Exploring Economic Indicators as Determinants of Medal Success in the Summer and Winter Olympics

Nate Ukoha – Tufts University

Economics of Sports – EC0149

May 06, 2024

Abstract: This paper investigates determinants of medal success at the Olympic Games. Because of the wide variety and large magnitude of participants involved at the Olympics, this study explores more holistic ways of predicting medal outcomes. Furthermore, it seeks to explain some countries' historical lack of success at both the Summer and Winter Games and uncover what intrinsic traits of a country might consistently shape how they will perform despite the Olympics being comprised of primarily individual competing athletes. The model evaluates economic determinants, host nation status, inequality measures, and government stability across a span of 72 years. Ultimately, there is clear evidence of host nation effect, especially during the summer Season. Meanwhile there are noticeably smaller and fluctuating levels of evidence justifying the influence of population and income across time.

1. Introduction:

` A robust methodology for predicting medal success at the Olympic Games would ideally involve an evaluation of talent and personnel for a single competition and repeating such analysis across all competitions for each nation. However, since the first modern Olympics dating back to the end of the 20th century, the Games have grown massively in scale of participation both at an individual and national level, as well as in terms of the breadth of events competed in. Therefore, the feasibility of this talent evaluation strategy faces many barriers, mainly pertaining to limited access to information as well as a high athlete turnover rate across years because of increased diversity in competition.

Now as attention turns to the upcoming 2024 Paris Games, which will anticipate over 200 National Olympic Committees and over 10,000 athletes, the value of overall medal outcomes, either perceived or actual, cannot be overstated(Pareja). Powerhouse countries such as the United States, Germany, and or some forms of the USSR are projected favorites to leave with the most medals totaling 219 medals. Historically, these three countries together have won 32% of all medals at the Olympic Games while 38% of participating NOCs involved in the 2020 Summer Games have never medaled(Reardon). Though it is ambiguous whether an outcome as trivial as standalone national medal count really matters relative to current socioeconomic issues that are prevalent globally, history has at least shown us that there are a wide variety of both underlying and overt social and financial implications of medal success, especially when it is associated within the context of a specific time period. For many athletes, the Olympics often represent either the culmination of their body of work as competitors, or potentially a launching point from which they can propel and advance the rest of their careers. Katie Ledecky's gold-medal winning time of ~8:14 minutes at only 15 years old in the 2012 London Games marked only the beginning of a successful, decorated and long-tenured career as a competitive swimmer (Katie). However, Wayde van Niekerk, already a renowned domestic track sprinter, set the record for the 400-meter dash during the 2016 Rio Games, marking his ultimate accomplishment as a professional.

Contrarily, the implications of Olympic medal counts can also be observed from a macro-level and often transcend beyond simply influencing an individual's career trajectory. Specifically, because of the international tensions leftover from World War II, and the formation of new nation states, the significance of the Olympics has taken on a somewhat redefined meaning. The Olympics now serve as a medium for intense displays of (hyper)nationalism, in some ways contradicting and detracting from its original purpose as a means of promoting international solidarity and cooperation as well as intellectual and physical human development between countries, the overall premise behind Olympism[1]. Furthermore, Ditmore et al. discusses the media's role as one of the driving forces behind these displays of nationalism, mentioning

[1] "Olympism is a philosophy of life, exalting and combining in a balanced whole the qualities of body, will and mind. Blending sport with culture and education, Olympism seeks to create a way of life based on the joy of effort, the educational value of good example, social responsibility and respect for universal fundamental ethical principles.

The goal of Olympism is to place sport at the service of the harmonious development of humankind, with a view to promoting a peaceful society concerned with the preservation of human dignity."

disproportionate coverage of American athletes on U.S. network providers such as NBC during the 2020 Olympics relative to America's actual levels of success, as well as padded medal count table data by CCTV stations in China (Dittmore). Thus, the prevalence of nation-grouped medal tables has led to the perception that a single country can win or lose the Olympics, thus perpetuating the ever-divisive "Us vs. Them" narrative.

In this sense, the medal count is seen as not just an individual award but also as a nation's shared accomplishments. Audiences from a specific nation might enjoy the medal results specifically because they hold the belief that unique characteristics about their own country, to which they also share and contribute, play a part in the success of an athlete.

Exploring medal outcomes is important because Olympic success is one of several factors that play a role in the amount of participant compensation, either in the form of a lump sum payment, or a sizeable wage increase. The size of their compensation often reveals how valuable a medal is not necessarily just for the individual, but also for the country they represent. Prior to the 2022 Winter Olympics, an article published by Forbes reported that Turkish and Hong-Kong athletes were eligible to receive up to $380,000 and $642,000 respectively if they won gold. Meanwhile, U.S. were paid athletes only $37,500 directly for gold. Considering that Hong-Kong (which operates independent of China) and Turkey at the time of this report had never won gold at the Winter Olympics, and the U.S. is a regularly top-performing country during this international competition, the level of compensation most likely reflects (in part) how rare a given medal is for a certain country. So, given that Hong Kong and Turkey are not the smallest, newest, or least-funded participating delegations, what factors may contribute to their historic lack of success?

This paper investigates to what extent socioeconomic qualities about a country determine success at the Olympic Games. Moreover, I am curious as to why bigger countries might not see the same level of success as other large countries since the expectation is that the pool of available athletic talent to choose from is also large.

In this paper, we analyze the effect of macro level indicators on the national Olympic Medal Count. I am curious if nation-specific attributes play any role in determining their overall success at the Games. Looking at both the Summer and Winter Olympics, I will attempt to explain total medal count outcomes, by leveraging a combination of economic, political, and geographic indicators as explanatory variables, while also creating distinctions between men's and women's medal count data. The ensuing sections will contain a literature review of similar and influential empirical approaches, an overview of the data including explanations of variable choices, an outline of the choice of estimation methodology, and a finally review of the findings from this analysis as well as conclusions.

