.993
	(0.168)	(0.275)	(0.267)	Countries	42	42	42
Y	0.640	-0.351	0.147	Observations	168	168	168
	(0.508)	(0.616)	(0.086)	F-test	55.168	50.611	367.328
D/1000	1.47E-05	2.04E-05	-3.58E-06	(p-value)	<0.001	<0.001	<0.001
	(2.27E-05)	(1.12E-05)	(3.24E-05)	Hausman test $\chi 2$	25.707	66.017	32.421
H	0.054	0.026	0.061	(p-value)	<0.001	<0.001	<0.001
	(0.015)	(0.004)	(0.008)				

Source: Own calculations.

editions (from Mar del Plata in 1995 to, and including, Rio de Janeiro in 2007) and includes all 42 countries listed in Table 2. The explanatory variables are not highly correlated (the biggest correlation coefficient, between the square root of population share and income share, is 0.71). The fixed effects estimator is consistent (the Hausman test rejects the null of uncorrelated effects). The main results are:

I) The population share has a positive effect on success: For a country that is 100 times bigger, the medal shares are 2 to 6 percentage points higher.
II) The income share has no effect on winning gold and silver medals; only bronze medal winning is higher: If income share is 10 percentage points higher, the country's share in bronze medals is only 1.5 percentage points higher (significantly different from zero at a 5% significance level in a one tailed test).
III) The hypothesis that a bigger distance from the home country to the country that organizes the Games reduces success is not supported by the estimates.
IV) Home advantage is significant and has the expected positive sign: organizing the Pan American Games results in an increase in the medal share of 2.6-6.1 percentage points.

The model is highly significant as the F-test indicates. The measure of fit (R^2) indicates that over 95% of the variation of the medal share is explained by the determinants.

■ *Forecasting success*

The estimates presented above are used to forecast medal winning at the 2011 Pan American Games in Guadalajara in Mexico.

Athletes from 42 countries competed in 46 sports, with 361 medal events, including sports that are not (or not anymore) on the list of Olympic sports, like bowling, basque pelota, racquetball, roller skating, squash, water skiing, and rugby sevens. Two bronze medals will be awarded for 63 events in boxing, judo, wrestling, taekwondo and karate.

Table 5
Medal forecasts for the 2011 Pan American Games

Rank	Gold	Silver	Bronze	Total
1 United States	103	105	68	276
2 Cuba	75	40	42	157
3 Canada	42	47	53	142
4 Mexico	37	29	55	121
5 Brazil	34	29	45	108
6 Argentina	16	21	30	67
7 Venezuela	14	20	28	62
8 Colombia	11	14	21	46
9 Chile	4	6	9	19
10 Guatemala	3	6	8	17
11 Puerto Rico	3	4	9	16
12 Ecuador	3	3	7	13
13 Jamaica	3	3	3	9
14 Dominican Republic	1	5	7	13
15 Peru	1	4	8	13
16 Bolivia	1	2	3	6
Honduras	1	2	3	6
18 Bahamas	1	2	2	5
Costa Rica	1	2	2	5
Nicaragua	1	2	2	5
Panama	1	2	2	5
Paraguay	1	2	2	5
23 Uruguay	1	1	2	4
24 Suriname	1	1	1	3
25 Netherlands Antilles	1	0	1	2
26 El Salvador	0	1	2	3
Haiti	0	1	2	3
Trinidad and Tobago	0	1	2	3
29 Belize	0	1	1	2
Cayman Islands	0	1	1	2
31 Bermuda	0	1	0	1
Virgin Islands	0	1	0	1
33 Antigua and Barbuda	0	0	1	1
Barbados	0	0	1	1
Grenada	0	0	1	1
Saint Lucia	0	0	1	1
37 Aruba	0	0	0	0
Dominica	0	0	0	0
Guyana	0	0	0	0
British Virgin Islands	0	0	0	0
Saint Kitts and Nevis	0	0	0	0
Saint Vincent and the Grenadines	0	0	0	0

Source: Own calculations.

The forecasts presented in Table 5 indicate that:

- In total, 30 countries win at least two medals and six countries are expected to win one medal.
- The Top-10 countries win 94% of the gold medals, and 89% of the total number of medals.
- The first eight countries are the same countries that led the medal tallies since 1995.
- The United States wins the medal count with Cuba in second place as in the four previous editions.
- Canada is again in third place, which is one place up from the 2007 results.
- Mexico takes advantage of the home advantage by winning 37 gold medals. This is about twice as much as in Rio de Janeiro in 2007.
- Brazil drops from third to fifth place.
- Argentina gradually has moved down the medal table from fourth place in 1995 (Mar del Plata) to eight place in 2007. In 2011, they are expected to be back in sixth place.

