NOTES

WHO WINS THE OLYMPIC GAMES: ECONOMIC RESOURCES AND MEDAL TOTALS

Andrew B. Bernard and Meghan R. Busse*

Abstract—This paper examines determinants of Olympic success at the country level. Does the United States win its fair share of Olympic medals? Why does China win only 6% of the medals even though it has one-fifth of the world's population? We consider the role of population and economic resources in determining medal totals from 1960 to 1996. At the margin, population and income per capita have similar effects, suggesting that both a large population and high per capita GDP are needed to generate high medal totals. We also provide out-of-sample predictions for the 2000 Olympics in Sydney.

I. Introduction

In this paper, we ask the straightforward question of how many Olympic medals countries should be expected to win by considering what factors influence national Olympic success.[1] Most Olympic medal predictions assess athletic talent sport by sport and predict winners in each event. We follow a different path by generalizing from individual sports. Although this has the disadvantage of missing nation-specific expertise in a particular event, it has the advantage of averaging the random component inherent in individual competition, enabling us to make more accurate predictions of national medal totals.

Even the most ardent xenophobes would not suggest that a single country should win all the medals at a given Olympic Games. The real question is how many medals qualify as a successful performance by a national team. Clearly, population should play a role in determining country medal totals. Larger countries have a deeper pool of talented athletes and thus a greater chance at fielding medal winners. We present and test a simple theory of medal success based on population but find that pure population levels are not sufficient to explain national totals. If they were, China, India, Indonesia, and Bangladesh, with just over 43% of the world's population, would have won more than 6% of total medals in 1996. To this end, we extend the population-based model to include a measure of resources per person in the form of GDP per capita.[2]

The addition of per capita GDP dramatically improves the ability of the model to fit the data. Whereas China, India, Indonesia, and Bangladesh have a huge share of world population, together they account for under 5% of world GDP in 1996, roughly equal to their medal share. Real GDP is the best single predictor of a country's Olympic performance. Population and per capita GDP contribute equally at the margin, implying that two countries with identical GDP but different populations and per capita GDP will win the same number of medals.

GDP is not the whole story. Host countries typically win an additional 1.8% of the medals beyond what would be predicted by their GDP alone. The forced mobilization of resources by governments clearly can also play a role in medal totals. On average, the Soviet Union and Eastern Bloc countries had medal shares more than 3 percentage points higher than predicted by their GDP.[3]

The rest of the paper is organized as follows. Section II presents two simple models for analyzing country medal totals. Section III describes the data and the empirical results. Section IV presents out-of-sample forecasts for the Sydney 2000 Olympics, and section V concludes.

II. A Simple Theory of Population and Olympic Success

We start by considering the underlying distribution of athletic talent. If we think of countries as being arbitrary divisions of the world population, then we should expect to find medal-caliber athletes in proportion to the country's share of world population. One caveat is that not every country participates in the Olympics. The actual relationship predicted by the talent distribution is that the expected medal share accruing to a country should be equal to its share of the total population of countries participating in the Olympics:

$$\mathrm{E}(medalshare_{it}) = \frac{medals_{it}}{\sum_j medals_{jt}} = \frac{population_{it}}{\sum_j population_{jt}} = popshare_{it}. \quad (1)$$

Equation (1) can be tested empirically in the form of a tobit, with the results shown in column I of table 1. Although the population share is positive and significant, the estimated coefficient is significantly below 1.[4]

There are several reasons related to the structure of the Olympics that help explain why the relationship in equation (1) does not hold. First, countries cannot send athletes in proportion to their populations for each event, for example, in team competitions, where each country has at most one entry. Second, in medal counts, team events count as one medal even though a country must provide a number of athletes. Finally, the number of athletes from each country is determined by the IOC in negotiation with the country's Olympic committee. As a result, not all the Olympic caliber athletes from a large country are able to participate.

A. *Economic Resources and Olympic Medals*

To augment our specification, we examine the role of economic resources in generating Olympic medals. We choose to frame our analysis in terms of a production technology. In the previous section, we assumed that talented athletes were randomly distributed in the

Received for publication June 5, 2001. Revision accepted for publication December 4, 2002.

* Tuck School of Business at Dartmouth and National Bureau of Economic Research; and Haas School of Business, UC Berkeley, respectively.

We thank Julio Teran for valuable research assistance. A version of the paper was distributed before the Sydney Olympics on August 28, 2000 and is available at http://www.andrew.bernard.org. Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

[1] All references to the Olympics or Games refer to the Summer Games.