2. Literature Review:

There have been extensive efforts to quantify the relationship between economic indicators and Olympic Success, with notable papers dating back to the 1970s. Existing literature tends to look at variables that may vary across time, country, or both. Frequently chosen factors with at least some level of significance include the advantages of host country status, per capita wealth, population size, and weather. With the existence of the Soviet Union, and their well-documented success as Olympic participants, non-traditional indicators that have also been chosen include corruption status and political categorization. Moreover, papers relevant to this topic vary in whether they identify distinctions between men's and women's medal count outcomes. There were three major studies that guided my expectations for the results of this paper as well as influenced my choice of methodology.

One notable paper by Noland et al. in 2017 does not just investigate historical economic indicators as determinants of success at the Olympics but does so within "the context of these events' growing pluralism, evolving from their aristocratic and largely European male roots to the more gender inclusive and geographically diverse showcase of athletic talent that we see today" (Noland). Essentially, as the number of countries and cultures involved in the Olympics have developed over time, the authors expect there to be a diminished impact of general economic indicators. Here, the authors use medal shares, rather than medal counts as a proxy for success and predicts medal share for individuals rather than nations as a percentage of all medals available in male of female specific events. Moreover, compared to similar studies, Noland's paper offers the most robust model on this topic by forming predictions using a sports-specific and gender specific underlying cross-sectional panel dataset. For each season, they employ a separate core OLS specification with fixed effects and a lagged dependent variable where both models look at population, education, income per capita, latitude, current and previous host nation status, Soviet Union membership, and notable boycotting years.

Consistent with most other studies, Noland finds a significant effect of host nation status. However, the net effect of increased diversity over time is insignificant in the Winter Olympics because the scope of participating teams involved has grown to a much lesser extent compared to the Summer Olympics. Thus, large, rich colder countries are still dominant in the Olympics, while a large scale of diversity has diminished the effect of these factors since the barrier to participation is much lower. In such cases, education has arisen as an important determinant of medal winnings rather than income, but overall, the net effect of these economic factors is found to be far smaller as time progresses. Finally, on an event-specific level, the authors find that larger, richer, and more educated countries succeed at a higher rate in sports whose infrastructure requires large capital expenditures, while poorer nations face less handicaps in more universally accessible sports such as “athletics and boxing” (Noland).

In another paper, Todd Potts is also interested in the role of corruption in predicting medal counts. In addition to including common economic indicators, he builds a Tobit regression model and uses a set of World Governance Indicators as a proxy for a government’s ability to control corruption between 1996 and 2016(Todd). This expectation that he has for corruption to be a significant determinant of medal count is based off Russia’s long-standing history of doping violations leading to the suspension of several athletes in recent decades as well as the suspension of the Russian NOC entirely in recent years, thus disrupting their track-record for winning a lot of medals. Potts expects such an incidence of disqualifications to take away the best athletes for a given nation, thus leading to a positive correlation between medal count outcomes and civil obedience. Here, the author uses the *Rule of Law* Word Governance Indicator provided by the World Bank, which “captures perceptions of the extent to which agents have confidence in and abide by the rules of society, and in particular the quality of contract enforcement, property rights, the police, and the courts, as well as the likelihood of crime and violence”(Home). Though Rule of Law does not pertain specifically to the Olympic rulebook, the assumption Potts makes here is that level of uncontrolled corruption in a country can be indicative of how often athlete’s cheat. For example, nations with lower *Rule of Law* scores have citizens who are not as able or willing to follow the rules of society, therefore they are less likely to abide by IOC rules, and more likely to be suspended. The findings show that high average rule of Rule of Law and Control of Corruption scores are more likely to result in significantly higher medals scores and less doping violations suspensions, with countries in the 67th percentile of scores having a 15% greater chance of medaling. However, unlike Noland’s study, this paper does not differentiate genders and events, thus not allowing us to observe the magnitude of impact on gender-or-event-specific events.

Another key paper explaining Olympic success is *(Non) determinants Olympic Success* by Johan Rewilak. This study uses a similar Tobit regression to Potts but incorporates male and female medal counts for each observation (John). Furthermore, unlike the two other papers, Rewilak also applies the *Mundlack Correction*[2] to his model by adding the group means of independent variables to his specification to best estimate random effects. Moreover, this model controls time-constant un-observables such as a country's landscape (John). The paper also measures the effect of gender equality (using secondary school enrollment info). By using this methodology and relaxing the "assumption that a single factor determines medal success", he finds that host nation status and size of country are the only significant predictors of medal success while the effect of political status which was significant pre-correction is no longer significant. Finally, the author also discovers that while there is a roughly equivalent significant effect of host nation status on both male and female sports, but the impact of population differs between genders.

Therefore, the analysis in this paper will attempt to contribute to existing literature by drawing from different aspects of analysis conducted in these three papers. Specifically, I will look to observe differences in male and female medal counts, looking at both predicted medal counts and percentage medal share, while also including governance indicators as well as common economic indicators and proxies for climate. Furthermore, this model will incorporate

[2] "Estimates random-effects regression models (xtreg, re) adding group-means of variables in indepvars which vary within groups. This technique was proposed by Mundlak (1978) as a way to relax the assumption in the random-effects estimator that the observed variables are uncorrelated with the unobserved variables".(Perales)

effects about current host status. Finally, I am looking to control the growing diversity seen in the Summer Games and to explore the role that income inequality plays in medal outcomes using the Gini Coefficient.

