Sport Management Review 15 (2012) 211–217

Contents lists available at ScienceDirect

Sport Management Review

journal homepage: www.elsevier.com/locate/smr

# Olympic medals and demo-economic factors: Novel predictors, the ex-host effect, the exact role of team size, and the "population-GDP" model revisited

George Vagenas *, Eleni Vlachokyriakou

*Dept. of Physical Education and Sport Science, National and Kapodistrian University of Athens, 41 Olgas, Dafne, Attiki 17237, Greece*

ARTICLE INFO

*Article history:*
Received 20 February 2011
Received in revised form 26 June 2011
Accepted 18 July 2011

*Keywords:*
Olympic medals
Demo-economic predictors
Log-linear regression models
Ex-host effect
Olympic team size
GDP-free models

ABSTRACT

The present study revisited the problem of estimating Olympic success by critical demo-economic indicators. The sample consisted of the 75 winner countries at the Athens 2004 Olympic Games (not previously analyzed). Medal totals were log-linearly regressed on land, population, GDP, urban population, inflation, growth rate, unemployment, labor force, health expenditures, ex-host, and team size. Multiple regression assumptions were tested with proper diagnostics including collinearity. Olympic team size was the best single predictors of Olympic medals ($R^2 = 0.690$, $p < 0.001$), and as an alternative criterion variable was significantly regressed on population, growth rate, health expenditure, and unemployment ($R^2 = 0.563$, $p < 0.001$). Medal totals were significantly regressed on population, ex-host, health expenditure, growth rate, and unemployment ($R^2 = 0.541$, $p < 0.001$). The classical population-GDP model extracted only 28% of the variance in total medals ($R^2 = 0.277$, $p < 0.001$), and this was slightly improved when combined with unemployment ($R^2 = 0.365$, $p < 0.001$). It appears that the size of the Olympic team plays the role of transmitting the composite impact of a country's size and economy to the end-phase of Olympic success. Winning Olympic medals depends on the combined potential of population, wealth, growth rate, unemployment, ex-host, and social-sport expenditures. Larger and wealthier countries win more medals by "producing" larger Olympic teams as a result of possessing more athletic talents and better support for social and sport related activities.



## 1. Introduction

The demographic and economic determinants of Olympic success have been investigated over the last four decades by researchers and analysts from diverse fields. Olympic medal counts were routinely analyzed in relation to critical aspects of the political, social, economic, and demographic profiles of the participating countries for both explanatory and predictive purposes (i.e. Hoffmann, Ging, & Ramasany, 2004). Early analysis correlated non-parametrically medal indexes to selected dichotomies reflecting political and economic systems (i.e. Ball, 1972), while the first prediction models were tested in an attempt to explain total Olympic medals by the linear combination of GDP with the dichotomy of socialist vs. non-socialist economy, land area (Levine, 1972), and population (Grimes, Kelly, & Rubin, 1974). Subsequent studies explored primarily population-GDP based specifications, enhanced their models with the additive contribution of other potential factors, examined several editions of the games, and applied more sophisticated statistical methods.

* Corresponding author. Tel.: +30 2107276164, fax: +30 2107276192.
*E-mail address:* gvagenas@phed.uoa.gr (G. Vagenas).

1441-3523/$ – see front matter © 2011 Published by Elsevier Ltd on behalf of Sport Management Association of Australia and New Zealand.
doi:10.1016/j.smr.2011.07.001

Artificial neural network techniques contrasted to ordinary and weighted least squares regressions were utilized for the concurrent assessment of many candidate correlates of medal counts from the Atlanta 1996 Games and verified the validity of land area, population, and per capita GDP as prime predictors of Olympic success (Condon, Golden, & Wasil, 1999). Logarithmic functions of population and GDP estimates combined with the previous Olympic medals and the positive effect of the home – neighboring advantage were combined into successful specifications of total and gold medal counts from all summer Games between 1952 and 2000 (Johnson & Ali, 2002, 2004). Similarly, total Olympic medals were correlated with population-based estimates of proportions in athletic talents and, then, log-linearly regressed on the population-GDP pair of predictors along with home advantage and the Soviet influence (Bernard & Busse, 2004). These two later factors were examined in a series of more inclusive analyses of all Games between London 1948 and Sydney 2000, and the resulting models combined certain indicators of climatic conditions with population and GDP (Hoffman, Ging, & Ramasany, 2002, 2004). In that respect, Western and Eastern block countries were compared on their Olympic efficiency by applying a theoretical model to estimate the added cost for each additional athlete and medal, thus confirming by econometric analysis the best way of planning an optimal level of Olympic medals "production" (Tcha, 2004).

