Working Paper Series

WP 14-7 OCTOBER 2014

What Goes into a Medal: Women's Inclusion and Success at the Olympic Games

Marcus Noland and Kevin Stahler

Abstract

This paper examines determinants of women's participation and performance in the Olympics. Female inclusion and success are not merely functions of size, wealth, and host advantage, but a more complex process involving the socio-economic status of women and, more weakly, broad societal attitudes on gender issues. Female labor force participation and educational attainment in particular are tightly correlated with both participation and outcomes, even controlling for per capita income. Female educational attainment is strongly correlated with both the breadth of participation across sporting events and success in those events. Host countries and socialist states also are associated with unusually high levels of participation and medaling by female athletes. Medal performance is affected by large-scale boycotts. Opening competition to professionals may have leveled the playing field for poorer countries. But the historical record for women's medal achievement is utterly distorted by the doping program in the former East Germany, which specifically targeted women. At its peak in the 1970s and 1980s, the program was responsible for 17 percent of the medals awarded to women, equivalent to the medal hauls of the Soviet or American team in 1972, the last Olympics not marred by widespread abuse of performance-enhancing drugs.

JEL Codes: J16, L83, F53, Z13

Keywords: women, gender, sports, Olympics

Marcus Noland, executive vice president and director of studies at the Peterson Institute for International Economics, has been associated with the Institute since 1985. He is also senior fellow at the Institute and at the East-West Center. **Kevin Stahler** has been a research analyst at the Peterson Institute since June 2013.

Note: EY provided financial support for this research. We thank the International Olympic Committee for its provision of data used in this paper; seminar participants at the Peterson Institute, particularly Arvind Subramanian; and Andrew Bernard for comments on an earlier draft. The views expressed herein and any errors or shortcomings are solely the responsibility of the authors.

Peterson Institute for International Economics
1750 Massachusetts Avenue, NW Washington, DC 20036-1903
Tel: (202) 328-9000 Fax: (202) 659-3225 www.piie.com

The Olympics are the largest regularly scheduled international mass gathering, with hundreds of thousands of onsite participants including athletes, team officials, press, and spectators, and billions more following via the media. The Olympic Games are among the few truly global events and as such wield an enormous ideational influence on popular culture and public attitudes. One of the striking aspects of the modern Olympics is the rise in female participation (figures 1 and 2). This expansion is partly a by-product of the overall growth of the Games due to the inclusion of new events and the increase of participating National Olympic Committees (NOCs) over time. But this is not the whole story: Since the exclusion of female participants in the 1896 Athens Games, the share of female participants has risen steadily, and women now make up nearly half the competitors in both the Summer (figure 1) and Winter Games (figure 2).[1]

Previous researchers have empirically investigated the determinants of success at the Olympic Games, modeling medal counts primarily as a function of country size and income, host country advantage, and select socioeconomic indicators (e.g., Bernard and Busse 2004; Klein 2004; Johnson and Ali 2004; Lui and Suen 2008; Andreff 2013; and Lowen, Deaner, and Schmitt 2014). Some of this work has looked at female athletic outcomes specifically. For example, Michael Klein (2004) finds that a higher labor force participation gender ratio led a country to win more medals in women's events at the 2000 Sydney Games, and Aaron Lowen and colleagues (2014) conclude that female athletic participation and success at the Summer Games between 1996 and 2012 are in part determined by a nation's score on the UN's Gender Inequality Index.

This paper expands on existing literature by modeling female participation and medal counts at the Summer Games using a broad range of available controls including educational attainment, labor force outcomes, and societal attitudes. As a departure from some previous work, we employ a wide sampling frame of available observations between 1960 and 2012, as well as using statistical techniques to account for an Olympic delegation's path dependency, or legacy effect. Perhaps not surprisingly we find that female inclusion and success are not merely functions of size, wealth, and host advantage, but a more complex process involving the socioeconomic status of women and, more weakly, broad societal attitudes on gender issues. Female labor force participation and educational attainment in particular are tightly correlated with participation and outcomes, even controlling for per capita income. Host countries and socialist states also are associated with unusually high levels of participation and medaling by female athletes.