3. Data Source:

The information analyzed in this study is a cross-sectional panel data set with observations ranging from 1948 (beginning of the post-war period) to 2022 (the most recent Winter Olympics). The full data set represents the consolidation of several underlying data sets from a few different sources. The medal count information was gathered using both IOC data and Olympedia. From these two websites, for each Olympic Games, I left-merged separate tables about the names of countries participating in a certain year, the year itself, and the medal counts for all NOCs that had earned at least one medal. Men's, women's, and total medal counts all occupied separate columns in the preliminary data set. To note, the sum of men's and women's medal counts do not necessarily always equal total medal counts likely because of omission of mixed and open gender events. This merger was necessary to also identify countries who had participated, but earned no medals that year so that their medal count could be imputed as 0. From here, a consolidated panel data was formed by horizontally merging information from all years into one table. Then, more explanatory indicators were added as columns to the rest of the data set.

NOCs (National Olympic Committees) are the representative body of athletes that in most cases compete for a designated country, territory, or are an independent body of athletes who might not have access to either (National). Latitude is also measured in degrees and is a time-constant effect that is included for each country. Another fixed effect variable is simply the season to denote whether that observation took place during the Winter or Summer Games.

The next characteristic of the data set is real GDP per capita using 2022 purchasing price parity. Thus, the effects of inflation and price-level fluctuations are controlled. RGDP per capita info is provided for a given country each year. Population data is also included, for every country each year. In all three papers, these two indicators were consistently included in each regression specification, regardless of their significance level. Thus, they are included here due to expectations developed from previous findings.

These major economic and political indicators were provided by the World Bank and include robust information dating back to 1960. For the years prior to 1960, information was manually searched and inputted when possible and when relevant, using a combination of Trading Economics, Statista, as well as Macrotrends.net.

Another indicator that I borrowed from the World Bank was the Gini coefficient which measures the level of income inequality in the nation. This variable is also country fixed but is only available between 1990 and 2022. Therefore, there are roughly the number of observations in the data set which will possess this information. I use the Gini Coefficient because according to a working paper by Kufenko et al., a 1% increase in the Gini Coefficient leads to a 1.47% decrease in the probability of winning at least one medal. Thus, I am curious if the effects of income inequality are still existent when included with a plethora of other variables within a longer-term data set.

For this study I also use one of the World Bank’s, World Governance Indicators (WGI). Instead of focusing only on corruption, I build on Government Effectiveness. According to the World Bank, this metric “Reflects perceptions of the quality of public services, the quality of the civil service and the degree of its independence from political pressures, the quality of policy formulation and implementation, and the credibility of the government's commitment to such policies" (Home). I include this variable because unlike corruption, which is more focused on violation of societal rules, GE is a more holistic measurement of the quality of government. Federal and local governments often play a huge role in the allocation (or lack thereof) of funding for sports at youth, adolescent, and professional levels. Thus, GE here could be a proxy for how easily the public can access the resources required to play sports, and in general enjoy a higher standard of living. Because GE is estimated as a score between -2.5 and 2.5, this model will instead use percentile rank of a given country's GE estimate, relative to the scores of other countries. Medal share is also included, and it is calculated as the share of medals earned by that country in a specific Olympic year and season. This metric better represents how large an impact that NOC made in a specific Olympic Game. Since the number of events and participants vary each year, the amount made of medals available depends on the circumstances of that year.

3.1. Summary Statistics Table:

Variable	Obs	Mean	Std. Dev.	Min	Max
year	3767	1992.619	20.506	1948	2022
fmedals	3767	1.822	5.866	0	73
mmedals	3767	2.684	7.856	0	131
totalmedals	3419	4.965	13.563	0	195
medalshare	3085	0.009	0.025	0	0.309
host	3767	0.01	0.1	0	1
population	3306	39,942,908.00	141,900,000.00	10,243.00	1,417,000,000.00
Gini	3283	9.563	16.706	0	63.4
GiniFixed	852	36.848	8.347	23.7	63.4
latitude	3566	24.357	25.532	-41	65
govE	1798	54.003	29.283	0	100
id	3767	108.286	63.259	1	228
rgdppc	3328	$ 11,499.74	$ 21,298.05	$ -	$ 240,862.19
summer	3767	0.703	0.457	0	1

As seen from the given graph, there are in total over 3700 observations across a 72-year time span. Over the years, men's events tend to garner more medals than women's events (1.8 vs. 2.6), and an average country's overall medal share is roughly 1% (0.009). The average population of a competing country is roughly 40 million, and the average Gini score is ~37. Furthermore, the average country is in the 54th percentile of government effectiveness rank and has an average GDP per capita of $11,500 USD. Finally, there tends to be more participants in the Summer Olympics, which hints at lower barriers to entry compared to the Winter Olympics.

4. Methodology

To best explain the relationship between economic indicators and olympic success, I employ 6 core specifications, along with supplemental regression analyses that will be included in an Appendix section.

4.3 Core OLS Specifications

1. $MedalShare = \alpha + B_1 * Host_{season*,year,country} + B_2 * ln(RDGPPC)_{season*,year,country} +$

   $B_3 * ln(Population_{season,year,country}) + \epsilon$

2. $Fmedals = \alpha + B_1 * Host_{season*,year,country} + B_2 * ln(RDGPPC)_{season*,year,country} +$

   $B_3 * ln(Population_{season,year,country}) + \epsilon$

3. $Mmedals = \alpha + B_1 * Host_{season*,year,country} + B_2 * ln(RDGPPC)_{season*,year,country} +$

   $B_3 * ln(Population_{season,year,country}) + \epsilon$

4. $MedalShare = \alpha + B_1 * Host_{season*,year,country} + B_2 * ln(RDGPPC)_{season*,year,country} +$

   $B_3 * ln(Population_{season*,year,country}) + B_4 * Gini_{season*,year,country} +$
   $B_5 * GovernmentEffectiveness_{season,year,country} + \epsilon$

5. $Fmedals = \alpha + B_1 * Host_{season*,year,country} + B_2 * ln(RDGPPC)_{season*,year,country} +$