[2] In a contemporaneous paper, Johnson and Ali (2000) also investigate the economic and political determinants of participation and medal totals at the summer games and make out-of-sample predictions. Earlier papers in the area include Ball (1972), Levine (1974), and Grimes, Kelly, and Rubin (1974).

[3] Shughart and Tollison (1993) argue that the change in the structure in economic incentives in the former Soviet countries is responsible for their lower medal totals in the 1992 Olympics.

[4] In all our estimated specifications we include year dummies and correct for heteroskedasticity and autocorrelation in the tobit standard errors. See Busse and Bernard (2002) for the econometric methodology.

TABLE 1.—TOBIT OF MEDAL SHARE ON POPULATION AND GDP

Independent Variable	Medal Share I	II	III	IV
Population share	0.5753			
	(0.1758)			
Log population		0.0152		0.0163
		(0.0024)		(0.0022)
Log GDP per capita			0.0115	0.0142
			(0.0018)	(0.0017)
LR test				4.435 (0.03)
Year dummies	Yes	Yes	Yes	Yes
Log likelihood	468.72	581.46	476.66	723.93
Observations	1254	1254	1254	1254

Note: Heteroskedastic- and autocorrelation-consistent standard errors are in parentheses (see Busse & Bernard, 2002). The LR test reports the test of the equality of the coefficients on log GDP per capita and log population (and the p-value). It is distributed *chi*-squared with 1 degree of freedom.

world population. However, developing Olympic caliber athletes requires considerable expenditure on facilities and personnel. Wealthier countries are more likely to have individuals, organizations, or governments willing to make such an investment. Wealthier countries are also more likely to have athletics as a part of schooling and to have leisure time to devote to sports. As a result we include a measure of real GDP per capita in a model of Olympic medal production.

Our production function for generating Olympic caliber athletes for a country i in year t requires people, money, and some organizational ability:

$$T_{it} = f(N_{it}, Y_{it}, A_{it}), \tag{2}$$

where N_{it} is the population, Y_{it} is the GDP, and A_{it} is the organizational ability of the country. The share of Olympic medals, M^*_{it}, won by a country is a function of the talent in it:

$$\mathrm{E}\left(\frac{medals_{it}}{\sum_j medals_{jt}}\right) = M^*_{it} = g(T_{it}).$$

There is no theoretical guidance on the precise form of either $f(\cdot)$ or $g(\cdot)$. We use a Cobb-Douglas production function in population (N_{it}) and national income (Y_{it}) for the production of Olympic talent, and a log function for the translation of relative talent to medal shares:

$$\begin{aligned} T_{it} &= A_{it}N^{\gamma}_{it}Y^{\theta}_{it}, \\ M^*_{it} &= \ln \frac{T_{it}}{\sum_j T_{jt}}. \end{aligned} \tag{3}$$

This yields the following specification for medal shares:

$$M_{it} = \begin{cases} \ln A_{it} + \gamma \ln N_{it} + \theta \ln Y_{it} - \ln \sum_j T_{jt} & \text{if } M^*_{it} \geq 0, \\ 0 & \text{if } M^*_{it} < 0. \end{cases} \tag{4}$$

Because national income can be expressed as the product of population and per capita income, we will actually estimate a specification of the form

$$M_{it} = \begin{cases} C + \alpha \ln N_{it} + \beta \ln (Y/N)_{it} + d_t + \nu_i + \epsilon_{it} & \text{if } M^*_{it} \geq 0, \\ 0 & \text{if } M^*_{it} < 0, \end{cases} \tag{5}$$

where d_t is a year dummy included to capture changes in the total pool of talent and in the number of countries participating, as well as the changing number of sports; ν_i is a country random effect; and ϵ_{it} is a normally distributed error term.

III. Data and Results

The data for this project consist of two main components: Olympic medal counts and socioeconomic indicators. We obtained the medal data from Wallechinsky (1992) and direct correspondence with the International Olympic Committee (IOC). We would prefer to have a range of socioeconomic indicators for each country. However, the difficulty of obtaining such measures for more than 150 countries over 30 years precludes us from considering anything but GDP and population. Our primary source for population and GDP data is the World Bank.[5] Population figures could be found fairly readily; GDP measures were more difficult. For some countries, it was necessary to interpolate or extrapolate using either reported or imputed growth rates. All GDP figures are converted to 1995 U.S. dollars using current exchange rates.

