Predicting Olympic Medal Counts: the Effects of Economic Development on Olympic Performance

Article · January 2005

Source: OAI

1 author:

Xun Bian
University of North Texas

Predicting Olympic Medal Counts: the Effects of Economic Development on Olympic Performance

Xun Bian

I. INTRODUCTION

The modern Olympics were conceived by their founder Pierre de Coubertin to be competition between individual athletes, not countries (IOC, 2000). The Olympic Spirit emphasizes participation rather than winning. In reality however, the success of a country's athletes is held to be an important source of national prestige. By-country medal tables are widely published. A glance at Olympic history will immediately tell us that not all nations have an equal ability to win medals. In this past August, 199 countries participated in the Athens Olympics, and 124 countries did not win a single medal. On the other hand, the top ten winners collectively took home 514 medals, more than 50% of the medals available at the Athens Olympics. Therefore, a natural question to ask is why some countries are able to enjoy a great success in the Olympic arena, while others never have the chance to do so.

"The success of a country's athletes is held to be an important source of national prestige."

The unequal distribution of Olympic Medal numbers might be explained by the relative strength of countries in different sports. For example, with a large number of high-quality basketball players, the United States should have a higher probability of winning a medal in basketball. We could then generate a prediction for a national medal total by a summation across sports. However, this paper takes a different perspective and attempts to predict a nation's Olympic performance by investigating the socioeconomic variables that have significant influence on a nations' Olympic performance. The influence of population size, economic resources, political and economic structure, and hosting advantage are estimated by using two different models.

The paper is organized in the following structure. Section II introduces the theoretical framework of the research and reviews previous literature related to this topic. The empirical model and data used to test the research hypothesis are described in detail in Section III. Section IV presents the regression results. Finally, Section V concludes the research by discussing possible policy implications.

II. THEORETICAL FRAMEWORK AND REVIEW OF LITERATURE

Starting from the post-World War II games, sociologists and economists began to analyze the impact of social and economic conditions on the number of Olympic medals won by different countries. Examples of those studies are Ball (1972), Grimes et al, (1974) and Levine (1974). Those early studies showed that population, income per capita, hosting advantage, and political system have a significant impact on a nation's medal counts. First, population is one of the fundamental determinants of Olympic success. A large population increases the group of potential athletes. As we can see, China wins more medals than most other nations, because having 1.3 billion people improves the odds of producing a Yao Ming. The second determinant is economic resources. Richer countries can usually afford to train athletes better, provide better medical care, and send a larger group of athletes to the Olympic Games. Hosting advantage is also significant. The hosting country is allowed to participate in all events. In addition, the crowd of home spectators will support the performing athletes. More resources are likely to be devoted to training in preparation for the game that will attract so much attention within the home country. The fourth determinant is political and economic structure. There is a large amount of evidence suggesting that communist countries perform better. This is probably because a centrally-planned economic system allows more specialization in sports, and more resources can be distributed to training and supporting athletes than in market-based economies. Moreover, the governments of communist countries not only have better capability to channel economic resources to sports but, also have a stronger incentive to do so. Since Olympic performance is so closely connected with national prestige, winning a large number of Olympic medals can help them obtain recognition internationally as well as stimulate patriotism domestically. Without having a democratic political system, international recognition and patriotism are extremely valuable to the government for maintaining political stability.

Surprisingly enough, the literature that models Olympic performance did not develop until the 1990s. An explanation of this might be that in the 1970s and 1980s the Olympic Games were disrupted by the Cold War. The first study that restarts the performance analysis is Slughart et al (1993), which analyzes the Olympic performance of transitional economies. Recently, two studies by Johnson and Ali (2000) and Bernard and Busse (2000) revived attention to this issue. Johnson and Ali (2000) assume the medal counts to be a linear function of GDP per capita, population, and two dummy variables indicating host country and political system respectively. They find that the home advantage adds a 12 percent chance of success, and communist countries outperform the others by 12 medals (5 gold medals).