Lozano, Villa, Guerrero, and Cortes (2002) estimated best athletes' performance for future prediction of Olympic success by using countries that won at least one medal in all Games from Los Angeles 1984 to Sydney 2000. The approach was rather complex in terms of the number of potential predictors tested, with this being indicative of the difficulty inherent to correctly specifying optimal econometric models (Studenmund, 2001, pp. 157–242). As too complex regression models tend to increase variances in the coefficients and the predictions, while oversimplified ones suffer from biased coefficients and predictions (Myers, 1990, pp. 178–180), the main methodological problem in this area has been the determination of optimal specifications. When a large number of candidate correlates of Olympic success are explored, some of them end-up to simply be distal covariates of minor or no relevance, as for example correlates such as expected life span, death rate, number of airports, and total railway length found in study (Condon et al., 1999). On the other hand, there are cases of studies in which oversimplified models of Olympic success are proposed with only two predictors (population & GDP), as for example the multiplicative function of Morton (2002) and the ordered-logit model of Andreff (2001), or even with solely one predictor (GNP), as the simplest ever published model proposed by Nevill and Stead (2003).

According to theory developed by Shughart and Tollison (1993) the problem of Olympic success is a pure economic one that requires a certain balance between expected profit (number of Olympic medals) and cost (i.e. sport financing). Therefore, a better interpretability of the phenomenon under study may result from theory-based macroeconomic specifications (Studenmund, 2001, p. 167, 171) of moderate complexity (Stevens, 1996, p. 78), as for example the one proposed by Andreff, Andreff, and Poupaux (2008), which successfully combined population and GDP per capita, with the political regime, the host effect, and the cultural differences of countries across regions worldwide. Accordingly, certain substantive improvements are tenable in this field of research if some theoretically relevant but so far untested correlates of Olympic success are incorporated into novel specifications. For example, economic growth (Alexander, 1997; Barro, 1991; Moldan & Billharz, 1995), labor force (Alexander, 1997; Maddison, 1995), inflation (Grier & Tullock, 1989; Levine & Renet, 1992; Osborne, 2006; Pollin & Zhu, 2006), unemployment (Alexander, 1997; Benerjee, Marcellino, & Masten, 2005), and health expenditure, as an essential element of social and developmental investments (Alexander, 1997; Barro, 1991; Gupta & Sommers, 1999), had not yet been examined, despite their critical role in the structure of national economies and in international economics (Alexander, 1997; World Bank, 1998).

In addition, the exact role of Olympic team size had not been fully determined statistically and interpretably, while certain originality is secured by using data from the Athens 2004 Olympics, since no study has yet examined this edition of the Games. Moreover, there is margin for methodological improvement by testing for potential collinearity in the proposed specifications (i.e. population with land), as this inherent problem of multiple regression leads, among other things, to inflated standard errors of the estimated coefficients and in turn to unstable prediction models (Aczel & Sounderpardian, 2009, p. 536; Pedhazud, 1997, p. 295; Stevens, 1996, p. 76; Studenmund, 2001, pp. 248 and 252). In light of these novel perspectives the present study aimed at investigating the correlational structure of Olympics success (medals) by exploring the combination of traditional and novel demo-economic predictors, analyzing the exact role of Olympic team size, testing novel specifications of optimal fitting and minimum multi-collinearity, and revisiting the traditional "population-GDP" solution for comparative purposes.

## 2. Methods

The sample consisted of the 75 countries that won at least one Olympic medal in the Athens 2004 Games. Only the medal winner countries were analyzed and this secured certain statistical requirements regarding raw data transformation (e.g. when 0 medals are reported) and the presence of outliers (Condon et al., 1999; Grimes et al., 1974; Morton, 2002).

Olympic success was assessed by the total number of medals won by each country (Ball, 1972; Bernard & Busse, 2004; Grimes et al. 1974; Shughart & Tollison, 1993). A total of 929 medals were shared in these Games (301 gold, 301 silver, 327 bronze). Medal tallies (gold, silver, bronze, total) per country were collected from sources available by the International Olympic Committee, the NBC, and the official book of Olympic Games. The countries that have hosted the Games were also identified and a dichotomy was added to reflect this effect.

Valid data available in World Bank and United Nations sources were collected for each participant country on selected national demographic and economic indicators for the year preceding the Games, according to relevant studies (Ball, 1972;

Bernard & Busse, 2004; Condon et al., 1999; Moosa & Smith, 2004). These indicators were land area (LAND) in $km^2$, population in mil., gross domestic product (GDP) per capita in 2000 standard $ values, % labor force, % inflation, % annual growth rate, health expenditures in % of GDP, % urban population, and % unemployment. The particular demo-economic indicators were chosen on the basis of their importance in reflecting national economic capacity and of their significant role in international comparative studies on the efficacy of the countries to be productive in major economic activities (Alexander, 1997; Banerjee, Marcellino, & Masten, 2005; Barro, 1991; Gupta & Sommers, 1999; Spangenberg, 2004).