A. *Results*

In this section, we report results on the relationship between medal shares and population, income per capita, and total GDP. Univariate specifications with log population and log GDP per capita, reported in columns II and III of table 1, raise the log likelihood over that with population shares alone.

Column IV estimates equation (5), including both log population and log GDP per capita. The variables are positive and significant at the 1% level and have similar magnitude. The log likelihood increases markedly over the univariate specifications. The similarity of the coefficients on log population and log GDP per capita suggests that log GDP is the relevant determinant of country medal shares. However, a likelihood ratio test rejects the equality of these coefficients at the 5% level.

The coefficients estimated in this section can be loosely interpreted to mean that if the average country were to double its total GDP, it could expect its medal share to rise by 1%–1.5% of the total medals awarded.

B. *Additions to the Model*

The empirical specification given in equation (5) shifts all country-specific information not included in GDP and population into the error term. In this section we explore some of the additional factors that might augment or diminish medal shares, including the advantages of hosting, the medal premium enjoyed by the former Soviet Union and its satellites, and the role of large-scale boycotts.

Hosts have several potential advantages over other Olympic participants. First, the cost of attending the Olympics for individual athletes is minimized. In addition, host countries can tailor facilities to meet the needs of their athletes and may gain an edge if home crowd enthusiasm sways judges. Individual athletes may be more motivated to achieve Olympic fame when the events are conducted in front of friends and family. Finally, host countries are influential in the addition of new sports to the Games themselves.

[5] We also used United Nations data sources, the *CIA Factbook,* the *Economist* magazine, and the *Taiwan Statistical Planning Book.*

TABLE 2.—TOBIT OF MEDAL SHARE ON EXPANDED EXPLANATORY SET

Independent Variable	Medal Share			
	I	II	III	IV
Log population	0.0128	0.0127	0.0083	0.0064
	(0.0016)	(0.0017)	(0.0011)	(0.0005)
Log GDP per capita	0.0126	0.0125	0.0098	0.0062
	(0.0014)	(0.0015)	(0.0010)	(0.0006)
Host	0.0605	0.0241	0.0122	0.0179
	(0.0221)	(0.0077)	(0.0077)	(0.0067)
Soviet	0.0666	0.0610	0.0300	0.0338
	(0.0088)	(0.0085)	(0.0136)	(0.0031)
Planned (non-Soviet)	0.0177	0.0161	0.0174	0.0101
	(0.0027)	(0.0031)	(0.0041)	(0.0051)
Lagged medal share				0.7333
				(0.0340)
LR test	0.063 (0.80)	0.059 (0.81)	8.400 (0.00)	0.031 (0.86)
Boycott years	Yes	No	No	No
Year dummies	Yes	Yes	Yes	Yes
Random effects	No	No	Yes	No
Log likelihood	862.06	738.23	984.04	794.03
Observations	1254	1036	1036	885

Note: Heteroskedastic- and autocorrelation-consistent standard errors are in parentheses (see Busse & Bernard, 2002). The LR test reports the test of the equality of the coefficients on log GDP per capita and log population (and the *p*-value). It is distributed *chi*-squared with 1 degree of freedom.

One of the most interesting questions regarding Olympic medal totals concerns the ability of countries, especially the former Soviet Union and Eastern European countries, to "manufacture" gold medals. Unconditional medal totals cannot tell us how successful they were at mobilizing resources. We create two dummy variables to capture these effects. The first covers countries distinctly inside the Soviet sphere of influence, and the second includes other nonmarket, typically communist, countries.[6] We consider the additional medals for these groups after controlling for income and population to provide an estimate of the power of central planning in the Olympic race. The resulting specification is

$$M_{it} = C + \alpha \ln N_{it} + \beta \ln \left(\frac{Y}{N}\right)_{it} + Host_{it} + Soviet_{it} + Planned_{it} + d_t + \nu_i + \epsilon_{it}. \quad (6)$$

Two Olympics were subject to large-scale boycotts: those in 1980 and in 1984.[7] The coefficients on the host dummy and the dummies for the centrally planned economies are likely to be particularly sensitive to the inclusion of these Games. Table 2 reports the tobit specification for medal shares with and without the boycotted years, in columns I and II.

The results for population and GDP per capita are largely unchanged in that they remain positive and significant. Now, however, we cannot reject the hypothesis that the coefficients are equal to each other. The Soviet countries have medal shares more than 6.1 percentage points higher than other countries. The other planned economies have shares that are higher by roughly 1.6 percentage points. Neither of these effects is sensitive to the exclusion of the boycotted Games.