Bernard and Busse (2000) estimate probit models for medal shares using data since 1960. They specify a Cobb-Douglas production function for medal shares, using population and economic resources (measured in GDP) as production factors. By specifying a Cobb-Douglas production function form, Bernard and Buss assume that both population share and economic resources should be subject to diminishing marginal returns. This assumption does make economic sense. Holding economic resources constant, additional talented athletes will inevitably decrease the fund attributed to each person, and some athletes might not be able to obtain the training conditions that are necessary for them to fully exert their potential. Therefore, the marginal contribution of population growth to the Olympic medal winning process tends to decline as the population size gets bigger. Conversely, holding population constant, additional economic resources attributed to sports should also yield diminishing returns, as more athletes deplete their potential. As we move down the list of athletes, we encounter more less talented athletes. Spending economic resources in training those average athletes will not produce any Olympic medals. In an extreme case, once all the talented athletes who are capable of competing for Olympic medals in a country reached their physical limits by having ideal training conditions, additional funding would not increase the country's Olympic medal share at all. In addition, Bernard and Busse also include a dummy variable for the hosting advantage, a soviet dummy, and a non-soviet but planned economy dummy. The hosting advantage is estimated to be 1.2 percentage point of the medal share. The effect of soviet dummy varies between 3-6 percentage points.

"Population, income per capita, hosting advantage, and political system have a significant impact on a nation's medal counts."

Some of the most recent studies go beyond medal counts and argue that not all Olympic medals are alike, and countries with different characteristics specialize in different sports. Tcha and Pershin (2003) investigates each country's performance and attempt to identify the determinants of this performance in each sport, while examining other issues related to specialization at these games, using the concept of revealed comparative advantage (RCA). Each country's RCA is explained by geographical, biological, and economic variables of the participating countries.

The analyses present the determinants of each country's specialization in sports and patterns of RCA, which are substantially different from those obtained by analyzing medal total. The authors found that high-income countries specialize less; in other words, they win medals in a more diversified range of sports.

"High-income countries specialize less; in other words, they win medals in a more diversified range of sports."

This paper will follow the two most recent studies on modeling national Olympic performance and using both the linear function and the Cobb-Douglas production function to estimate the influence of population size, economic resources, political and economic structure, and hosting advantage on nations' Olympic performance. Based on the results of previous studies, I expect population size and economic resources are positively correlated with a country's medal share, and being a socialist country or a hosting country increases a country's medal share. Moreover, if the diminishing marginal return of population size and economic resources indeed exist, the Cobb-Douglas production function should generate a better prediction than the simple linear function.

III. EMPIRICAL MODEL AND DATA

Model one uses a linear function which is easy to construct and interpret. The coefficient of each variable represents the marginal effect of that particular variable on Olympic medal counts. The actual model estimated is shown by the following equation (1):

$$M_t = C + \alpha_1 N_t + \alpha_2 (Y_t / N_t) + \alpha_3 P + \alpha_4 H_t + \varepsilon$$

M_t denotes the medal number for a country at a particular Olympic Game. In this research I do not distinguish between gold, silver, and bronze medals, because the difference between the best and the second best is usually so minute that the rank of medalists depends more on luck rather than athletic talents. N_t is the population size of the country at the year t when a particular Olympic Game is held. Y_t denotes the GDP of the country at the same Olympic year. Y_t / N_t is therefore the per capita GDP of the country at the Olympic year. P and H_t are dummy variables for political and economic structure and hosting countries respectively. P takes the value 1 if the country has socialist background, which means the country is or was a socialist country, and it takes 0 if otherwise. Similarly, if the country is hosting the Olympics in that year, H_t takes the value of 1, and 0 if otherwise. For the research hypothesis to be true, the coefficients of all independent variables need be positive.

The second model follows the same notations as Model 1, but uses the Cobb-Douglas production function. It views the medal winning process similar to a production process, and the two key factors are population size (N) and economic resources (Y), which are both subject to diminishing marginal return. Hence, the medal winning process could be modeled in the following way:

$$M_t = A_t (N_t)^{\gamma} (Y_t)^{\theta} \qquad (2)$$

Equation 2 indicates the production of talented athletes requires people (N), economic resources (Y), and some other influential factors, which are captured by A as a whole. One important property of equation 2 is that increases in medals should be less than one-to-one in both population and economics resources. Hence, γ and θ should be both positive and less than one. By taking natural log of both sides of equation 2, I yield the following specification for Olympic medal counts:

$$\ln M_t = \ln A_t + \gamma \ln N_t + \theta \ln Y_t + e \qquad (3)$$

Since A_t captures other aspects that are influential on a country's Olympic performance, we can replace $\ln A_t$ with the constant C, the communist dummy variable P, and the hosting dummy variable H_t. Therefore, the actual equation I used takes the following form, in which $\alpha_1 = \gamma$ and $\alpha_2 = \theta$.