1. The Summer Games are officially known as the "Games of the *n*th Olympiad," and the Winter Games are the "*n*th Winter Olympics." For expository convenience we will follow the colloquial practice of referring to them as the Summer Olympics and Winter Olympics.

We then extend the analysis in four ways, examining the impact of boycotts, doping, the opening of competition to professional athletes, and societal attitudes. Like previous researchers, we find that medal performance is affected by large-scale boycotts, as occurred at the 1980 Moscow and 1984 Los Angeles Games. We also find that the opening of the competitions to openly professional athletes had some impact, in effect helping athletes from poorer countries. But the historical record for women's medal achievement is completely distorted by the East German program of systematic administration of performance-enhancing drugs (PEDs) that was applied particularly intensely to female athletes. At its peak in the 1970s and 1980s, we estimate that the East German doping program was responsible for 17 percent of the medals awarded to female athletes, equivalent to the total female medal share that the Soviet and American teams each earned separately in 1972, the last year the Summer Games were not marred by widespread doping.

PARTICIPATION AND PERFORMANCE

With Saudi Arabia fielding a mixed men and women's delegation at the 2012 London Games, mixed national delegations have effectively become universal for the Summer Games (table 1). (The percentage of national teams including women remains less than 100 percent due to some very small national delegations that have only one or two participants.) We begin our analysis with correlates of inclusion of women athletes and predictors of the share of women competitors. Previous researchers, notably Andrew B. Bernard and Meghan R. Busse (2004), have focused on country size and income level as the main predictors of Olympic success. Specifically, they posit a model in which countries produce Olympic caliber athletes using people, money, and some organizational capacity using Cobb-Douglas technology,

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

where T is talent, N is population, Y is national income, A is organizational capacity, and the subscripts i and t refer to country and year, respectively. A country's share of Olympic medals is a function of talent, and a log function translation of talent into medal shares:

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

The following yields a specification for medal shares:

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

Because national income can be expressed as the product of population and per capita income, the previous condition can be restated as

$$M_{it} = \begin{cases} C + \alpha \ln N_{it} + \beta \ln (Y/N)_{it} + d_i + v_i + \epsilon_{it} & if\ M_{it}^* \geq 0 \\ 0 & if\ M_{it}^* < 0, \end{cases}$$

yielding an estimatable model

$$M_{it} = C + \alpha \ln N_{it} + \beta \ln \left(\frac{Y}{N}\right)_{it} + Host_{it} + Soviet_{it} + Planned_{it} + d_i + v_i + \epsilon_{it}$$

where the NOC share of total medals won is a function of log population and GDP per capita, together with dummy controls for Olympic host countries and whether it was a Soviet or planned economy.

Below we demonstrate that the simple income-based model does not adequately capture the more complex process of generating female Olympic athletes. Table 2 replicates a simplified version of the Bernard and Busse model using our data. The resulting coefficients on population and GDP per capita are very similar in model 2.1 to those obtained by Bernard and Busse despite our wider sampling frame and use of GDP in purchasing power parity (PPP) terms. However, the estimated coefficients diverge when we separate the dependent medal share variable into male-specific and female-specific events, as well as adding an additional control for average years of schooling in the total population. The results provide initial evidence for our contention: In both models GDP per capita appears to be highly collinear with proxies for education levels and, most likely, other related social indicators. Additionally, determinants of success differ noticeably: For female medaling, income matters little, but the impact of education is more than twice as powerful as compared to that of male athletes.

Our initial specification search for additional determinants of female-specific inclusion and success at the Summer Games revealed a much larger pool of potential correlates. Table 3 shows pair-wise correlations for nearly a dozen country-specific measures against the four dependent variables (female participation share, female medal share, NOC participant gender share, and NOC medal gender share) used throughout this study. Variable definitions, data sources, and relevant notes on data limitations are reported in appendix A.

