Survival of the Fittest? An Econometric Analysis in to the Effects of Military Spending on Olympic Success from 1996-2012.

Mark Frahill

The Olympics are the world's greatest sporting celebrations, seen as a celebration of sporting values and a unified spirit. However, what determines the winners? In this essay Mark Frahill takes a novel perspective and assesses the relationship between a country's militarisation and its medal haul, discussing the possible mechanisms of this relationship and carrying out a comprehensive econometric analysis of the relationship. He finds that while there is evidence of a relationship it is not conclusive, which shows awareness for the level of estimation and uncertainty in econometric analysis, something often overlooked.

Introduction

What does it take to produce an Olympic champion? A nation naturally needs significant economic resources to boost health outcomes and be capable of investing in long-term training and infrastructure. The determinants of success in the Olympic Games would be expected to be the wealth of a nation and its population size. However, the interaction between Olympic success and the militarization of nations is important in understanding issues in society and global context today.

The regression analysis I will use is dependent on the assumption that changes in military spending cause changes in the obtainment of Olympic medals. I will hypothesise that military spending will cause higher receipt of medals because those countries with higher military spending would see the Olympic Games as an opportunity to gain soft power, influence and an opportunity to raise national pride. Militarised countries would have an interest in maximising this type of power and would inevitably see the Olympic Games as a way of obtaining this power. I propose that it would be difficult for the receipt of Olympic medals to increase military spending.

As an example, consider the fact that India, with a population of over 1.2 billion people has just 28 medals, while the USA (approximately 350 million people) has 2520 medals. We can see rich countries such as Monaco and Singapore that have gathered few medals and the massively populous India has less than even the small island of Ireland, while the United States overwhelmingly dominates. I will present this snapshot of the 1996 Summer Olympics in Atlanta in which the USA won 101 medals, Ireland won 4, while Cuba won 25 and India picked up a solitary medal. Here, again, Ireland won more than India and the US topped the board, but Cuba (which spends more than average on military) won 25 medals. This paper asks the question of whether countries can expect to win more medals by increasing military spending.

Literature Review

The Summer Olympic Games is, globally, the largest sporting event and a source of both national pride and lifetime success for its athletes. Bernard and Busse (2004) establish the link between economic strength and Olympic success. They used pooled data from 1960-1996 and found that population and income per capita are needed to generate high medal totals. Johnson and Ali (2004) examined the 1952-2000 Summer and Winter Olympic Games. They concluded that socioeconomic factors explain Olympic participation rates particularly well. Income is a key driver of success with wealthier more populous countries being more capable of sending athletics to compete and therefore succeed.

Of course, no econometric model can capture every political and economic factor involved in medal winning. Research from Ho man, Ging and Rama (2002) found that many inherent characteristics and cultural factors have an impact on the receipt of medals, although to a limit. These factors are very difficult to quantify in practice and this paper has strictly included only quantitative factors in its analysis.

Research about Olympic success has largely focused on the characteristic determinants of each economy such as GDP per capita and population, without considering the distribution of the spending. The research presented in this paper provides a fresh perspective on the Olympics, analysing data from recent games during 1996-2012, and adds to the literature by considering new variables of interest such as military and healthcare spending in terms of current US dollars. This paper seeks to prove that if medal distributions act in a way that is consistent with countries that have a high proportion of military and healthcare spending.

Data and Expectations

The data for this paper, broadly speaking, consistent of two dimensions; Olympic Medal Counts and Economic Indicators. Five Olympic Games (1996, 2000, 2004, 2008, and 2012) and 190 countries were analysed.

The 2016 Olympic Games was excluded as 2016 economic data was not yet available for countries at the time of this study. Olympic Games prior to 1996 were not analysed as data was not available comprehensively or readily available for countries. In addition, due to the fall of Communism and breakup of communist countries such as the Soviet Union and Yugoslavia in the early 1990s, older currencies would be difficult to compare, and many newly created countries would be problematic to analyse.