   $B_3 * ln(Population_{season*,year,country}) + B_4 * Gini_{season*,year,country} +$
   $B_5 * GovernmentEffectiveness_{season,year,country} + \epsilon$

6. $Mmedals = \alpha + B_1 * Host_{season*,year,country} + B_2 * ln(RDGPPC)_{season*,year,country} +$

   $B_3 * ln(Population_{season*,year,country}) + B_4 * Gini_{season*,year,country} +$
   $B_5 * GovernmentEffectiveness_{season,year,country} + \epsilon$

The purpose of the first three specifications is to isolate primary economic and social indicators. Since these are the most commonly significant explanatory variables across examined literature, they are included here to understand if their effect is still significant over a longer period of time. Moreover, these first three equations looking at host status, GDP per Capita(an income proxy), and population offer the widest range of available information. Due to the recency with which the World Bank started recording them (~1990), metrics such as Government effectiveness (GE) and the Gini Coefficients (GiniFixed), cause a larger drop in the number of observations. Thus, the first three and last three groups represent differences in the amount of explanatory *X* values included. For all specifications, I include season, year, and country as fixed effects. Though certain modifications are not listed out explicitly, other specifications will be generated at times, ignoring or including any combination of these fixed effect variables; however, for the sake of simplicity, they are shown here to represent the *option* of using them.

While overall medal count predictions are represented as a share of all medals earned, gender-specific medal count dependent variables are expressed as raw counts. Finally, to better interpret the impact of income per capita and population, the natural log of both is taken on the right-hand side to simplify our interpretation of their effect in terms of percentage increases or decreases. The next three equations perform similarly to the first three but include the role of government quality, expressed as a percentile rank, and Gini Coefficients, expressed as a decimal. The following sections will discuss results from running these specifications.

5. Results

5.1. (1) Full Specification: Men's and Women's

Medalshare	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
host	0.047	0.01	4.54	0	0.026	0.067	***
log_pop	0.007	0.002	4.52	0	0.004	0.01	***
log_income	0.006	0.002	2.72	0.008	0.002	0.011	***
govE	0.0002	0	0.25	0.806	0	0	
GiniFixed	-0.0004	0	-2.41	0.017	-0.001	0	**
1996	0	.	.	.	.	.	
1998	0.009	0.007	1.32	0.189	-0.004	0.022	
2000	-0.001	0.003	-0.45	0.656	-0.007	0.005	
2002	0.005	0.004	1.23	0.221	-0.003	0.013	
2004	-0.004	0.003	-1.45	0.149	-0.009	0.001	
2006	-0.002	0.004	-0.61	0.542	-0.01	0.005	
2008	-0.008	0.003	-2.67	0.009	-0.013	-0.002	***
2010	-0.005	0.003	-1.83	0.069	-0.011	0	*

2012	-0.008	0.003	-2.98	0.003	-0.013	-0.003	***
2014	-0.006	0.004	-1.67	0.098	-0.013	0.001	*
2016	-0.009	0.003	-2.88	0.005	-0.015	-0.003	***
2018	-0.008	0.004	-2.24	0.027	-0.016	-0.001	**
2020	-0.01	0.003	-3.17	0.002	-0.017	-0.004	***
2022	-0.021	0.005	-4.48	0	-0.03	-0.012	***
Constant	-0.14	0.038	-3.71	0	-0.215	-0.065	***
**** p<.01, ** p<.05, * p<.1*							

Mean dependent var	0.014	SD dependent var	0.027
R-squared	0.419	Number of obs	711
F-test	6.15	Prob > F	0
Akaike crit. (AIC)	-3483.93	Bayesian crit. (BIC)	-3397.16

From a smaller set of observations, we find that host nation status is an extremely significant determinant of Olympic Success. Over 700 observations across the span of 14 different Olympics show that at the 1% level, being the host nation increases that nation's overall share of medals by 4.7%. The population effect is also significant in this model at a 1% level, but the real effect is quite smaller than that of host status. **Table 5.1** shows that a 1% increase in population results in only a 0.7% increase in medal share. Considering that the average medal share is already only ~1% of all attainable medals, a less than one percent increase is unlikely to be considered by a home audience or government as an economically significant increase in success. Similarly, the effect of income per capita is also significant at the 1% level, but the economic significance is subtle. Here a 1% increase in real GDP per capita corresponds to a 0.6% increase in total medal share. Finally, the Gini Coefficient is also found to be a statistically significant determinant of a NOC’s medal share at the 5% level. Here, the negative coefficient highlights how higher levels of economic inequality translates to a slightly smaller share of all medals. However, this outcome is economically negligible as a 1% change in Gini is associated with a practically 0% decrease in medal share. Overall, this model explains 41% of the variance in predicted medal share. On the other hand, government effectiveness seems to have no impact on medal share. Moreover, the p-value of 0 for the model's F-statistic tells us that this overall specification is statistically significant.

Though the above specification does not control seasonal effects, in Appendix table **7.1**, I also observe seasonal effects for the summer. Though the magnitudes of impact are slightly varying, we find a nearly identical impact of hosting, income, and population. However, population is only significant at the 5% level. Moreover, one significant change is that both government effectiveness and the Gini Coefficients are insignificant. I suspect that this is because there are less barriers to entry to summer Olympic events compared to winter Olympic events which usually require a higher level of individual investment in gear, infrastructure, and training. This justification is further strengthened in table **7.2** where there is a statistically significant and negative relationship between the Gini Coefficients and medal shares. Additionally, for the Winter Olympics, there appears to be an even stronger effect of host status, which now increases medal-shares by nearly 6%. This outcome could be attributed to fewer participants in the winter compared to the summer.

Women's Vs. Men's Medal Counts: Full Specification – Year effects included but not shown.