The raw data were log-transformed to improve their distributional structure and their conformity to the statistical assumptions of multiple linear regression (i.e. Morton, 2002; Nevill & Stead, 2003). A series of regression models of the form $Y = \alpha X_1^{b1}. X_2^{b2} \ldots X_k^{bk} e$ were log-linearized to $\ln Y = \ln(\alpha) + b_1 \ln(X_1) + b_2 \ln(X_2) + \ldots + b_k \ln(X_k) + \ln(e)$, and tested for significance and per demo-economic factor ($X$'s) contribution to the variance of the Olympic medals ($Y$).

The degree to which the regression standardized residuals showed satisfactory levels of normalcy, linearity, and homoskedasticity was assessed by visual inspection of the respective residual scatterplots (Tabachnick & Fidell, 1989, pp. 130–133). The degree of collinearity among the demo-economic predictors was assessed by tolerance and variance inflation factor (VIF) values (Pedhazur, 1997, pp. 294–318; Stevens, 1996, p. 77; Studenmund, 2001, pp. 255–258). Tolerances were computed as $1 - R_j^2$, with $R_j^2$ being the % of variance of each predictor extracted by the linear combination of the other predictors, and variance inflation factor (VIF) the inverse of tolerance ($1/1 - R_j^2$). The proposed regression models were tested for over- and/or under-fitting according to Mallow's criterion (Stevens, 1996, p. 81). All statistical analyses were carried out in SPSS17 and included correlation, multiple regression, and relevant diagnostics (scatters, residual plots, collinearity). Significances were tested at the $\alpha = 0.05$ probability level. A portion of the raw data for the top and last five countries of the final Olympic medals rank is presented in Table 1.

## 3. Results

### 3.1. Correlational profile of Olympic medals

The correlations among the log-transformed values of medal tallies were high and significant ($p < 0.01$). Gold medals correlated highly with bronze ($r = 0.786$), silver ($r = 0.743$), and total ($r = 0.911$). The correlations among the log-transformed values of the demo-economic variables were lower in degree and significance. Land correlated significantly with population ($r = 0.725$) and inflation ($r = 0.379$), and health expenditures with GDP ($r = 0.493$) and urban population ($r = 0.341$).

The correlations between the log-transformed medal tallies and demo-economic factors are given in Table 2 and show several interesting trends. The three categories of medals correlated significantly with team size, land area, health expenditure, unemployment, and the ex-host dichotomy ($p < 0.05$). GDP correlated significantly only with gold medals ($r = 0.261$). The total medals showed high positive correlations with team size ($r = 0.831$), ex-host ($r = 0.529$), population ($r = 0.481$), land ($r = 0.387$), and health expenditures ($r = 0.329$), and negative correlation with unemployment ($r = -0.326$). GDP, growth rate, labor force, inflation, and urban population did not correlate significantly with medal totals ($p > 0.05$).

### 3.2. Statistical importance of Olympic team size

The full (1st) log-linear regression included all 11 independent variables at once, with total medals being the dependent variable. As expected, due to overfitting the analysis produced a highly significant model ($R^2 = 0.739$, $R^2$ *adj.* = 0.694, $SE = 0.559$, $F = 16.248$, $p = 0.000$), with team size being the only significant predictor ($\beta = 0.709$, $t = 6.559$, $p = 0.000$) and unemployment the next predictor that approached, but not reached, significance ($\beta = -0.132$, $t = -1.888$, $p = 0.064$). A second run of this model after excluding ex-host produced again a highly significant model ($R^2 = 0.733$, $R^2$ *adj.* = 0.691, $SE = 0.562$, $F = 17.587$, $p = 0.000$), with team size ($\beta = 0.761$, $t = 7.629$, $p = 0.000$) and unemployment ($\beta = -0.141$, $t = -2.015$, $p = 0.048$)

**Table 1**
Medal tallies and demo-economic indicators for the top and bottom five medal-winner countries of the Athens 2004 Olympics ($N = 75$).