$$\ln M_t = C + \alpha_1 \ln N_t + \alpha_2 \ln Y_t + \alpha_3 P + \alpha_4 H_t + e \qquad (4)$$

For the diminishing marginal return of population

TABLE 1
Definitions of Key Variables and Hypothesized Signs

Dependent Variable:	Definitions:	Hypothesized Sign
Medal Counts (M_t)	The number of medals won by a country in a particular Olympics	N/A
Independent Variables:		
Population(N_t)	The population size of a country at a particular Olympic year.	Model One: + Model Two: + and ? 1
GDP per capita (Y_t / N_t)	The per capita GDP (measured in P PP current international dollars) of a country at a particular Olympic year.	Model One: + Model Two: + and ? 1
Socialist Background (P)	1 if the country is or used to be a socialist country 0 otherwise	+
Hosting Country (H_t)	1 if the country is the hosting country of the year 0 otherwise	+

size and economic resources to be present, α_1 and α_2 need to be both positive and less than one. Moreover, similar to the linear function form, we should also expect α_3 and α_4 to be positive and statistically significant

Data used for this research are primarily from two sources. Data of Olympic medal counts and information of hosting countries are obtained by direct correspondence with International Olympic Committee (IOE). The data of population and per capita GDP (measured in PPP current international dollars) are extracted from World Development Report (World Bank, 2004). Table 1 gives the definition of each variable used in both models.

In this research, I used data from the last four Olympics (1988, 1992, 1996, and 2000). Athens Olympics are not included because the data of economic resources and population are not available until the end of the year. The reason why I do not include Olympics before 1988 is largely because Olympic performances in many of those games were affected by non-socioeconomic factors. For example, due to the Cold War, the United States did not attend the Moscow Olympics in 1980. Together with many other socialist countries, the Soviet Union boycotted

TABLE 2
Descriptive Statistics of Key Variables

Years	Variable:	Observations:	Mean:	Std. Dev.:	Min:	Max:
1988	Medal Numbers	52	14.33	26.10	1	132
	GDP per Capita	40	9320.75	6468.75	830	20520
	Population	46	57389.76	163763.40	105	1101630
1992	Medal Numbers	64	12.73	22.42	1	112
	GDP per Capita	53	10183.58	7413.01	440	24700
	Population Size	57	55741.74	157464.50	262	1164970
1996	Medal Numbers	73	10.70	16.39	1	94
	GDP per Capita	60	11295.00	8737.98	650	29770
	Population	63	68075.37	193656.90	284	1217550
2000	Medal Numbers	79	11.70	18.25	1	96
	GDP per Capita	73	12900.00	9905.70	710	35130
	Population	77	62985.83	184837.20	267	1262645

TABLE 3
Regression Results of Model 1 (Equation 1)

Years	Population	GDP per capita	Socialist Background	Host	Adjusted R^2
1988	0.0000298	0.0011436	14.17467	29.39847	0.3034
	(2.14)***	(3.21)***	(1.73)**	(2.1)***	
1992	0.0000576	0.0012806	8.459771	8.50263	0.3598
	(4.18)***	(4.11)***	(0.53)	(1.37)	
1996	0.0000271	0.0007394	8.313714	67.16248	0.5131
	(3.25)***	(3.54)***	(2.14)***	(5.12)***	
2000	0.0000386	0.0007787	10.53783	40.98336	0.3108
	(3.94)***	(3.81)***	(2.42)**	(2.58)**	

*** indicates si gnificance at 0.01 level
** indicates significance at 0.05 level

TABLE 4
Regression Results of Model 2 (Equation 4)

Years	ln Population	ln GDP	Socialist Background	Host	Adjusted R^2
1988	-0.4474455	0.8214528	2.11128	1.976354	0.501
	(-2.44)***	(4.96)***	(2.24)***	(3.88)***	
1992	-0.1952765	0.6266003	1.311737	1.311737	0.3949
	(-1.30)	(4.22)***	(3.48)***	(1.00)	
1996	-0.2673551	0.6951747	0.8200749	1.056373	0.4671
	(-1.94)	(5.39)***	(2.76)***	(1.02)**	
2000	-0.1499807	0.6355092	1.20652	1.947337	0.5556
	(-1.43)	(6.09)***	(5.27)***	(2.29)***	

*** indicates significance at 0.01 level
** indicates significance at 0.05 level

the Los Angeles Olympics in 1984. Nations that won at least one medal on a selected Olympic Game are selected as sample countries for that year. By omitting countries with zero medals, which are the majority of participating countries, the impacts of population size and economic resources should be more readily measurable. Descriptive statistics of data of each Olympics are provided in Table 2.