5.2 Men's Events vs. Women's – Full Specification

Men's

mmedals	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
host	13.918	3.867	3.60	0	6.264	21.571	***
log_pop	2.323	.5	4.65	0	1.334	3.312	***
log_income	1.993	.778	2.56	.012	.453	3.532	**

govE	-.007	.035	-0.19	.849	-.076	.063	
GiniFixed	-.162	.049	-3.32	.001	-.259	-.065	***
Mean dependent var		4.203	SD dependent var			8.551	
R-squared		0.397	Number of obs			711	
F-test		11.807	Prob > F			0.000	

Women's

fmedals	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
host	9.902	4.266	2.32	.022	1.458	18.346	**
log_pop	2.58	.616	4.19	0	1.361	3.799	***
log_income	1.682	.846	1.99	.049	.007	3.356	**
govE	.006	.038	0.15	.881	-.07	.081	
GiniFixed	-.136	.056	-2.41	.017	-.247	-.024	**
Mean dependent var		3.972	SD dependent var			8.892	
R-squared		0.367	Number of obs			711	
F-test		6.673	Prob > F			0.000	

Many of magnitude differences observed in predicted success between this model and the first equation are because we now predict and explain medal counts rather than share of medals. Consistent with previous findings, the host effect is extremely positive and significant at the 1% level. According to the model, hosting results in an average increase of 14 total medals earned during that Olympic Games. Thus, the effects of homefield advantage are seemingly the strongest economically significant determinants of Olympic Success. We also find that population and income are both statistically significant at the 1% and 5% level respectively. Here, only 1% increases in population or income correspond to roughly an additional 2 medals earned for that nation, which is important considering on average, most men's teams earn an aggregate of only 3 medals during the Olympic season. Hosting seems to have a smaller level of impact as only 9 additional medals are predicted to be won for women's events when that country is hosting. However, the effect of income and population are nearly the same (1 & 2 more medals), as the men's impact. And in both cases, there is a small, statistically significant effect of inequality in both men's and women's events.

Looking further into the women's medal count, in tables **7.3** and **7.4**, we find that the host nation effect is significantly larger in the summer than in the winter. During this season, 13 additional medals are won for hosting teams, however, only 5 additional medals are typically earned for the host during the winter games. Likewise, for 1% increments in income and population, 2-3 additional medals are earned in the summer while only 1-2 additional medals are earned in the winter for the same percentage increases in independent variables.

This outcome is also consistent for men in the full specification, where summer hosts see an average of 18 additional medals, and 9 additional medals in the winter. Marginal population and income increases, lead to an additional ~3 medal gain each, but only in the summer.

5.3 Men's Events and Women's – Reduced Specification

medalshare	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
host	.056	.014	4.03	0	.029	.084	***
log_pop	.004	.001	4.89	0	.003	.006	***
log_income	.006	.001	5.69	0	.004	.008	***
1972b	0	.	.	.	.	.	
1976	-.004	.002	-2.08	.039	-.008	0	**
1980	-.004	.003	-1.30	.194	-.01	.002	
1984	-.005	.003	-1.48	.139	-.011	.002	
1988	-.008	.002	-3.31	.001	-.013	-.003	***
1992	-.009	.003	-2.74	.007	-.016	-.003	***

1994	-.006	.005	-1.28	.203	-.016	.003	
1996	-.01	.003	-2.90	.004	-.017	-.003	***
1998	-.009	.005	-1.98	.049	-.018	0	**
2000	-.01	.003	-3.06	.003	-.017	-.004	***
2002	-.01	.004	-2.35	.02	-.018	-.002	**
2004	-.012	.004	-3.35	.001	-.019	-.005	***
2006	-.013	.005	-2.83	.005	-.022	-.004	***
2008	-.015	.004	-3.76	0	-.023	-.007	***
2010	-.015	.004	-3.58	0	-.024	-.007	***
2012	-.016	.004	-3.81	0	-.024	-.008	***
2014	-.016	.005	-3.43	.001	-.025	-.007	***
2016	-.016	.004	-3.82	0	-.024	-.008	***
2018	-.017	.005	-3.52	.001	-.026	-.007	***
2020	-.016	.004	-3.87	0	-.024	-.008	***
2022	-.017	.005	-3.52	.001	-.027	-.008	***
Constant	-.094	.018	-5.34	0	-.129	-.06	***

Mean dependent var	0.009	SD dependent var	0.024
R-squared	0.307	Number of obs	2765
F-test	6.227	Prob > F	0.000

The model's findings are still consistent in the reduced specification. The purpose of this implementation of this version of the model was to verify the significance of the core explanatory variables on a wider set of information. Even as we increase the number of observations to 2700, the effect of host nation status, income, and population are all significant at either the 5% or 10% levels. Here, the host nation effect is even stronger; being a host nation is associated with a 5.6% increase medal share, as opposed to only a 4.7% increase in medal share observed prior. One feasible justification is that host nations are chosen based on their level of economic stability to ensure that they can sustain the influx of tourists as well as the construction of new infrastructure needed to hold the Olympics. Since one aspect of economic stability is income inequality, it could be that already economic stable and equitable countries are being chosen as host nations and the effect of the Gini coefficient is being captured within the host-nation dummy variable. Therefore, the inclusion of Gini could be important to avoid increases in omitted variable bias.

5.4 Men's Events and Women's Separate Counts – Reduced Specification

Men

mmedals	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
host	15.876	4.316	3.68	0	7.365	24.387	***
log_pop	1.294	.261	4.96	0	.78	1.808	***
log_income	1.618	.312	5.19	0	1.004	2.233	***
Constant	-28.37	5.747	-4.94	0	-39.702	-17.038	***

Mean dependent var	2.452	SD dependent var	7.386
R-squared	0.262	Number of obs	3031
F-test	.	Prob > F	.
Akaike crit. (AIC)	19855.144	Bayesian crit. (BIC)	20017.594

Women:

fmedals	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
host	9.638	2.657	3.63	0	4.398	14.878	***
log_pop	1.125	.252	4.47	0	.629	1.621	***
log_income	1.312	.273	4.81	0	.774	1.85	***
Constant	-25.704	5.579	-4.61	0	-36.706	-14.702	***

Mean dependent var	1.862	SD dependent var	5.952
R-squared	0.259	Number of obs	3031
F-test	.	Prob > F	.