Additionally, Olympic participants that are territories of countries were not included, for instance places such as Guam, Cayman Islands or Macao. Neither were countries with unreliable figures such as North Korea or Somalia. Countries that had no military or health expenditure were dropped from the dataset as they were not of interest to the study.[1]

The dependent variable of this examination was Olympic performance by total medals won per Olympics per participant country. This was taken from an aggregate number of bronze, silver and gold medals won by each country that year and treating each medal as worth the same. The data for each country was readily available online.

The independent variables were a range of economic indicators about each country. The data for military expenditure in current US dollars, healthcare expenditure in current US dollars, GDP in current US dollars, GDP per capita in current US dollars, population, and country area in total kilometres squared, were all readily available online and were taken from the World Bank and CIA World Factbook.

Empirical Approach

Using a simple regression model:

$$\text{medals} = \beta_0 + \beta_1 \text{armyVOL}$$

The variable "medals" represents the number of medals a country obtains at a given Olympic games, while the dependent variable "ArmyVOL" represents the volume of military spending in the country. This equation should estimate the gross effect of military spending on medal hauls through the coefficient β_1.

Military spending, however, is not the only factor affecting Olympic medal hauls, so it is highly probable that this model would suffer from omitted variable bias. Therefore, I will control for the other main likely determinants of Olympic medals won using a multiple regression model:

$$medals = \beta_0 + \beta_1 armyVOL + \beta_2 gdp + \beta_3 gdpcap + \beta_4 size + \beta_5 pop + \beta_6 healthVOL$$

This model attempts to describe the net effects of each independent variable on Olympic medal hauls, controlling for the other variables. This model is much less likely to suffer from omitted variable bias, as most of the key factors influencing medals are controlled for in this way.

Panel data estimation methods were applied to this study's data. Its advantage over time series is that it controls for unobservable and time-invariant factors. It also overcomes the omitted variable bias that is likely to exist in the Olympics, for instance cultural and geographic influences, which Hoffman, Ging and Rama (2002) demonstrate have an influence.

Both fixed effects and random effects were considered. A Hausman Test was run to decide between fixed effects and random effects estimation, concluding that due to a p-value of 0.0035, fixed effects estimation was appropriate. Using FE estimation allows for correlation between the unobserved effects and the independent variables, as opposed to RE which requires these to be uncorrelated.

To test the quadratic relationship between what might cause an Olympic medal to be won and military spending the following initial model is specified:

$$medal_{it} = \beta_0 + \beta_1 armyVOL_{it} + \beta_2 armyVOL_{it}^2 + \alpha_i + u_{it}$$

Then a more comprehensive model is examined to try and incorporate a full examination using what has been learned from the background literature, such that:

$$medals_{it} = \beta_0 + \beta_1 armyVOL_{it} + \beta_2 armyVOL_{it}^2 + \beta_3 gdp_{it} + \beta_4 gdp_{it}^2 + \beta_5 gdp_{it} + \beta_6 gdp_{it}^2 + \beta_7 size_{it} + \beta_8 size_{it}^2 + \beta_9 pop_{it} + \beta_{10} pop_{it}^2 + \beta_{11} healthVOL_{it} + \beta_{12} healthVOL_{it}^2 + \alpha_i + u_{it}$$

Table 1 provides a breakdown of the summary statistics for the data and the number of observations.

Table 1: Summary Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
medals	716	6.072626	14.71407	0	111
gdp	716	23.98341	2.328683	16.32794	30.41327
gdp^2	716	580.619	112.4467	266.6016	924.9668
gdpcap	716	8.06341	1.651255	4.782983	11.63383
$gdpcap^2$	716	67.74142	26.98323	22.87692	135.3459
size	716	12.0322	2.0998	5.755742	16.65449
$size^2$	716	149.1769	48.32073	33.12857	277.3719
pop	716	16.12071	1.653799	11.24396	21.02389
pop^2	716	262.6084	53.48263	126.4266	442.0038
armyVOL	716	19.84806	2.472037	13.14339	26.69813
$armyVOL^2$	716	400.0477	98.57866	172.7487	712.7902
healthVOL	716	21.10588	2.444592	12.78831	27.81892
$healthVOL^2$	716	451.4257	104.4983	163.5408	773.8923