IV. RESULTS

The OLS regression results of the simple linear function (equation 1) and the Cobb-Douglas production function (equation 4) are presented in Table 3 and Table 4 respectively. Table 3 shows that population size and GDP per capita are consistently significant over time, though the magnitudes of them differ from year to year.

The socialist dummy variable and the hosting dummy variable are also statistically significant for the majority of Olympics, except the 1992 Los Angeles Game. Moreover, the magnitudes of those two variables also fluctuate drastically over time. For example, the hosting country of 2000 gave Australia approximatelt 41 additional medals, while the same position in 1992 only increased the medal counts of the United States by eight medals. One possible explanation for the coefficient of hosting dummy variable to be relatively small and less significant in 1992 is because it was not the first time for the United States to host the Olympics, and Americans were probably less excited than the citizens of the other three countries that were hosting the Olympics for the first time.

Model 2 tests the diminishing marginal return of both population size and economic resource. Following the standard Cobb-Douglas production function, which uses aggregate capita, I decide to use aggregate GDP as a measure of economic resources. Comparing the results to my research hypothesis, $\ln Y_t$, the dummy variable of socialist background, and the dummy variable of hosting countries are statistically significant have the expected signs. In addition, the coefficients of $\ln Y_t$ are consistently positive and less than one. However, we also see from Table 4 that the coefficients of $\ln N_t$ are negative, which indicates a negative correlation between a country's population size and its Olympic performance. Moreover, most of the coefficients of $\ln N_t$ are not statistically significant. These results contradict my research hypothesis as well as the results yielded from model 1. More importantly, it does not make economic sense. Given the strong correlation between $\ln N_t$ and $\ln Y_t$ (notice that GDP is just the product of per capita GDP and population size), it is reasonable to guess that the regression results of equation 4 might be distorted by multicollinearity. To test for the existence of a co-linearity problem, I conducted the correlation test between $\ln N_t$ and $\ln Y_t$. The results are provided in Table 5. It is obvious that there exists a considerably high level of co-linearity between $\ln N_t$ and $\ln Y_t$.

TABLE 5
Correlation Test Between $\ln N_t$ and $\ln Y_t$

	1988	1992	1996	2000
Correlation	0.7927	0.8039	0.8106	0.8277

One of the most commonly used strategies to correct a multicollinearity problem is to use an alternative function specification. In this case, a different function form, without using both $\ln N_t$ and $\ln Y_t$, will be desirable to estimate the diminishing marginal return of both population size and per capita GDP. In order to achieve this goal, I used a more restricted version of the Cobb-Douglas production function, which is shown by the following equation.

$$M_t = A_t (N_t)^{1-\theta}(Y_t)^{\theta} \qquad (5)$$

Equation 5 assumes the constant returns of scale, which implies that doubling both population size and GDP simultaneously will double a nation's medal counts. Although there are no solid theoretical justifications for the assumption of constant return of scale to be true, in this case it might be best to econometrically eliminate the multicolinearity problem. Taking the natural log of both sides of equation 5 and substitute $\ln A_t$ with P and Ht yield the following equation, in which $\alpha_2 = \theta$.

$$\ln M_t = C + (1-\alpha_2) \ln N_t + \alpha_2 \ln Y_t + \alpha_3 P + \alpha_4 H_t + e \qquad (6)$$

A simple mathematical transformation of equation 6 will give us a more appropriate equation:

$$\ln (M_t / N_t) = C + \alpha_2 \ln (Y_t / N_t) + \alpha_3 P + \alpha_4 H_t + e \qquad (7)$$

Notice that equation 7 achieves the goal by eliminating the co-existence of $\ln N_t$ and $\ln Y_t$. The coefficient of $\ln(Y_t / N_t)$ is the same as the coefficient of $\ln Y_t$ in equation 6, and the coefficient of $\ln N_t$ can be obtained indirectly by subtracting α_2 from 1. Moreover, the t-statistics of the coefficient of $\ln N_t$ can be calculated by dividing $(1-\alpha_2)$ by the standard error of α_2. The regression results of equation 7 are shown on Table 6. As we can see, the adjusted R-square is significant reduced from the original regression results of equation 4. This should be expected, because equation 7 is a more restricted function form than equation 4 by having an additional assumption of constant return of scale. Moreover, this restriction also makes the dummy variable of hosting advantage less statistically significant in predicting a country's Olympic performance. However, equation 7 makes more economic sense by having coefficients of $\ln N_t$ that are positive and less than one.