Table 4 reports multivariate regressions on female participation and performance for the full sample period (1960–2012), and for reasons of data availability, a "modern" subsample period (1996–2012). The dependent variables are defined in binary terms: Did the NOC delegation contain a female athlete? Did a female athlete medal? Did a female athlete earn a gold medal? Year dummies are included, but there are no country fixed effects. The reported standard errors are heteroskedasticity-robust.

Female participation is strongly positively correlated with the country size (population), average years of female schooling, and whether the country was a member of the Communist bloc.[2] Participation negatively correlated with the Muslim population share. For reasons of data availability, the ratio of female to male labor force participation is included only in the modern subsample, and in this regression it is positively and significantly correlated with female medaling but not with female inclusion. As among the few regressors with generally complete geographical and temporal coverage, adolescent fertility and the urbanization rate were included as proxies for women's health status and general living environment, but they are not robustly correlated with the dependent variables. Interestingly, NOCs from small states[3] suffer an additional handicap in sending female athletes and winning medals throughout the full period, but this effect may have dissipated in more recent games.[4]

Country population, average years of female schooling, and membership in the Communist bloc are also positively correlated with medaling. And like the participation regression, female labor force participation is correlated with medaling in the modern subsample. However, unlike the participation regression, the level of per capita income is more statistically and substantively significant. And while the Muslim population share is correlated with participation, it is uncorrelated with performance, once women are allowed to compete.[5]

So far women's participation and performance has been treated as a binary outcome in which each Olympic Game is a separate event with the past not influencing the future. However, in tables 5 through 8 the analysis is treated differently in a number of ways. In table 5, the dependent variable is defined as the NOC's share of female participants in the Games; in table 6 it is defined as the share of female

2. The Communist bloc dummy is not included in specification 4.5 because every Communist country in the subsample (China, Cuba, Laos, and Vietnam) sent female athletes. Similarly, the host country dummy is not included because it is a perfect predictor of success in all estimations.

3. The Commonwealth Secretariat/World Bank (2000) defines 45 countries as "small states," under the general rule that the population is less than 1.5 million. In our data, this variable is treated as a fixed effect.

4. In specifications 4.3 and 4.6 (table 4), the small state dummy was intentionally left out of the regression because the variable perfectly predicted zero female gold medals won. In fact, two countries designated as small states have won female gold medals during the sampling period (Bahamas in 2000 and 2004, Estonia in 1992), but are not included in the regression because of one or more missing independent variable values.

5. The Muslim population share was largely insignificant in subsequent applications and is not included in the regressions reported in the remaining tables.

participants within each NOC delegation. In table 7, the dependent variable is defined as the NOC's share of female medal winners in the Games, and in table 8, it is defined as the share of medals specific to female events out of all medals won at the Summer Games within each NOC delegation.

Second, we include lagged dependent variables to take into account the apparent path dependence of these outcomes. The regressions in table 5 are estimated four ways: as simple pooled ordinary least squares (OLS) and three variants of the generalized methods of moments (GMM) estimation. As explained pedagogically by David Roodman (2009), GMM estimators accommodate dynamic linear models in which outcomes depend on their past realizations, independent variables that may not be strictly exogenous, and individual-level controls in "small T-large N" panel data. For our models, we see little concern for endogeneity in our vector of independent controls and treat all variables except the lagged dependent variable as strictly exogenous. One-step system GMM with orthogonal deviations[6] and robust standard errors is employed for all models, and in most cases we calibrate for instrument proliferation by limiting or collapsing all available dependent variable lags of two and greater.[7]

As seen in table 5, regardless of the estimator used, much like table 3, female participation is strongly correlated with population across all models, though the magnitude of the coefficients vary, sometimes considerably, depending on the estimation technique. Average female years of schooling is also positive and significant in all models except the collapsed system GMM iteration where it falls just shy of the 10 percent p-value cutoff. Being the host of the Olympic Games boosts female participation significantly, but that effect goes into reverse in subsequent Games.[8] Being a member of the Communist bloc matters for the full sample but not the modern subsample, perhaps because the membership changes from being mostly European Soviet bloc in the full sample to almost entirely Asian in the recent subsample. The coefficients on the first lagged dependent variables, ranging from 0.4 to 0.6, indicate that past outcomes have a nontrivial effect, though perhaps not an overwhelming one. If the collapsed instrument system results (specifications 4.4 and 4.6) are considered preferred, the robust correlates are population and status as the current host. Average years of female schooling and the labor force gender ratio also matter in the unrestricted system GMM models (5.2 and 5.5), though instrument proliferation, which may over fit endogenous variables, could be an issue.