Although the impact of being the host nation is greater for men then it is for women's events, men's events historically have constituted a larger share of medal count then women's overall medal count. Thus, it cannot be concluded that the host effect is economically greater for men then it is for women. In fact, relative to the share of total medals won, it is possible that the women's percentage increase of medal share is greater than the men's overall increase in medal share. However, because we are tracking changes in the number of medals awarded in this case, we are limited to inferences and restricted from making such a definitive conclusion.

More generally about the impact of income and population, the effect of log income is economically significant. Since the average income was reported at ~$12,000USD, a 1% change in net income is a $120 increase in income that corresponds to anywhere from a 1 -3 difference in total medal production. However, the impact on population is not as strong. Since the average country size was 39 million, this means a 1% or ~300,000 increase in population is needed for an additional 1-3 medals.

Finally, in the appendix tables **7.7** and **7.8**, when we create distinction between the summer and winter Olympics in the simplified specification, home field advantage is a larger contributor to medal share than in the wintertime. Compared to the summer, the marginal increase in medal share is 1.2% greater than it would have been in the winter, all else equal. This discrepancy is both economically and statistically significant. Considering that there are 339 medal events at the Summer Games, homefield advantage is at least 4.5 - 5 medals more impactful than it would have been in the winter.

6. Conclusion:

This study leverages mainly economic and political factors and host status to determine success at the Olympic Games. This paper borrows from previous reports by including common and extensively documented indicators such as GDP per capita, population and host status. An OLS model is used to analyze the impact of these potentially explanatory factors. Furthermore, this model builds upon those studies by incorporating popular World Governance and Economic Development Indicators, specifically **government effectiveness** and the **Gini coefficient**. Both are employed to better understand how and if quality of government and levels of income inequality affect a nation's success at the Olympic Games. Moreover, we define success in two distinct ways: 1.) as percentage or decimal share of medals earned out of all medals available, and 2.) simply the number of medals earned.

Original intentions for this analysis included plans to also explore the effect of latitude as a proxy for climate, on medal success. However, due to how unique elevation appeared to be for each country, and fixed nature over time, it soon became a pseudo-identifier for country. Thus, it was omitted from our specification because it would not be possible to observe statistically significant effects from it while also controlling for NOC.

The results show that host nation effect is consistently positive and statistically significant with the largest magnitude impact on both medal count and percentage medal share. Furthermore, the host nation's effect is stronger in the summer seasons, most likely because of ease of travel and terrain familiarity for athletes. Moreover, GDP per capita and population are also consistently significant, but as we control for season, population especially, tends to be economically insignificant.

We find that these outcomes are consistent between men's events medal counts and women's event medal counts. Hosting seems to have a stronger positive relationship with men's raw medal counts; however, it is important to consider that these raw counts are not necessarily the best interpretation of success.

Moving forward, an improvement to this model might involve quantifying gender-specific success to medal-share, rather than using medal shares for just overall success prediction.

Finally, between the Gini value and Government Effectiveness percentile rank, only Gini is a statistically significant determinant of Olympic Success. Unlike the other explanatory variables, it possesses a negative, small relationship with medal counts and medal shares. Therefore, the effect of income inequality may be somewhat negligible due to its focus mainly on income and not on infrastructure. Thus, income may be irrelevant for elite talent pools who are likely not on the poverty line. Government effectiveness is likely insignificant for several reasons such as only having 30 years of relevant data and the infrequency of the Olympics. Moreover, participants in the Olympics are likely products of elite privately funded youth and professional training programs, thus the impact of publicly funded programs are most likely not relevant to their personal success.

7. Appendix:

7.1 Summer Olympics – no gender effect

medalshare	**Coef.**	**St.Err.**	**t-value**	**p-value**	**[95% Conf**	**Interval]**	**Sig**
host	.036	.013	2.79	.006	.011	.062	***
log_pop	.007	.002	4.13	0	.004	.01	***
log_income	.006	.002	2.53	.013	.001	.011	**
govE	0	0	-0.74	.459	0	0	
GiniFixed	0	0	-1.50	.136	-.001	0	
1996	0	.	.	.	.	.	

2000	-.001	.003	-0.26	.793	-.006	.005	
2004	-.003	.002	-1.20	.233	-.007	.002	
2008	-.006	.003	-2.18	.031	-.011	-.001	**
2012	-.006	.003	-2.54	.012	-.011	-.001	**
2016	-.007	.003	-2.54	.012	-.012	-.002	**
2020	-.008	.003	-2.63	.01	-.014	-.002	***
Constant	-.14	.041	-3.41	.001	-.222	-.059	***

Mean dependent var	0.011	SD dependent var	0.021
R-squared	0.462	Number of obs	435
F-test	5.159	Prob > F	0.000

7.2 Winter Olympics Effect – no gender effect

medalshare	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
host	.058	.015	3.84	0	.028	.088	***
log_pop	.008	.002	4.03	0	.004	.013	***
log_income	.007	.003	2.36	.021	.001	.012	**
govE	0	0	1.38	.17	0	.001	
GiniFixed	-.001	0	-3.40	.001	-.002	-.001	***
1998b	0	.	.	.	.	.	
2002	-.004	.005	-0.70	.485	-.014	.007	
2006	-.016	.006	-2.81	.006	-.027	-.005	***
2010	-.019	.007	-2.76	.007	-.033	-.005	***
2014	-.02	.008	-2.64	.01	-.035	-.005	**
2018	-.022	.007	-2.88	.005	-.036	-.007	***
2022	-.025	.008	-3.03	.003	-.041	-.009	***
Constant	-.14	.036	-3.84	0	-.213	-.067	***