The host effect on medal totals is also positive and significant. The bump in medal share from hosting a nonboycotted Olympics is more than 2 percentage points. During the boycotted Games the host effects were enormous, on the order of 19 percentage points, suggesting that the United States and Soviet Union were the prime beneficiaries from each other's boycotts in terms of medal counts. In column III, we report results for nonboycott years with country random effects to control for any persistent country-specific ability to produce medals. The results are broadly similar, although we reject the equality of the coefficients on population and GDP per capita.

We finish by considering the adequacy of our sparse specification for the purposes of prediction by presenting the results visually. Figure 1 shows the relationship between the predicted medal shares and actual medal shares for 1996 from our augmented tobit excluding the boycotted years (column II of table 2). The model underpredicts medal shares at both the low and high ends of the range and overpredicts in the middle. Although the additions of log GDP per capita and several dummies have improved the fit substantially, the current model is lacking in overall predictive power.

C. *Time to Build*

Until now we have implicitly modeled the production of Olympic athletes and medals as a within-period flow process with potentially persistent country-specific organizational capabilities. However, it is quite likely that Olympic athletes are more similar to durable capital goods in that they may provide medal potential over several Olympics. This would suggest that investments for one Olympics may increase the chance of winning medals in subsequent Olympics. To capture such effects we add lagged medal shares to our empirical specification[8]:

$$M_{it} = C + (1 - \delta) M_{it-1} + \alpha \ln N_{it} + \beta \ln \left(\frac{Y}{N}\right)_{it} + \delta_t + \epsilon_{it}. \quad (7)$$

Results from this specification are given in column IV of table 2. Because of the inclusion of the lagged dependent variable, we omit the boycotted games from the sample, and also the 1988 games, because the 1984 medal shares are distorted by the Soviet-led boycott. The coefficients on population and per capita GDP are again significant and statistically equal to each other. Lagged medal share has a coefficient of 0.73 and is strongly significant. The estimated host effect is 1.8 percentage points; the Soviet effect and planned-economy effect are 3.4 and 1.0 percentage points respectively.

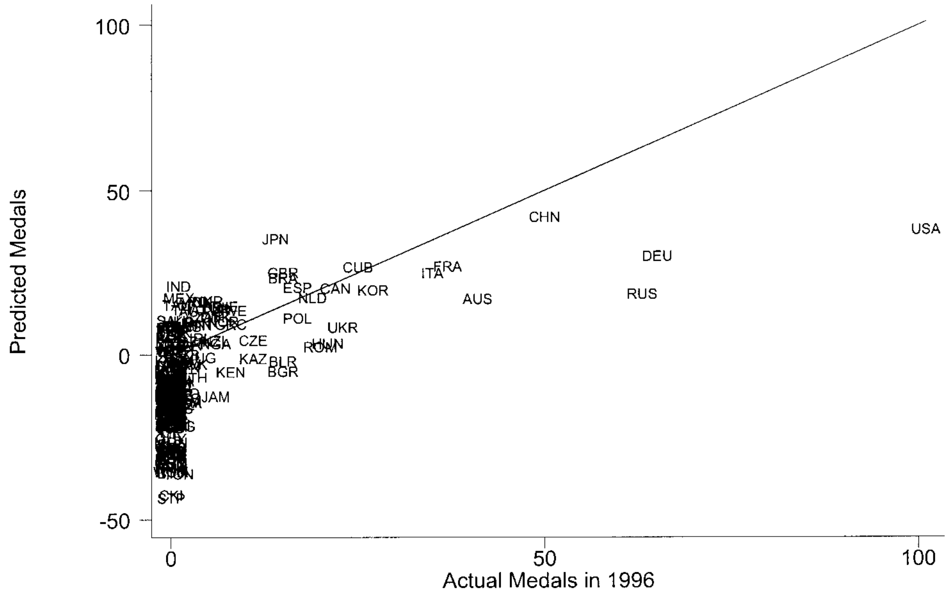
The statistical significance of GDP per capita indicates, perhaps not surprisingly, that economic resources are important in producing Olympic medalists. More surprising is the persistent similarity of the coefficients on log population and log GDP per capita. This suggests that it is a country's total GDP that matters in producing Olympic athletes. This in turn has the implication that two countries with the same GDP will win approximately the same number of medals, even if one is more populous with lower per capita income and the other is smaller with higher per capita GDP. Furthermore, this section has identified some important country characteristics that boost medal totals, including Soviet and host effects. Finally, there is strong evidence for durability to a country's Olympic investments. Past success is an indicator of current success; including lagged medal share further improves the fit of the model.