Descriptive Statistics

The variables are as follows:

Table 2: Variable Descriptions

medals	Number of Olympic Medals won by a country at an Olympic Games
gdp	GDP in current US Dollars (natural log)
gdp^2	gdp squared
gdpcap	GDP per capita in current US Dollars (natural log)
$gdpcap^2$	gdpcap squared
size	Total area of a country in km^2 (natural log)
$size^2$	size squared
pop	Total population of a country (natural log)
pop^2	pop squared
armyVOL	The volume of military expenditure in current US dollars (natural log)
$armyVOL^2$	armyVOL squared
healthVOL	The volume of health expenditure in current US dollars (natural log)
$healthVOL^2$	healthVOL squared

Table 3: Variable Correlations

	medals	gdp	gdp^2	gdpcap	$gdpcap^2$	size	$size^2$	pop	pop^2	armyVOL	$armyVOL^2$	healthVOL	$healthVOL^2$
medals	1.0000												
gdp	0.3941	1.0000											
gdp^2	0.4117	0.9977	1.0000										
gdpcap	0.3191	0.4460	0.4487	1.0000									
$gdpcap^2$	0.3268	0.4476	0.4508	0.9950	1.0000								
size	0.3585	0.2445	0.2566	-0.1450	-0.1441	1.0000							
$size^2$	0.4008	0.2716	0.2840	-0.1154	-0.1151	0.9906	1.0000						
pop	0.4435	0.4226	0.4359	-0.1203	-0.1157	0.7233	0.7230	1.0000					
pop^2	0.4642	0.4356	0.4498	-0.1121	-0.1083	0.7120	0.7176	0.9968	1.0000				
armyVOL	0.4100	0.9271	0.9265	0.4510	0.4508	0.2757	0.3014	0.3974	0.4066	1.0000			
$armyVOL^2$	0.4325	0.9320	0.9360	0.4526	0.4530	0.2876	0.3138	0.4167	0.4273	0.9967	1.0000		
healthVOL	0.4236	0.9863	0.9854	0.4764	0.4800	0.2371	0.2610	0.4018	0.4119	0.9155	0.9209	1.0000	
$healthVOL^2$	0.4436	0.9821	0.9865	0.4800	0.4844	0.2502	0.2742	0.4147	0.4257	0.9138	0.9245	0.9968	1.0000

The expected results are as follows:

armyVOL: Is expected to have a positive effect, as the Olympics are a display of national superiority which presumably is more important to countries that spend more on their military power.

gdp: The background literature establishes the significance, and it makes economic sense that countries with more resources can invest in sports programs and infrastructure and increase medals won.

gdpcap: Is expected to have a positive effect as richer countries can afford to invest more in individuals training and athletics career opportunities.

pop: Is expected to have a positive effect as countries have a greater talent pool from which to draw competitors.

size: Is expected to have a positive effect, as there is more area for infrastructure and diversity of sport.

healthVOL: Is expected to have a positive effect as citizens would presumably be healthier and in better physical form to compete in the Olympics

Empirical Results

Results of the simple regression were as follows:

$$medals = 42.36894 + 2.44062armyVOL$$

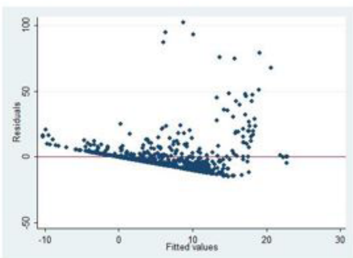
Table 4: Simple Regression Results

Simple Regression Results	Medals
armyVOL	2.4406*** (.2031)
Constant	-42.3689*** (4.0636)
Observations	716
Number of countries	162
R-squared	0.1681

Standard errors in parentheses*** p<0.01, **p<0.05, *p<0.1

The simple regression indicates that the relationship between a country's military expenditure and their Olympic success is statistically significant. An R-squared value of 0.1681 was obtained, implying that armyVOL can explain 16.81% of the variation in the model within a simple linear regression. According to the t-test, armyVOL is significant at the 1% level. This confirms this paper's working assumption that military expenditure influences Olympic success and that the relationship is positive. However, as we can see from a plot of the residuals against the fitted values in Figure 1 there is a clustered pattern which should not be true for linear data, implying there is violations of the least squares assumptions as the residuals are not homoskedastic. This opens the model up to further examination such as quadratic forms and heteroskedasticity analysis through robust tests.