Mean dependent var	0.018	SD dependent var	0.033
R-squared	0.460	Number of obs	276
F-test	8.527	Prob > F	0.000

7.3 Women's Events – Summer

fmedals	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
host	13.543	6.564	2.06	.041	.547	26.538	**
log_pop	3.45	.911	3.79	0	1.646	5.254	***
log_income	2.743	1.214	2.26	.026	.339	5.146	**
govE	-.029	.051	-0.56	.579	-.13	.073	
GiniFixed	-.081	.082	-0.99	.326	-.243	.081	
Constant	-70.607	21.68	-3.26	.001	-113.531	-27.682	***
Mean dependent var		5.021	SD dependent var			10.679	
R-squared		0.411	Number of obs			435	
F-test		4.792	Prob > F			0.000	

7.4 Women's Winter

fmedals	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
host	5.643	2.186	2.58	.012	1.288	9.997	**
log_pop	1.207	.307	3.93	0	.596	1.819	***
log_income	.673	.346	1.94	.056	-.017	1.362	*
govE	.035	.022	1.60	.114	-.009	.078	
GiniFixed	-.161	.053	-3.02	.003	-.267	-.055	***
Constant	-18.968	4.57	-4.15	0	-28.07	-9.865	***
R-squared		0.412	Number of obs			276	
F-test		7.111	Prob > F			0.000	

7.5 Men's Summer

mmedals	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
host	18.206	5.725	3.18	.002	6.871	29.54	***
log_pop	3.186	.729	4.37	0	1.742	4.629	***
log_income	3.026	1.147	2.64	.009	.755	5.296	***
govE	-.042	.051	-0.83	.409	-.142	.058	
GiniFixed	-.132	.072	-1.84	.068	-.274	.01	*
1996b	0	.	.	.	.	.	
2000	.436	1.328	0.33	.743	-2.194	3.066	
2004	-1.039	1.135	-0.91	.362	-3.287	1.209	
2008	-2.28	1.285	-1.77	.078	-4.825	.264	*
2012	-3	1.218	-2.46	.015	-5.412	-.588	**
2016	-3.336	1.298	-2.57	.011	-5.906	-.765	**
2020	-4.24	1.739	-2.44	.016	-7.684	-.797	**
Constant	-64.359	18.352	-3.51	.001	-100.695	-28.024	***

Mean dependent var	5.425	SD dependent var	10.174
R-squared	0.444	Number of obs	435
F-test	6.563	Prob > F	0.000
Akaike crit. (AIC)	3020.766	Bayesian crit. (BIC)	3069.670

**** p<.01, ** p<.05, * p<.1*

7.6 Men's Events Winter

mmedals	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
host	8.969	2.372	3.78	0	4.244	13.694	***
log_pop	.901	.235	3.83	0	.433	1.369	***
log_income	.994	.428	2.32	.023	.141	1.847	**
govE	.023	.019	1.24	.218	-.014	.061	

GiniFixed	-.131	.038	-3.47	.001	-.206	-.056	***
1998b	0	.	.	.	.	.	
2002	.193	.779	0.25	.805	-1.359	1.744	
2006	-.772	.677	-1.14	.258	-2.12	.577	
2010	-1.16	.844	-1.37	.173	-2.841	.521	
2014	-1.188	.81	-1.47	.147	-2.802	.427	
2018	-1.476	.742	-1.99	.05	-2.954	.002	*
2022	-2.05	.784	-2.61	.011	-3.612	-.488	**
Constant	-17.947	5.229	-3.43	.001	-28.362	-7.532	***

Mean dependent var	2.275	SD dependent var	4.397
R-squared	0.407	Number of obs	276
F-test	10.932	Prob > F	0.000
Akaike crit. (AIC)	1479.478	Bayesian crit. (BIC)	1522.922

7.7 Reduced Specification: Summer Games – No Gender Effect

medalshare	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
host	.065	.024	2.73	.007	.018	.111	***
log_pop	.004	.001	4.40	0	.002	.005	***
log_income	.004	.001	4.44	0	.002	.006	***
1972b	0	.	.	.	.	.	
1976	-.003	.001	-2.58	.01	-.006	-.001	**
1980	-.003	.003	-0.83	.409	-.009	.004	
1984	-.003	.002	-1.48	.14	-.007	.001	
1988	-.006	.002	-3.00	.003	-.01	-.002	***
1992	-.006	.003	-2.16	.032	-.011	0	**
1996	-.006	.003	-2.00	.047	-.012	0	**
2000	-.006	.003	-2.17	.031	-.012	-.001	**
2004	-.008	.003	-2.48	.014	-.014	-.002	**

2008	-.01	.003	-2.84	.005	-.016	-.003	***
2012	-.01	.004	-2.85	.005	-.017	-.003	***
2016	-.01	.003	-2.92	.004	-.017	-.003	***
2020	-.011	.004	-2.89	.004	-.018	-.003	***
Constant	-.074	.016	-4.59	0	-.105	-.042	***

Mean dependent var	0.006	SD dependent var	0.018
R-squared	0.329	Number of obs	1943
F-test	6.693	Prob > F	0.000