[6] The *Soviet* dummy includes Bulgaria, Czechoslovakia, Poland, the USSR, East Germany, Hungary, and Romania from 1960 to 1988; the Unified Team in 1992; and Cuba throughout the period. The *Planned* dummy includes China, Albania, Yugoslavia (through 1988), and North Korea.

[7] All previous results are robust to the omission of these boycotted Olympics.

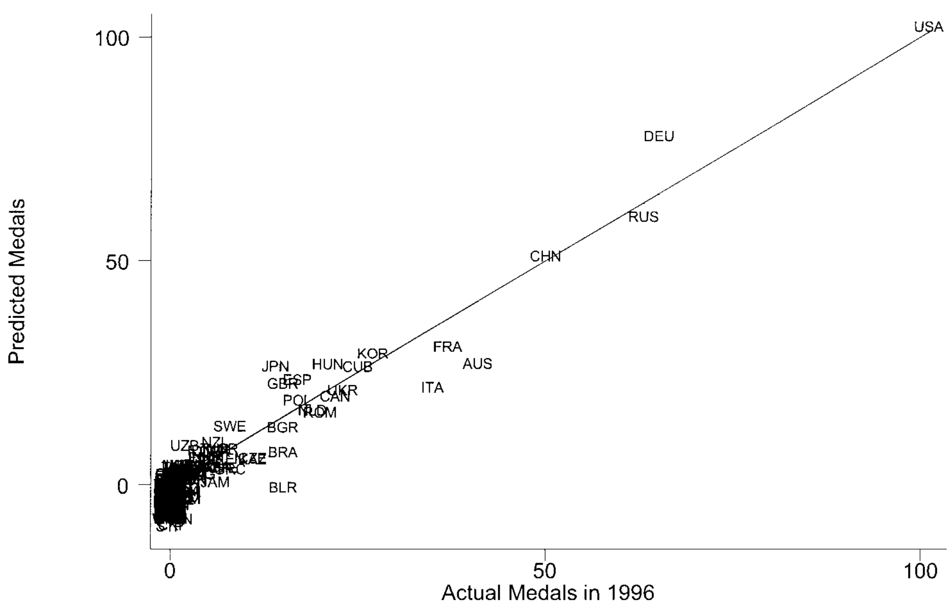
[8] Bernard and Busse (2000) present a talent investment model that generates equation (7).

FIGURE 1.—PREDICTED AND ACTUAL MEDAL TOTALS FOR 1996 FROM THE SPECIFICATION IN TABLE 2, COLUMN II



Points on the diagonal line represent perfect predictions.

FIGURE 2.—PREDICTED AND ACTUAL MEDAL TOTALS FOR 1996 FROM THE SPECIFICATION IN TABLE 2, COLUMN IV



Points on the diagonal line represent perfect predictions.

IV. Predicting Medals in Sydney

To provide a sterner test of our framework, we evaluate the out-of-sample performance. We made public predictions based on the model several weeks before the 2000 Sydney Games by estimating equation (7) on the 1996 cross section.[9] This cross-section specification cannot estimate a host effect, so we employ the coefficient on the host dummy for the same specification pooled over all nonboycott years (column IV of table 2). Figure 2 shows actual numbers of medals won in 1996 and the predicted numbers from this specification. The model does quite well in predicting totals for a number of countries, including the United States, Russia, and China. However, France, Italy, and Austria won more medals than predicted, and Germany won fewer than predicted.

For the 2000 games, we predicted total medal counts for the 36 countries that won at least five medals in 1996. Table 3 contains two sets of predictions for the Sydney Games as well as the actual medal totals for the 36 countries. Columns 1 and 2 represent the predictions and standard errors the model would have made if it had been

[9] See the *New York Times,* September 9, 2000. Johnson and Ali (2000) also predicted country medal totals for the Sydney Games, using different dependent and independent variables.