| Rank/country | Gold | Silver | Bronze | Total | Team size | Ex-host | Land area | Population | Urban pop. | Labor force | Gross D.P. | Growth rate | Health expen. | Inflation | Unemploy. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 01. USA | 35 | 39 | 29 | 103 | 546 | 1 | 9161.9 | 290.8 | 80.1 | 75.2 | 35313 | 1.9 | 15.2 | 2.0 | 6.0 |
| 02. China | 32 | 17 | 14 | 63 | 392 | 0 | 9326.4 | 1288.4 | 38.6 | 82.5 | 5019 | 9.3 | 4.8 | 2.6 | 4.3 |
| 03. Russia | 27 | 27 | 38 | 92 | 457 | 1 | 16381.0 | 144.6 | 73.2 | 69.6 | 8373 | 7.8 | 6.3 | 14.0 | 7.9 |
| 04. Australia | 17 | 16 | 16 | 49 | 489 | 1 | 7682.3 | 19.9 | 87.8 | 73.6 | 29408 | 2.8 | 9.2 | 3.0 | 5.7 |
| 05. Japan | 16 | 9 | 12 | 37 | 321 | 1 | 364.5 | 127.7 | 65.6 | 72.5 | 26063 | 1.6 | 8.0 | -1.6 | 5.2 |
| 72. Colombia | 0 | 0 | 1 | 1 | 55 | 0 | 1109.5 | 43.7 | 72.1 | 74.5 | 6512 | 2.3 | 7.7 | 8.2 | 14.1 |
| 72. Eritrea | 0 | 0 | 1 | 1 | 4 | 0 | 101.0 | 4.1 | 18.8 | 75.1 | 1046 | 1.4 | 4.7 | 11.9 | 11.6 |
| 73. Mongolia | 0 | 0 | 1 | 1 | 20 | 0 | 1566.5 | 2.5 | 56.7 | 70.1 | 1649 | 4.8 | 6.5 | 12.7 | 14.2 |
| 74. Syria | 0 | 0 | 1 | 1 | 6 | 0 | 183.8 | 18.1 | 50.4 | 63.5 | 3260 | -0.9 | 5.1 | 2.2 | 12.3 |
| 75. Trinidad | 0 | 0 | 1 | 1 | 23 | 0 | 5.1 | 1.3 | 11.6 | 67.0 | 11473 | 13.1 | 3.8 | 5.4 | 10.4 |

*Note*: land in $10^3$ $km^2$, population in millions, GDP in $ per capita, all other demo-economic variables are in % values.

**Table 2**
Correlations between the log-transformed values of medal tallies and demo-economic indicators ($N = 75$).

| | Olympic medal tallies | | | |
|---|---|---|---|---|
| Demo-economic factors | ln(gold) | ln(silver) | ln(bronze) | ln(totals) |
| Ex-host | **0.446 | **0.566 | **0.486 | **0.529 |
| ln(team size) | **0.719 | **0.772 | **0.728 | **0.831 |
| ln(land area) | **0.373 | **0.365 | **0.352 | **0.387 |
| ln(population) | **0.452 | **0.469 | **0.458 | **0.481 |
| ln(urban population) | 0.163 | 0.109 | 0.152 | 0.168 |
| ln(labor force) | 0.093 | 0.085 | 0.083 | 0.096 |
| ln(GDP) | *0.261 | 0.078 | 0.154 | 0.162 |
| ln(growth rate) | 0.069 | 0.070 | 0.045 | 0.082 |
| ln(health expenditure) | **0.322 | **0.297 | *0.266 | **0.329 |
| ln(inflation) | −0.082 | −0.085 | 0.046 | −0.044 |
| ln(unemployment) | **−0.298 | **−0.316 | −0.225 | **−0.326 |

* $p < 0.05$.
** $p < 0.01$.

being the only two significant predictors. Further log-linear regression analysis revealed that these results were due to the highly significant dependence of total medals on team size ($R^2 = 0.690$, $R^2$ *adj.* = 0.686, $SE = 0.566$, $F = 162.747$, $p = 0.000$), which almost completely masked the potential contribution of some of the 9 demo-economic variables in the initial analysis.

To clarify this point, a 2nd log-linear regression analysis was carried-out with team size as dependent and 9 demo-economic indicators as independents. The produced model was highly significant (Table 3), extracted 56.3% of the variance in team size, and the four predictors were significant and almost free of collinearity (tolerance $\geq$ 0.928, VIF $\leq$ 1.078).

## 3.3. Novel log-linear regression models

After excluding team size from further analysis, a 3rd log-linear (stepwise) regression (the main one of this study) was carried-out with total medals as dependent and the 9 demo-economic indicators along with ex-host as independents. The produced model was highly significant (Table 4), extracted 54.1% of the variance in total medals, and the five predictors were significant and of very low collinearity (tolerance $\geq$ 0.739, VIF $\leq$ 1.353). In light of the rather high ratio (15) of cases ($N = 75$) to predictors ($k = 5$), the validity of this novel specification seems to be secured.

To regress total medals only on pure demo-economic indicators a 4th log-linear regression model was tested after excluding ex-host. This model was also highly significant ($F = 15.548$, $p < 0.001$), extracted 47% of the variance in total medals ($R^2 = 0.470$, adj. $R^2 = 0.440$, $SE = 0.756$), and took the form (predictors in order of relative importance): ln(total medals) = −11.347 + 0.327ln(population) + 1.160ln(health expenditure) + 1.894ln(growth rate) − 0.494ln(unemploy-0.494ln(unemployment) + ln($\varepsilon$). The four predictors were significant and almost free of collinearity (tolerance $\geq$ 0.928, VIF $\leq$ 1.078).