In the next section, we compare the expected results with the final results at the Pan American Games to identify over performing and underperforming countries.

■ *Evaluation of the 2011 Pan American Games*

Table 6 shows the results of the 2011 Pan American Games in Guadalajara. The results account for the failed drug test of Canadian wakeboarder Aaron Rathy. He was disqualified and his silver medal went to Marcelo Giardi of Brazil, and the bronze medal to Alejo de Palma of Argentina.

The Cayman Islands won its first ever gold medal (Brett Fraser won the men's 200m freestyle event), while Kim Collins, at age 35, won the silver medal in the men's 100m final, which is Saint Kitts and Nevis' first ever Pan American Games medal.

The results in Table 6 show that 29 countries won at least one medal. It also shows that the Top-10 countries won 94% of the gold medals, and 89% of the total number of medals. This is exactly as predicted (see Table 5).

Table 6
Results for the 2011 Pan American Games

Rank	Country	Gold	Silver	Bronze	Total
1	United States of America	92	79	65	236
2	Cuba	58	35	43	136
3	Brazil	48	35	58	141
4	Mexico	42	41	50	133
5	Canada	30	40	49	119
6	Colombia	24	25	35	84
7	Argentina	21	19	35	75
8	Venezuela	12	27	33	72
9	Dominican Republic	7	9	17	33
10	Ecuador	7	8	9	24
11	Guatemala	7	3	5	15
12	Puerto Rico	6	8	8	22

Rank	Country	Gold	Silver	Bronze	Total
13	Chile	2	17	24	43
14	Jamaica	1	5	1	7
15	Bahamas	1	1	1	3
	Cayman Islands	1	1	1	3
17	Netherlands Antilles	1	0	1	2
18	Costa Rica	1	0	0	1
19	Uruguay	0	3	2	5
20	Peru	0	2	5	7
21	Trinidad & Tobago	0	2	2	4
22	Saint Kitts & Nevis	0	2	0	2
23	El Salvador	0	1	0	1
24	Barbados	0	0	2	2
	Bolivia	0	0	2	2
	Paraguay	0	0	2	2
27	Dominica	0	0	1	1
	Guyana	0	0	1	1
	Panama	0	0	1	1
Totals		361	363	453	1177

Source: http://www.guadalajara2011.org.mx/sports

Mexico finished in fourth place, as we predicted, but it wins more gold medals than expected. We conclude that Mexico performed better than expected, partly because of the re-introduction of racquetball and basque pelota at the Pan American Games. In these two sports, Mexico won 10 out of a possible 16 gold medals.

Compared to our forecasts, Brasil (3rd place), Colombia (6th place) and the Dominican Republic (9th place) performed better than expected, while Canada (5th place) and Chile (13th place) underperformed.

Table 7
Our forecasting performance of the 2011 Pan American Games

	Gold	Silver	Bronze
Mean Absolute Error	2.60	3.10	2.57
Mean Squared Error	26.02	32.63	20.83

Source: Own calculations.

If we compare our forecast performance in terms of the Mean Absolute Error and the Mean Squared Error (Table 7), with our performance for the Olympic Games (Appendix B), we have to conclude that our overall predictions for the Pan American Games are worse. This is mainly caused by the United States of America winning 40 medals less than expected and Colombia winning 38 medals more than expected. One reason for the lower forecast performance is that it is difficult to foresee whether a country sends their best athletes, or their 2nd tier athletes. This may probably hold true for countries like the United States of America with an abundance of top athletes. Another reason is that we have used a stripped down version of our model for the Olympic Games. This means that the forecasts for the Pan American Games are based on smaller information set due to lack of available data. So, the lower forecast performance for the Pan American Games does not surprise us.

■ *Summary and conclusion*

Our paper is the first to describe the impact of demographic, economic and geographical conditions on medal winning success at the Pan American Games. Our main quest is to establish the determinants of medal winning. We show that the population size and the home advantage are important determinants for medal success. The effect of income is marginal.

Also, we are the first to forecast medal winnings at the 2011 Pan American Games in Guadalajara. Our forecasts indicate that the US leads the table with Mexico in fourth place taking a huge advantage of the home advantage effect. This home advantage effect is also reflected in the final results of the Pan American Games.

The forecasting performance of success at the Pan American Games is lower than similar forecasts we have published for the Olympic Games. The forecasts can be improved if historical data about the number of athletes for each country that have participated at the earlier editions of the Pan American Games is made available.

■ *Appendix A – Data sources*

The dependent variables are *Medal shares.*
The main source for gold, silver and bronze medals is the official pages of Guadajara 2011 Pan American Games.