7.8 Reduced Specification: Winter Games – No Gender Effect

medalshare	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
host	.043	.012	3.48	.001	.019	.068	***
log_pop	.006	.001	4.73	0	.003	.008	***
log_income	.01	.002	5.12	0	.006	.014	***
1972b	0	.	.	.	.	.	
1976	-.007	.006	-1.24	.218	-.019	.004	
1980	-.011	.005	-2.06	.041	-.021	0	**
1984	-.01	.01	-1.02	.312	-.029	.009	
1988	-.015	.007	-2.09	.039	-.03	-.001	**
1992	-.02	.008	-2.59	.011	-.036	-.005	**
1994	-.018	.008	-2.11	.037	-.034	-.001	**
1998	-.021	.008	-2.62	.01	-.036	-.005	**
2002	-.021	.008	-2.60	.01	-.037	-.005	**
2006	-.027	.009	-2.99	.003	-.044	-.009	***
2010	-.03	.009	-3.28	.001	-.048	-.012	***
2014	-.031	.009	-3.33	.001	-.049	-.012	***
2018	-.032	.009	-3.47	.001	-.05	-.014	***
2022	-.033	.01	-3.44	.001	-.052	-.014	***

Constant	-.147	.029	-5.00	0	-.205	-.089	***

Mean dependent var	0.015	SD dependent var	0.032
R-squared	0.306	Number of obs	822
F-test	4.408	Prob > F	0.000

Citations

"Athletes." Olympedia, OlyMADMen 2006 — 2023, www.olympedia.org/. Accessed 6 May 2024.

Bian '05, Xun, "Predicting Olympic Medal Counts: the Effects of Economic Development on Olympic Performance" (2005). Honors Projects. 13. https://digitalcommons.iwu.edu/econ_honproj/13

Dittmore S and Kim K (2022) Nationalistic Media Obsession With Olympic Medal Counts: The Case of the 2020 Tokyo Olympic Games. Front. Sports Act. Living 4:848071. doi: 10.3389/fspor.2022.848071

Perales, Francisco. "Mundlak - Estimates Random-Effects Regressions Adding ..." Help Mundlak, The University of Queensland - Brisbane: School of Social Science, fmwww.bc.edu/RePEc/bocode/m/mundlak.html. Accessed 6 May 2024.

Statistisches Bundesamt. "Population in The Former Territories of The Federal Republic of Germany and The German Democratic Republic from 1950 to 2016." Statista, Statista Inc., 1 Jun 2017, https://www.statista.com/statistics/1054199/population-of-east-and-west-germany/

Johan Rewilak, 2021. "The (non) determinants of Olympic success," Journal of Sports Economics, , vol. 22(5), pages 546-570, June.

"Katie Ledecky." Encyclopædia Britannica, Encyclopædia Britannica, inc., 5 May 2024, www.britannica.com/biography/Katie-Ledecky#:~:text=2012%20London%20Olympics,-Olympics%20medals&text=The%20following%20year%20Ledecky%20set,called%20World%20Aquatics)%20world%20championships.

Knight, Brett. "These 12 Countries Will Pay Their Olympians Six-Figure Bonuses for Winning Gold Medals." Forbes, Forbes Magazine, 16 Feb. 2022, www.forbes.com/sites/brettknight/2022/02/04/these-12-countries-will-pay-their-olympians-six-figure-bonuses-for-winning-gold-medals/.

"National Olympic Committees (NOC) - Olympic Movement." National Olympic Committees, © 2024 International Olympic Committee. All Rights Reserved., olympics.com/ioc/national-olympic-committees. Accessed 6 May 2024.

Noland, M., & Stahler, K. (2017). An old boys club No more. Journal of Sports Economics, 18(5), 506–536. https://doi-org.ezproxy.library.tufts.edu/10.1177/1527002515588138

Pareja, Marcos Menocal, and Contact Marcos Read more of Marcos' articles. "Paris 2024 Facts and Figures Six Months Out." Olympic, Paralympic and Commonwealth Games News, Dunsar Media Company Limited, 22 Jan. 2024, www.insidethegames.biz/articles/1143455/numbers-paris-2024-

nine-budget-
spectator#:~:text=The%20main%20protagonists%20will%20be,agents%20a%20day%20providing%20security.

Reardon, Logan. “Tokyo Olympics by Numbers: Participating Country Stats and Facts.” NBC New York, NBC New York, 22 July 2021, www.nbcnewyork.com/news/sports/tokyo-summer-olympics/tokyo-olympics-by-numbers-participating-country-stats-and-facts/3169369/.

“Russia Indicators.” Trading Economics, Copyright ©2024 TRADING ECONOMICS All Rights Reserved, tradingeconomics.com/russia/indicators. Accessed 6 May 2024.

Todd B. Potts, 2022. "Does it pay to Play by the Rules? Respect for Rule of law, Control of Corruption, and National Success at the Summer Olympics," Journal of Sports Economics, , vol. 23(2), pages 222-245, February.

“What Is Olympism?” Olympism and the Olympic Movement Back , 2024 International Olympic Committee. All Rights Reserved., olympics.com/ioc/faq/olympism-and-the-olympic-movement/what-is-olympism. Accessed 6 May 2024.

“Home: Worldwide Governance Indicators.” World Bank, World Bank Group, 23 Jan. 2024, www.worldbank.org/en/publication/worldwide-governance-indicators.

“Uganda GDP 1960-2024.” MacroTrends, MacroTrends, www.macrotrends.net/global-metrics/countries/UGA/uganda/gdp-gross-domestic-product?__cf_chl_tk=AHroHWFGdvuAhfCgM0f1DTnx_wLo.DKPZwtxuWcQsYQ-1714943001-0.0.1.1-1727. Accessed 6 May 2024.

Kufenko, Vadim; Geloso, Vincent (2019) : Who are the champions? Inequality, economic freedom and the olympics, Hohenheim Discussion Papers in Business, Economics and Social Sciences, No. 13-2019, Universität Hohenheim, Fakultät Wirtschafts- und Sozialwissenschaften, Stuttgart, https://nbn-resolving.de/urn:nbn:de:bsz:100-opus-16961

“Indicators.” World Bank Open Data, 2024 The World Bank Group, All Rights Reserved., data.worldbank.org/indicator. Accessed 6 May 2024.