**Table 3**
Statistics for the log-linear regression of Olympic team size on four significant non-collinear demo-economic predictors ($N = 75$).

| Predictors | $b$ | SE | $\beta$ | $t$ | Sig. |
|---|---|---|---|---|---|
| Constant | −12.996 | 2.932 | | −4.432 | 0.000 |
| ln(health expenditure) | 1.748 | 0.264 | 0.543 | 6.626 | 0.000 |
| ln(population) | 0.351 | 0.056 | 0.501 | 6.309 | 0.000 |
| ln(growth rate) | 2.488 | 0.708 | 0.288 | 3.512 | 0.001 |
| ln(unemployment) | −0.310 | 0.146 | −0.169 | −2.128 | 0.037 |

Dep. variable: ln(team size), $R^2 = 0.563$, Adj. $R^2 = 0.538$, SE = 0.720, $F = 22.527$, $p < 0.001$.
Tolerance (& VIF) ranges from 0.928 (1.078) for growth rate to 0.990 (1.010) for population.

**Table 4**
Statistics for the log-linear regression of total Olympic medals on five significant non-collinear predictors ($N = 75$).

| Predictors | $b$ | SE | $\beta$ | $t$ | Sig. |
|---|---|---|---|---|---|
| Constant | −10.578 | 2.896 | | −3.652 | 0.001 |
| ln(population) | 0.274 | 0.0597 | 0.410 | 4.804 | 0.000 |
| Ex-host | 0.759 | 0.232 | 0.310 | 3.264 | 0.002 |
| ln(health expenditure) | 0.806 | 0.281 | 0.263 | 2.866 | 0.006 |
| ln(growth rate) | 2.009 | 0.692 | 0.244 | 2.878 | 0.005 |
| ln(unemployment) | −0.395 | 0.147 | −0.226 | −2.694 | 0.009 |

Dep. variable: ln(total medals), $R^2 = 0.541$, Adj. $R^2 = 0.508$, SE = 0.709, $F = 16.284$, $p < 0.001$.
Tolerance (& VIF) ranges from 0.739 (1.353) for ex-host to 0.945 (1.058) for unemployment.

An alternative 5th model was then tested with index 3Gold + 2Silver + Bronze as dependent, according to previous studies (i.e. Lozano, Villa, Guerrero, & Cortes, 2002). This model was also highly significant ($F = 14.332$, $p < 0.001$), extracted 51% of the variance in the composite medals index ($R^2 = 0.509$, adj. $R^2 = 0.474$, $SE = 0.931$), and took the form (predictors in order of relative importance): ln(3G + 2S + B) = −12.914 + 0.330ln(population) + 1.111ln(health expenditure) − 0.591ln(unemployment) + 0.801(ex-host) + 2.457ln(growth rate) + ln($\varepsilon$). The five predictors were significant and of very low collinearity (tolerance ≥ 0.739, VIF ≤ 1.353).

These three novel specifications were checked for proper fitting according to Mallow's criterion $C_p = p + (s^2 - \sigma'^2)(N - p)/\sigma'^2$, with $s^2$ the residual variance of the evaluated model, $\sigma'^2$ an the residual variance estimate of the full model, $N$ the sample size, $p = k + 1$, and $k$ the total number of predictors (Stevens, 1996, p. 81). If $C_p = p$ then the model is optimally fitted. In our case $p = 11 + 1 = 12$, $s^2 = 0.3102$, $\sigma'^2 = 0.4956$, $N = 75$, and $C_p = p = 12$, with this indicating optimal fitting in specifications with five or four predictors.

### 3.4. The GDP-based regression models

To test the traditional "population-GDP" solution a standard log-linear regression of total medals with GDP as the starting variable was applied. The produced model was significant ($F = 13.587$, $p < 0.001$), extracted only 36.5% the variance in total medals ($R^2 = 0.365$, adj. $R^2 = 0.338$, $SE = 0.823$), and took the form (predictors in order of relative importance): ln(total medals) = −3.911 + 0.324ln(population) + −0.519ln(unemployment) + 0.170ln(GDP) + ln(e). The three predictors were significant and almost free of collinearity (tolerance ≥ 0.986, VIF ≤ 1.015).

Finally, the "population-GDP" model alone was also tested and proved to be significant ($F = 13.808$, $p < 0.001$), extracted the much lower variance of 27.7% in total medals ($R^2 = 0.277$, *adj.* $R^2 = 0.257$, $SE = 0.8741$), and took the form (predictors in order of relative importance): ln(total medals) = −5.231 + 0.336ln(population) + 0.171ln(GDP) + ln(e).