The explanatory variables are:

I. *GDP share in regional GDP*
The definitions and the main sources the GDP share in regional GDP for the period: 1980-2011 are as follows:
- Gross domestic product, current prices; U.S. dollars; Billions; including estimates for 2009-2015 from International Monetary Fund (2010).
  For AHO, ARU, BER, CAY, CUB, ISV, IVB, PUR (see the country codes below) the data are from the United Nations Statistics Division (UN).
- For AHO, ARU, BER, CAY, CUB, IVB, PUR, we extrapolated the shares for 2010-2011

II. *Population share in regional population*
The definitions and the main sources the population share in regional population for the period 1980-2011 are as follows:
- Population, million persons, including estimates for 2009-2015, are from the International Monetary Fund (2010).
  For AHO, ARU, BER, CAY, CUB, ISV, IVB, PUR, we use United Nations (UN) population growth rates for the period 2005-2010.
- Early 1990s data for CRC, DMA, NCA, SUR, TRI are from the U. S. Census Bureau, International Data Base (U. S. Census Bureau).
- For TRI, we extrapolated population shares for 2010-2011

III. *Distance*

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

(B1) $h = \text{haversin}(d/R) = \text{haversin}(lat_1\text{-}lat_2)+\cos(lat_1)\cos(lat_2)\text{haversin}(lon_1\text{-}lon_2)$, where $\text{haversin}(x) = \sin^2(x/2)$ 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):

(B2) $d = 2R \arcsin(\sqrt{h})$.

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.

IV. *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.

Contry codes

Country	Code	Country	Code
Netherlands Antilles	AHO	Guatemala	GUA
Antigua and Barbuda	ANT	Guyana	GUY
Argentina	ARG	Haiti	HAI
Aruba	ARU	Honduras	HON
Bahamas	BAH	Virgin Islands	ISV
Barbados	BAR	British Virgin Islands	IVB
Bermuda	BER	Jamaica	JAM
Belize	BIZ	Saint Lucia	LCA
Bolivia	BOL	Mexico	MEX
Brazil	BRA	Nicaragua	NCA
Canada	CAN	Panama	PAN
Cayman Islands	CAY	Paraguay	PAR
Chile	CHI	Peru	PER
Colombia	COL	Puerto Rico	PUR
Costa Rica	CRC	Saint Kitts and Nevis	SKN
Cuba	CUB	Suriname	SUR
Dominica	DMA	Trinidad and Tobago	TRI
Dominican Republic	DOM	Uruguay	URU
Ecuador	ECU	United States	USA
El Salvador	ESA	Venezuela	VEN
Grenada	GRN	Saint Vincent and the Grenadines	VIN

■ *Appendix B – Evaluation of Olympic Games forecasts*

Table B1 summarizes our forecasting performance for the Olympic Winter Games of Turin in 2006 and Vancouver 2010, and the Olympic Summer Games of Beijing in 2008. The table reports two measures of forecast performance. The mean absolute error indicates that, on average for all countries in the samples, our predictions are off by about one gold medal, one silver medal, and one bronze medal. The mean squared error penalizes big deviations from the realizations more severely, but is not easy to interpret. We also 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. There are differences with our method. 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. Despite these differences, we sometimes outperform SI, as illustrated in Table B1.

Table B1
Our forecasting performance (KS) at the three most recent Olympic Games compared with those of Sports Illustrated (SI)

	Gold		Silver		Bronze	
	KS	SI	KS	SI	KS	SI
Mean Absolute Error						
Turin 2006	0.91	1.36	0.73	0.59	0.95	1.00
Beijing 2008	0.87	0.86	0.92	1.15	1.24	1.23
Vancouver 2010	1.05	0.63	1.24	1.03	1.18	1.18
Mean Squared Error						
Turin 2006	4.14	7.45	2.41	1.55	2.64	3.65
Beijing 2008	3.22	3.77	2.60	4.31	6.24	5.23
Vancouver 2010	3.32	1.16	3.08	2.55	3.34	3.61

Source: Own calculations.

■ *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 Development and Medal Totals", *Review of Economics and Statistics*, 86, 413-417.

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

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

Guadajara Pan American Games, http://www.guadalajara2011.org.mx/home

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

Johnson, D. K. N., Ali, A. (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., Sterken, E. (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.

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

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

Olderr, A. (2003). "The Pan American Games: A Statistical History, 1951-1999, Los Juegos Panamericanos: Una Historia Estadística, 1951-1999", McFarland, Jefferson (North Carolina).

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

Sterken, E. (2006). "Growth Impact of Major Sporting Events", *European Sport Management Quarterly*, 6, 375-389.

UN, United Nations Statistics Division, http://unstats.un.org/unsd/snaama/dnlList.as U.S. Census Bureau, International Data Base, http://www.census.gov/ipc/www/idb/

Wallechinsky, D., Loucky, J. (2008). "The Complete Book of the Summer Olympics: 2008 Edition", Aurum, London.