6. In the case of unbalanced panels with many gaps such as our own, orthogonal deviations subtract the average of all future available observations of a variable to minimize data loss. See Roodman (2009) for details.

7. Only specifications 5.2 through 5.4 use all available dependent variable lags of three and greater, because using shallower lags led to potentially hazardous issues of AR(2) autocorrelation in first differences.

8. The host dummies were not included in the table 3 regressions because every host included female athletes in its NOC delegation. Daniel K. N. Johnson and Ayfer Ali (2004) and Wladimir Andreff (2013) find, in addition to the host effect, broader neighborhood or regional effects; we did not.

Table 6 looks at the gender makeup of individual Olympic delegations; theoretically, this measure should highlight how much a country values the potential of its female athletes relative to males, regardless of a delegation's absolute size. Interestingly, per capita income is completely insignificant across the board. Indeed many of the delegations with the highest share of female athletes—or entirely female delegations in a few cases—are smaller NOCs from less wealthy countries. Instead, population and the lagged dependent variable are the most robust correlates. Female labor force participation is strongly correlated with the share of females included in NOCs in the modern subsample where this data are available. Female schooling, status as the current host, and membership in the Communist bloc are significant in the full sample but not the modern subsample.

In tables 7 and 8 we move from participation to performance. Here the issue of censoring becomes more acute. In our previous uncensored regressions on participation, approximately 20 percent of the observations reported female participation shares of zero. However in terms of medaling, closer to 60 percent of NOCs that sent female athletes reported zero female medals won. This pronounced clumping at the zero lower bound motivates the use of tobit models.

Table 7 reports estimates of female medals won by an NOC as a share of total female-event medals available that year among the countries that sent female participants.[9] Female schooling, female labor force participation (in the subsample where data are available), population, and per capita income are positively correlated with medaling. Membership in the Communist bloc is a statistically significant correlate in the full sample. Status as a current host is significant in all specifications. The inclusion of lagged dependent variables affects the magnitude of the coefficients on the other included regressors, but generally not their level of statistical significance.

In table 8 the dependent variable is defined as the female share of medals out of total medals won by the NOC delegation, excluding delegations that won zero total medals (which reduces sample size and dispersion).[10] Like the gender makeup of individual NOC participants, the gender makeup of medals is not determined by factors of per capita income. Female schooling is positively correlated with the female medal share regardless of the estimation technique for the full sample, but it drops out in the modern subsample. Conversely, female labor force participation is positive and significant in the recent subsample where the data are available. As shown in the other tables, being a member of the Communist bloc has

9. Throughout the sampling period 1960–2012, 20.6 percent of participating NOCs did not send female athletes. In all female medaling regressions, these observations were not included, as they would confound the effects of those countries that sent female competitors but did not medal.

10. We must exclude NOCs that did not win medals in any events for the purpose of this model. Indeed, we would not want to confuse the comparison between a country that won multiple total medals and zero female medals with a country that won medals in neither male, female, or mixed events. In addition to limiting the size, this also by definition skews the sample toward successful countries.

a pronounced effect on female medals relative to total medals in the full sample. Past studies have found positive Communist bloc effects on total medals won overall, which makes it hard to say to what extent women were affected differently than men. However, this finding shows a clear gender-specific difference for centrally planned economies: Communist countries large and small have sent delegations with higher proportions of female athletes (table 6) who in turn win higher proportions of medals (table 8).