## 4. Discussion

### 4.1. Correlates of Olympic medals and the mediating team size

A trend for high correlations with some of the demo-economic variables (land, population, GDP, health expenditure, urban population) was observed for Olympic team size. GDP, the classical prime correlate of Olympic success, showed a significant correlation exclusively to the gold medals, with this being setting the basis for re-assessing its role in interpreting a country's efficiency in winning Olympic medals. Besides team size, total medals were also correlated positively to the ex-host dichotomy, the population and the geographical size of the participating countries, their expenditures on health and, negatively, to the critical demographic variable of unemployment (Table 2).

Our analysis initially led to a single predictor (team size) model which explained one of the highest ever published percent variance in total medals (74%). However, since this result was partly due to overfitting and multi-collinearity, team size was treated as the dependent in the next step of the analysis. The log-linear combination of four non-collinear highly significant predictors (population, health expenditure, growth rate, unemployment) explained a rather large proportion of the variance in this factor (56%). Considering that team size correlated highly with the three categories of medal tallies as well (Table 2), these findings taken together seem to imply the deterministic mediating role of this factor in the "production" process of winning Olympic medals. There is a possibility though for several athletes that pass the performance criteria set by the International Athletic Federations to have been excluded from their national Olympic teams for various unknown reasons. However this variation may not alter this factor's statistics, as it is expected to be small in size and proportional to the final Olympic team sizes.

The importance of team size has been noticed previously in Baimbridge's (1998) analysis of Olympic uncertainty, where the number of athletes determined significantly the numbed of Olympic medals won. Taking into consideration Szymanski's (2000) theoretical explanation, it can be inferred that large sized countries possess the resources to cover expenses proportionally sufficient to those of the health related sector, and due to the lesser long-term unemployment-based structural problems that may affect the well-being of societies, they are possessing a higher competence for athletic talents to join their Olympic teams and, therefore, a higher probability to win medals.

### 4.2. Novel interpretations of Olympic success and the population-GDP model

The exclusion of team size from subsequent exploratory regression steps permitted the determination of a novel log-linear specification consisting of five predictors: population, ex-host, health expenditure, growth rate, and unemployment, negatively. This model extracted 55% of the variance in Olympic medals, which is higher than those previously published (e.g. Morton, 2000). The same five predictors resulted to a less cohesive model when total medals were substituted by the classical indexing scheme "3G + 2S + B", as proposed elsewhere (i.e., Lozano et al., 2002).

These results confirm the primary role of population and not of GDP, differentiate the traditional population-GDP based theory of Olympic success, and leave some space for a more accurate and comprehensive explanation of the phenomenon. Besides the definite role of the novel economic determinants of growth rate and unemployment, along with the additive effects of ex-host, this study proposes in place of GDP a more efficient explanatory role by health expenditure. The countries that have hosted the Games have a definite advantage over the rest in terms of lesser costs and better conditions for Olympic preparation (mid-term effect), and (partially) better sporting systems, programs, and Olympic traditions (log term effect). On the other hand, health expenditure has a better proximity than GDP to the process of winning Olympic medals, as it constitutes a strong indicator of the national economy's capacity to invest in social and health related activities (Moosa & Smith, 2004). In a previous study by Butter and van der Tak (1995) an attempt to establish the level of welfare as a major determinant of Olympic success was statistically verified, however, not by a direct health related indicator (as in the present study), but by using either national income estimates or some multidimensional welfare indicators, which, however, did not outperform national income in this respect. Our results confirm Butter and van der Tak's (1995) expectation of "performance in sports to be dependent upon the state of health of the population".

The secondary statistical importance of GDP was a bit of a surprise, since this factor has been a prime explanatory one in the majority of previous models of Olympic success (i.e., Andreff, 2001; Andreff et al., 2008; Hoffman et al., 2002; Johnson & Ali, 2004; Morton, 2002). Apparently, Olympic success is too complex as an empirical problem to be dealt with and ultimately interpreted by the classical population-GDP solution. Within the limitations of this edition of the Games, this finding can be explained by the fact that GDP correlated significantly only with team size ($r = 0.275$) and health expenditures ($r = 0.493$). This verifies our theoretical expectation for health expenditures to be a more proximal economic correlate of Olympic success than GDP. Its presence in the respective log-linear regression analyses shifted GDP out of the estimated collinearity – free specifications. In the absence of accurate and systematic estimates of sport expenditures (Taks & Kesenne, 2000) or of other more relevant meso-level factors (De Bosscher, De Knop, van Bottenburg, & Shibli, 2006) this finding confirms its efficiency as the best available estimator of the countries' capacity to support sports and particularly the costs of Olympic preparation per se.