Figure 1: Fitted values against residuals for Equation 5



Results of the multiple regression were as follows:

$$medals = 62.45367 + 1.0719armyVOL - 7.48gdp + 2.709gdpcap + 0.6439size + 6.612healthVOL$$

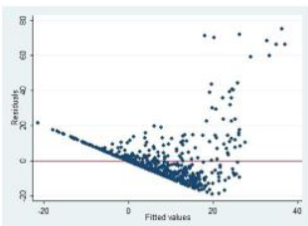
Table 5: Multiple Regression Results

Multiple Regression Results	medals
armyVOL	1.0729** (.4815)
Gdp	-7.4808*** (1.256)
gdpcap	2.7089*** (0.3305)
size	.6439** (.3076)
healthVOL	6.6129*** (1.115)
Constant	-62.454*** (7.06)
Observations	716
Number of countries	162
R-squared	0.3808

Standard errors in parentheses*** p<0.01, **p<0.05, *p<0.1

Exploring more independent variables in the multiple regression leads to new results as can be seen in Table 5, with an R-squared explaining 38.04% of the variation in medal-winning. Four variables (gdp, gdpcap, pop, healthVOL) were significant at the 1% level and two at the 5% level (size, armyVOL). The relationship between armyVOL and medals is still positive. All the other relationships are positive as expected, except for gdp which has a large negative coefficient which is strange as the background literature demonstrated a positive effect. Looking at a plot of the fitted values against the residuals we can see a clustered pattern, concluding the linear assumptions needed for OLS are not present which may explain this negative coefficient.

Figure 2: Fitted values against residuals for Equation 6



The results from the FE estimation are shown in Table 6. An R-squared of 0.2469 was obtained. 24.69% of the variation in medal winning is explained by the volume of military spending by a country. The armyVOL and $armyVOL^2$ are seen to be statistically significant at the 10% and 1% levels. The t-values are all significant. The Prob > F is 0.0057 which implies the model is alright and all the coefficients are statistically significant from zero. The coefficients of the independent variables are listed as -5.858 and 0.1604 such that:

medals=5.858+0.5208armyVOL

The (maximum) turning point for this value should be armyVOL=5.858/0.3208= 18.2606, which is equivalent to $86,318,076.43.2 This seems to confirm the working assumption that medal-winning can be function of only countries with high military spending, though naturally much more investigation is required. However, the rho value is 0.92944 implying that 93.94% of the variance is due to differences across panels. As we can see from Table 6, after running the robust fixed estimation each variable is no longer statistically significant.

Table 6: Fixed Effects and Robust Fixed Effects

	Fixed Estimation	Robust Fixed Estimation
armyVOL	-5.858* (2.336)	-5.858 (7.835)
armyVOL2	0.1604*** (0.593)	0.1604 (0.2092)
Constant	58.190* (22.906)	58.190 (71.829)
Observations	716	716
Number of countries	162	162
R-squared	0.2469	0.2469

Standard errors in parentheses*** p<0.01, **p<0.05, *p<0.1

Extending our panel data analysis to our full range of chosen variables we get a full picture of the effects. The R-squared is 0.2842, implying the explanatory variables explain 28.42% of the variation. The Fixed Effects estimation showed only four variables to be statistically significant armyVOL, $armyVOL^2$, healthVOL and $healthVOL^2$, while using robust measures resulted in no statistically significant variables. This is different from what we expected. The Prob > F is 0.0135 which implies the model is not alright and all the coefficients are not statistically significant from zero. The coefficients of the independent variables armyVOL and armyVOL2 are listed as -7.85391 and 0.2255 such that:

$$medals=7.85391+0.451armyVOL$$

The (maximum) turning point for this value should be armyVOL=7.85391/0.451= 17.41, which is equivalent to $36,558,877. This seems to confirm our working assumption that medal winning can be function of only countries with high military spending, though naturally much more investigation is required. However, the rho value is 0.9147 implying that 91.47% of the variance is due to differences across panels.