Finally, in table 9, we use sport-specific data to construct Herfindahl indices to examine the breadth of female participation and medaling. A commonly cited measure in industrial organization literature to judge industry concentration, the Herfindahl index is calculated by summing the squared market share values of all intra-industry firms.[11] In this application, we assign Herfindahl index values to each NOC by summing the squared share of the country's total participants/medals in each sporting category. For example, in the 2012 London Summer Games, females could compete in 26 sports. Therefore, on one extreme end, a country that allocates its share of female participants/medal winnings equally among all sports would receive a Herfindahl index value of $\sum_1^{26}\left(\frac{1}{26}\right)^2 = 0.038$. On the other end, a country that places all female participants/medals into a single sport would receive a value of 1.

Table 9 shows the top and bottom 15 NOCs ranked by the Herfindahl index of female sport–specific participation, lowest (least concentrated) to highest (highly concentrated), at London 2012. Only NOCs that won at least one medal in a female event are considered in order to create an adequate set of comparison countries and root out extremely small NOCs. France, South Korea, and Japan appear to be least concentrated in terms of female participation—though not completely evenly distributed—with other top ten countries very close on their tail. Indeed, these top three countries entered female participants in 18 to 21 of 26 sporting categories, with no pronounced clumping in any single sport. Moreover, these countries do not appear to concentrate their female athletes much more or less than its male athletes (see the "Herfindahl ratio" column) except for Azerbaijan, which concentrated its male talent in wrestling events, but evenly distributed female athletes among a dozen sports.

On the opposite end of the spectrum, Kenya, Ethiopia, and Jamaica are all highly concentrated in terms of their female participation and medaling. In fact, each of these countries placed more than 90 percent of their female athletes in "athletics" sporting events (as they did with male athletes). The perfectly concentrated 1 on the Herfindahl index for medaling shows that these countries won medals only in a single sport (here, "athletics").

The fourth column of table 9 reports the Herfindahl index for medaling. Not surprisingly, breadth of medaling is correlated with breadth of participation (the simple correlation coefficient is 0.55,

11. For example, in the case of a single monopolistic producer, the Herfindahl index would be 1 (100 percent ^ 2) indicating complete concentration (and therefore no competition); on the opposite end of the spectrum, in an industry where many small firms compete with similar market share (e.g., 1 percent ^ 2 + 1 percent ^2 …), the Herfindahl index would be near the zero lower bound.

significant at the 1 percent level)—one cannot medal without participating. But the correlation is less than perfect and may be affected by competitive strategic choices: Subject to resource constraints, NOCs may concentrate their delegations in events where the country has a perceived comparative advantage.

Table 10 reports multivariate modeling of these indices. The sample period 2000–2012 was chosen because the types of events open to female competition are comparatively diverse.[12] As opposed to our previous models, the "legacy effect" does not appear to be a major issue, but there tends to be some mild clumping at the dependent variable's higher upper bound of 1.[13] Therefore, we employ tobits and random effects tobits, and include lagged dependent variables only in regressions 10.2 and 10.4. We should not invite too much comparison between regression results in 10.1–10.3 and 10.4–10.6, since by methodological design they are estimating different sampling frames.[14]

Breadth of participation and medaling is robustly associated with country size, per capita income, and female educational attainment. In particular, in specification 10.6, one extra year of female schooling would appear to lead to a roughly 12 percent absolute decrease in the Herfindahl medaling index, which is rather large. However, breadth of participation or medaling is not associated with host country status. This result may be at least partly explained by our sample excluding the 1980 Moscow and 1984 Los Angeles Games, which were marred by large-scale boycotts that conferred undue advantage on the hosts. There is only weak evidence that female labor force participation or status as a Communist country affects the breadth of participation or medaling.

To sum up, table 11 reports the substantive effect of a one standard deviation (SD) shock for selected variables, and the effect of being a host or a Communist country, for the entire sample period of 1960–2012, except in the case of the Herfindahl index regressions that use the 2000–2012 period. In addition to reporting the relative increase in shares, table 11 demonstrates the absolute effect of these shocks at the 2012 London Games. Again, we need to be careful not to draw direct comparisons from these results, since sampling frames can sometimes differ widely depending on the estimation. Nevertheless, being the host of the Games has by far the largest positive effect on female participation share at 3.3 percent, or about 155 extra British female athletes at London. Indeed, this is not too far off from the 120 or so extra women the United Kingdom sent in comparison to the 2008 Games. Being a member of the Communist bloc would imply a relative increase in participation of 0.7 percent, or

12. Between 1996 and 2000, there was a large jump in the number of events open to women competitors (from 97 to 120). In the 2000 to 2012 period, the number of women and mixed events expanded by roughly 17 percent.