In this regard, among the novel aspects of this study is the significant explanatory merit of three previously untested economic factors. These were growth rate (Alexander, 1997; Barro, 1991; Moldan & Billharz, 1995), health expenditures, as part of national social expenditure (Alexander, 1997; Barro, 1991; Gupta & Sommers, 1999), and unemployment (Alexander, 1997; Banerjee, Marcellino, & Masten, 2005). Unemployment, in particular, exhibited a constant negative explanatory role in our models, outperformed GDP, and certainly reflects a substantial part of the negative side of the national economies. High rates of it lead to lower productivity, lesser capacity for investing in the social sectors, and, therefore, to lower capital funds directed to the sports sector and particularly to the Olympic team support (Levine, 1972). These three factors are closely related to the wealth of the countries (Alexander, 1997) and, their composite impact in predicting Olympic success, as determined by the final medal counts rank (i.e., Moosa & Smith, 2004), is superior, for example, to that of the more general macroeconomic index of GDP.

In spite of their importance as demo-economic elements of the participating countries, labor force (Maddison, 1995; Alexander, 1997) and inflation (Grier & Tullock, 1989; Levine & Renet, 1992; Osborne, 2006; Pollin & Zhu, 2006) appear not to constitute correlates of Olympic success, probably due to their yearly and especially among consecutive Games variation. No significant contribution was also shown by urban population and land (Condon et al., 1999; Levine, 1972), with this being due to their high multi-collinearity with some of the other demo-economic variables tested in this study (population).

## 5. Conclusions and recommendations

Based on the findings of this study several conclusions can be drawn about the problem of Olympic success. Firstly, it appears that Olympic team size is the best single predictor and at the same time a unique ultimate transmitter of the size and the socio-economic impact of the participating countries to the process of winning medals. Secondly, the novel determinants of growth rate, unemployment, and health expenditures (as a gross analogy to public spending in supporting social and sports related sectors) are clearly significant predictors of Olympic medals, in log-linear combination to the size of the population and the ex-host effect. Thirdly, the novel specifications proposed clearly improve previous ones methodologically and interpretably by better inclusiveness in terms of the number of distinct aspects needed for interpretation, higher proportions of extracted variance in the dependent variable (Olympic medals), lower standard error of prediction estimates, and the low levels of collinearity among the predictors retained. Fourthly, the traditional "population-GDP" approach proved to be inferior to the ones proposed in this study, and, therefore, insufficient in explaining Olympic success on the basis of the number of medals won by the participating countries.

Future studies may incorporate into this methodological approach data relating to previous Olympic Games and particularly the history of the participating countries in medal shares, their team sizes over a number of consecutive Olympiads, and, perhaps, the number of different events represented by their athletes. It appears that the GDP based solutions need to be replaced by the inclusion of some gross estimates of Olympic preparation financing or even of sport related investments especially in the public sectors. Economic figures more proximal to the medals' "production" are required to enhance the statistical search of the problem, as there are certain constants and variants in this problem. For example, athletic tradition or the level of culture and education, are some constants to be taken into consideration (Lui & Suen, 2008; Tan & Green, 2008). On the other hand, the demo-economic correlates of Olympic success are variants and will always show some patterns of change among consecutive Games. These factors may as well be analyzed by proper statistical methods (e.g. non-parametric or even non-linear analysis), given their difficulty to be accurately quantified as parameters of a comprehensive model for the interpretation of Olympic success.

## References


Aczel, A., & Sounderpardian, J. (2009). *Complete business statistics* (7th ed.). McGraw-Hill: New York.

Alexander, W. R. J. (1997). Inflation and economic growth: Evidence from a growth equation. *Applied Economics, 29*, 233–238.

Andreff, W. (2001). The correlation between Economic Underdevelopment and Sport. *European Sport Management Quarterly, 1*, 251–279.

Andreff, M., Andeff, W., & Poupaux, S. (2008). Economic Determinants of the Olympic Performance: Forecasts of the medals to be won at Beijin Olympic Games. *Revue d' Economie Politique, 118*(2), 135–169.

Baimbridge, M. (1998). Outcome uncertainty in sporting competition: The Olympic Games 1896–1996. *Applied Economics Letters, 5*(3), 161–164.

Ball, D. W. (1972). Olympic games competition: Structural correlates of national success. *International Journal of Competive Sociology, 15*, 186–200.

Barro, J. R. (1991). Economic growth in a cross-section of countries. *Quarterly Journal of Economics, 106*, 407–444.

Benerjee, A., Marcellino, M., & Masten, I. (2005). Leading indicators for euro – Area inflation and GDP growth. *Oxford Bulletin of Economics and Statistics, 67*, 785–812.

Bernard, A. B., & Busse, M. R. (2004). Who wins the Olympic games: Economic resources and medal totals. *The Review of Economics and Statistics, 86*, 413–417.

Butter, A. G. F. den, & van der Tak, C. M. (1995). Olympic medals as an indicator of social welfare. *Social Indicators Research, 35*, 27–37.