Table 7: Fixed vs Robust Estimation

	Fixed Effects Estimation	Robust Fixed Effects
armyVOL	-7.85391* (4.093)	-7.85391 (5.751)
armyVOL²	.2255687** (0.1081)	.2255687 (0.1579)
gdp	17.27902 (11.254)	17.27902 (23.228)
gdp²	-.3873933 (0.4210)	-.3873933 (0.5120)
gdpcap	2.276458 (2.078)	2.276458 (3.695)
gdpcap²	-.1794398 (0.1336)	-.1794398 (0.2575)
pop	-13.56585 (15.589)	-13.56585 (17.1883)
pop²	.510821 (0.5)	.510821 (0.6158)
healthVOL	-12.94962* (7.533)	-12.94962 (15.019)
healthVOL²	.3218201* (0.1820)	.3218201 (0.376)
Constant	88.61725 (123.27)	88.61725 (104.1906)
Observations	716	716
Number of countries	162	162
R-squared	0.2842	0.2842

Standard errors in parenthesis*** p<0.01, **p<0.05, *p<0.1

The variable size indicating a countries area in km^2 in naturally unchanged from year to year and so was run in a separate random effects regression to avoid collinearity within the fixed effects estimation results. The results were as expected. An R-squared of 0.2398 was obtained, therefore 23.98% of the variation in medal winning is explained by the landmass of a country. The size and $size^2$ are seen to be statistically significant at the 0.01% level, with a negative coefficient for the linear size -14.043 and a positive for the quadratic $size^2$ 0.7191 (see Table 10).

An important point to consider is that after running the Woolridge test for autocorrelation in the panel data, it was concluded that there is first order autocorrelation within the panel data models.

Table 10:

Random Effects Estimation Results	medals
Size	-14.043***
	(3.41)
size²	.7191***
	(.1481)
Constant	67.5161***
	(19.382
Observations	716
Number of countries	162
R-squared	0.2398

Standard errors in parenthesis
*** p<0.01, **p<0.05, *p<0.1

Possible Extensions

As the model presented does not account for all variables that create Olympic success there is much scope for extension. The dataset I used, though large at 716 observations, was limited to the games between 1996-2012. Data for 2016 was not yet available however should be to certain organisations, and to the future public. Perhaps extensions could include an entire analysis of every Olympic Games to get a complete picture.

There are also several outlier countries that may be influencing the data significantly such as United States, China, Great Britain, Russia, France, Germany and Australia and it might be useful to examine the data excluding these countries. Doing this however removes a significant proportion of medals won each year.

Conclusion

There may be a link between how much spending in terms of total volume in current US dollars on healthcare or military spending and Olympic success. However, depending on which statistical testing methods are used it is difficult to say for certain what is significant. A simple regression showed a positive link between military spending and medals won, and this relationship continued in the multiple regression. Using a panel data analysis with fixed effects regression also showed evidence of this relationship, though under robust conditions the chosen metrics for military spending were not statistically significant. This perhaps highlights the need for rigour in econometric analysis and conveys the need for caution when interpretation results.

Though the effect on medals was significant for both simple and multiple regression, it lost its significance in the more advanced panel data model. This is a useful example of the need for correct model specification and an appreciation for the level of estimation involved in econometric results. Unfortunately, the specific effect of military spending on Olympic medal hauls remains somewhat unknown.

Reference List:

1. Bernard, A. and Busse, M. (2004). Who Wins the Olympic Games: Economic Resources and Medal Totals. Review of Economics and Statistics, 86(1), pp.413-417.
2. Hoffmann, R., Ging, L. and Ramasamy, B. (2002). Public Policy and Olympic Success. Applied Economics Letters, 9(8), pp.545-548.
3. Johnson, D. and Ali, A. (2004). A Tale of Two Seasons: Participation and Medal Counts at the Summer and Winter Olympic Games*. Social Science Quarterly, 85(4), pp.974-993.