V. CONCLUSION

Consistent with previous studies on national Olympic performance, this paper finds that socioeconomic variables, including population size, economic resources, hosting advantage, and political structure have a significant impact on a country's Olympic performance. In general, population size and economic resources are positively correlated with medal counts. The larger the population size, the more likely a country is going to do better in the Olympics; the richer a country is, the more Olympic medals it will likely win. Being a hosting nation and having a communist background both have a favorable influence on a country's Olympic performance. Due to exogenous factors, e.g. international political atmosphere, the magnitude of those two variables differs significantly from one game to the other. Generally, my results are consistent with that of the studies carried out by Johnson and Ali (2000) and Bernard and Busse (2000). All the influential factors identified by those two studies are verified to be significant. However, due to the different data structure used (both of those two studies use panel data sets and measure population and GDP in terms of shares), it is difficult to compare the magnitude of each variable with their results.

Although this paper provides some insights on the correlation between a country's economic development and its Olympic performance, a major shortcoming of using cross-sectional data is the regression results are not quite useful in predicting countries' future medal counts, since coefficient of socioeconomic variables differs from one year to another. Hence, a more appropriate method to predict Olympic medal numbers would be regression using panel data. With knowledge of this obvious shortcoming, the reason I still decide to use cross-sectional data is that the prediction of future Olympic performance based on socioeconomic variables is not necessarily not as meaningful (as most people would think). Good Olympic performance is generally a byproduct of large population size and abundant economic resources. The logic that a country should increase its population size and its per capita GDP only because it wants more Olympic medals is impractical. Moreover, no country would ever want to change its political structure from democracy to communism simply for the sake of better Olympic performance.

As for future research, I would suggest including all countries with zero medal counts and using a probit model to estimate the impact of various socioeconomic variables on national Olympic performance across years. In this case, although the Cobb-Douglas production function is still valid in modeling Olympic performance, a conceptual problem is that the total number of medals available is exogenous (the number of medals in each Olympics is decide by the IOC). This problem might be solved by reconstructing the function and looking at the dependent variable and the explanatory variables in terms of shares. Moreover, future research should also take into consideration the countries that discourage women from participating in international sports events (examples will be some Middle-East countries). Therefore, variables like Political Freedom Index might be appropriate to be included to capture some cultural factors.

REFERENCES

Andrew B. Bernard and Meghan R. Busse. "Who Wins the Olympic Games: Economic Resource and Medal Totals". NBER working paper, August 2000.

Daniel K.N. Johnson and Ayfer 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.

Elmer Sterken and Cerard Kuper. "Participation and Performance at the Olympic Summer Games." Department of Economics, University of Groningen. 2000. <http://www.eco.rug.nl/medewerk/sterken/download/economysports.pdf >

Robert Hoffmann, Lee Chew Ging, and Bala Ramasamy. "Public Policy and Olympic Success". *Applied Economics Letter*, 545-548. 2002.

--------. "Olympic Success and ASEAN Countries: Economic Analysis and Policy

Implications". Nottingham University Business School, Malaysia Campus. 2004.

Shughart, William F., and Robert D. Tollison. "Going for the Gold: Property Rights and Athletic Effort in Transitional Economies" KYKLOS; v46 n2 1993, pp. 263-72. 1993.

Tcha, Moonjoong; Pershin, Vitaly. "Reconsidering Performance at the Summer Olympics and Revealed Comparative Advantage." *Journal of Sports Economics*, v. 4, pp. 216-39. August 2003.

Wallechinsky, David. *The Complete Book of the Olympics*, 1992 Edition, Little, Brown and Company, Boston. 1992.

Wing Suen, "Olympic Games and Economic Behavior" Working Paper. Department of Economics, University of Hong Kong. http://www.econ.hku.hk/~hrneswc/games.pdf >

World Bank, World Development Indicators 2004 (CD ROM). Published by World Bank. 2004.