TABLE 3.—COUNTRY MEDAL PREDICTIONS FOR SYDNEY 2000

	Predicted 2000 Medals (1)	Standard Error Estimate	Predicted 2000 Medals (2)	Actual 2000 Medals
Australia	57	5.1	52	58
Belarus	12	5.0	12	17
Belgium	7	5.1	7	5
Brazil	18	5.1	17	12
Bulgaria	11	5.1	10	13
Canada	24	5.1	23	14
China	51	6.2	49	59
Cuba	21	7.1	20	29
Czech Republic	10	5.0	9	8
Denmark	7	5.1	7	6
France	39	5.1	38	38
Germany	66	5.5	63	57
Greece	8	5.1	8	13
Hungary	19	5.1	18	17
Italy	37	5.1	35	34
Jamaica	1	5.1	1	7
Japan	20	5.2	19	18
Kazakhstan	9	5.0	8	7
Kenya	5	5.1	5	7
Netherlands	20	5.1	19	23
New Zealand	5	5.1	5	4
Nigeria	5	5.1	5	3
North Korea	3	6.2	3	4
Norway	7	5.1	7	10
Poland	17	5.0	16	14
Romania	18	5.1	17	26
Russia	62	5.4	59	88
South Africa	6	5.1	6	5
South Korea	28	5.1	27	28
Spain	19	5.1	18	11
Sweden	9	5.1	9	12
Switzerland	8	5.1	8	9
Turkey	7	5.1	7	4
U.K.	18	5.1	18	28
Ukraine	22	5.9	21	23
U.S.	102	5.1	97	97

implemented without error.[10] Column 3 gives the ex ante predictions made public before the Sydney Games, and column 4 gives the ex post medal totals.

Both sets of predictions do quite well in matching the outcomes from the games. Although the error-free predictions hit only one total exactly, 9 countries are within one medal of their actual total, and 23 are within three medals. Using more formal distance metrics to measure our forecast performance, the predictions have an R^2 of 0.96, and 35 of the 36 countries were within 2 standard errors of the predicted values. The mean absolute error is 4.3 medals.

[10] The major mistake we made in implementing the model was ignoring double bronzes in some sports, which resulted in an incorrect total of 888 medals instead of the actual 929.

V. Conclusions

In this paper, we examine the question of how many Olympic medals a country should win. We begin with a simple hypothesis that medals should be proportional to population. We also consider a production function for Olympic medals that encompasses resources, population and other national characteristics.

Although the simple population hypothesis does have explanatory power, it fails to adequately explain the distribution of medals across countries. We find evidence that economic resources are important for producing Olympic athletes. Per capita income and population have identical effects at the margin, suggesting that total GDP is the best predictor of national Olympic performance. As a larger point, these results suggest that having resources to invest in human ability is important in producing success. This paper has emphasized athletic success, an endeavor in which natural talent probably plays a more important role than in more commonplace endeavors. That economic resources play a very significant role in determining success in this realm suggests that training resources are likely to be valuable for producing success in more mundane undertakings as well.

GDP is not the whole story in Olympic success. Host countries typically win an additional 1.8% of the medals beyond what would be predicted by their GDP alone. This host bounce led us to predict that Australia would win 17 extra medals in 2000, putting our medal prediction only 1 short of the actual total. The forced mobilization of resources by governments can also play a role in medal totals. On average, the Soviet Union and Eastern Bloc countries had medal shares more than 3 percentage points higher than that predicted by their GDP and past performance during the 1960–1996 period.

We finished by exposing our simple specification to an out-of-sample test by predicting medal totals for the 2000 summer games in Sydney. The model does quite well by most statistical metrics.

REFERENCES

Ball, Donald W., “Olympic Games Competition: Structural Correlates of National Success,” *International Journal of Comparative Sociology* 15 (1972), 186–200.

Bernard, Andrew B., and Meghan R. Busse, “Who Wins the Olympic Games: Economic Development and Medal Totals,” National Bureau of Economic Research working paper no. 7998 (2000).

Busse, Meghan R., and Andrew B. Bernard, “Consistent Standard Errors in Panel Tobit with Autocorrelation,” Tuck School of Business mimeo (2002).

Grimes, A. Ray, William J. Kelly, and Paul H. Rubin, “A Socioeconomic Model of National Olympic Performance,” *Social Science Quarterly* 55 (1974), 777–782.

Johnson, Daniel K. N., and Ayfer Ali, “Coming to Play or Coming to Win: Participation and Success at the Olympic Games,” Wellesley College mimeo (2000).

Levine, Ned, “Why Do Countries Win Olympic Medals? Some Structural Correlates of Olympic Games Success: 1972,” *Sociology and Social Research* 58 (1974), 353–360.

Shughart, William F., and Robert D. Tollison, “Going for the Gold: Property Rights and Athletic Effort in Transitional Economies,” *Kyklos* 46:2 (1993), 263–272.

Wallechinsky, David, *The Complete Book of the Olympics, 1992 Edition* (Boston: Little, Brown, 1992).