Condon, E. M., Golden, B. L., & Wasil, E. A. (1999). Predicting the success of nations at the summer Olympic using neural networks. *Computers and Operations Research, 26*, 1243–1265.

De Bosscher, V., De Knop, P., van Bottenburg, M., & Shibli, S. (2006). A conceptual framework for analysing sports policy factors leading to international sporting success. *European Sport Management Quarterly, 6*(2), 185–215.

Grier, K. B., & Tullock, G. (1989). An empirical analysis of cross-national economic growth, 1951–80. *Journal of Monetary Economics, 24*, 259–276.

Grimes, A. R., Kelly, W. J., & Rubin, P. H. (1974). A socioeconomic model of national Olympic performance. *Social Science Quarterly, 55*, 777–782.

Gupta, S. K., & Sommers, P. M. (1999). A simple cross-section model of economic growth stands the test of time. *Applied Economic Letters, 6*, 601–603.

Hoffman, R., Ging, L. C., & Ramasany, B. (2002). Public policy and Olympic success. *Applied Economics Letters, 9*, 545–548.

Hoffmann, R., Ging, L. C., & Remasany, B. (2004). Olympic success and ASEAN Countries: Economic analysis and policy implications. *Journal of Sports Economics, 5*, 262–276.

Johnson, D. K. N., & Ali, A. (2004). A tale of two seasons: Participation and medal counts at the summer and winter Olympic Games. *Social Science Quarterly, 85*, 974–993.

Levine, N. (1972). Why do countries win Olympic medals? Some structural correlates of Olympic games success. *Sociology and Social Research, 58*, 353–360.

Levine, R., & Renet, D. (1992). A sensitive analysis of cross-country growth regressions. *The American Economic Review, 82*, 942–963.

Lozano, S., Villa, G., Guerrero, F., & Cortes, P. (2002). Measuring the performance of nations at the summer Olympics using data envelopment analysis. *Journal of the Operational Research Society, 53*, 501–511.

Lui, H. K., & Suen, W. (2008). Men, money, and medals: An econometric analysis of the Olympic games. *Pacific Economic Review, 13*(1), 1–16.

Maddison, A. (1995). Explaining the economic performance of nations, 1820–1989. *Explaining the Economic Performance of Nations,* 9–199.

Moldan, B., & Billharz, S. (1995). Indicators and their use: information for decision-making. In J. Wiley (Ed.), *Sustainability indicators* (pp. 13–27). Chichester, NY.

Moosa, I. A., & Smith, L. (2004). Economic development indicators as determinants of medal winning at the Sydney Olympics: An extreme bounds analysis. *Australia Economic Papers, 43*, 288–301.

Morton, H. R. (2002). Who won the Sydney 2000 Olympics? An allometric approach. *The Statistician, 51*, 147–155.

Myers, R. H. (1990). *Classical and modern regression with applications* (2nd ed.). Duxbury Press (PWS-Kent Publishing Company).

Nevill, A., & Stead, D. (2003). The relationship between national sporting success and gross national product: A law of diminishing returns. *Journal of Sports Sciences, 21*, 283–284.

Osborne, E. (2006). The sources of growth at different stages of development. *Contemporary Economic policy, 25*, 536–547.

Pollin, R., & Zhu, A. (2006). Inflation and economic growth: A cross-country non-linear analysis. *Journal of Post Keynesian Economics, 28*, 593–614.

Pedhazur, E. J. (1997). *Multiple regression in behavioral research – Explanation and prediction* (3rd ed.). Wadsworth – Thomson Learning: London.

Shughart, W. F., & Tollison, R. D. (1993). Going for the gold: Property rights and athletic effort in transitional economies. *Kyklos, 46*, 263–272.

Spangenberg, J. H. (2004). Reconciling sustainability and growth: Criteria, indicators, policies. *Sustainable Development, 12*, 74–86.

Stevens, J. (1996). *Applied multivariate statistics for the social sciences* (3rd ed.). Lawrence Erlbaum Associates, Publishers: New Jersey.

Studenmund, A. H. (2001). *Using econometrics – A practical guide* (4th ed.). Addison, Wesley, Longman: New York.

Szymanski, S. (2000). The market for Olympic Gold Medals. *World Economics, 1*(4), 2007–2214.

Tabachnick, B., & Fidell, L. (1989). *Using multivariate statistics* (2nd ed.). New York: Harper & Row Publ.

Taks, M., & Kesenne, S. (2000). The economic significance of sport in Flanders. *Journal of sport Management, 14*, 342–365.

Tan, T.-C., & Green, M. (2008). Analyzing China's Drive for Olympic Success in 2008. *International Journal of the History of Sport, 25*(3), 314–338.

Tcha, M. (2004). The color of medals - an economic analysis of eastern and western blocs' performance in the olympics. *Journal of Sport Economics, 5*(4), 311–328.