13. For example, the correlation coefficient of female medal and participation shares on their t-1 lagged value throughout this period is over 90 percent. However, the correlation coefficient of our Herfindahl index variables on their t-1 lagged values is approximately 50 percent.

14. For the participation Herfindahl index (10.1–10.3), we exclude countries that did not send female athletes since the squared share of zero is meaningless. For similar reasons, we exclude NOCs that did not win a medal in female events in regressions 10.4–10.6.

33 women. Additionally, although female education and the labor force ratio are just outside of the acceptable statistical cutoffs, a one-standard-deviation positive shock would imply modest gains.

It is in the other models that female schooling and labor force participation shine more brightly. A one-standard-deviation shock of 3.2 years of average schooling for women would lead to a 3 percent higher share of women in individual Olympic delegations, a 7 percent rise in the share of female medals relative to total delegation medals won, and an extra eight total female medals won at the 2012 Olympics. Significantly, the same shock to education leads to a 0.11 drop in the Herfindahl index for participation and a 0.375 drop for medaling, implying more educated NOCs have much more diversified female delegations, controlling for the level of per capita income and other influences. A positive shock to relative female participation in the labor force implies gains of similar magnitude, save medals won, where the effect is more muted.[15]

Clearly, female athletic inclusion and success at the Summer Olympics are due to more than a country's size and per capita income. In line with past studies, we find that some of the basics certainly do matter: Larger countries, Olympic hosts, and members of the Communist bloc tend to send more athletes and claim a larger share of the glory. Like Klein (2004) and Lowen and colleagues (2014) we find that indicators of socioeconomic gender equality, such as the labor force participation gender ratio, also matter. Notably, gains in female educational attainment evidence a modest, but consistently significant, contributor to athletic inclusion and success. This adds subtlety to our understanding of the Olympic movement: The story of female athleticism at the Games is intrinsically tied to that of a woman's status in society.

We next turn to some extensions of the basic model, examining the role of boycotts, doping, amateurism, and the possibility that societal attitudes have a direct impact on participation and performance.

BOYCOTTS

How one places in competition depends in part on who one competes against, and the Olympics have a long history of politically related exclusions and boycotts. There were threatened boycotts at the 1936 Berlin and 1956 Melbourne Games, but the first actual boycotts by NOCs came at the 1964 Tokyo Games when the International Olympic Committee (IOC) banned South Africa in response to the country's apartheid policies of racial segregation, and North Korea and Indonesia boycotted the Games for their own reasons. The first significant mass boycott occurred in 1976 when, in response to the IOC's

15. The impact of female labor force participation appears to be more modest than that found by Klein (2004), but this is explained by the inclusion of educational attainment in the models that soaks up some of the influence otherwise attributed to labor force participation.

refusal to ban New Zealand, which had carried on rugby matches with South Africa, 28 African NOCs withdrew from the Montreal Games, presumably affecting the results in distance running in particular. However, the biggest boycotts involved the 1980 Moscow Games, when 62 NOCs stayed away in response to the Soviet Union's invasion of Afghanistan, and the following Games, when in retaliation the Soviet Union and 16 of its allies refused to compete in Los Angeles (Senn 1999 and Guttmann 2002).

To get a sense of the impact of these boycotts, we examine the results from the 1980 and 1984 Games. To calibrate an NOC's underlying competitiveness, we use its resulting female medal shares from the 1988 Seoul Games. Table 12 reports three regressions: 12.1 is a tobit of the NOC's female medal share (equivalent to 7.1) estimated for the 1980 and 1984 competitions, on the sample of countries that participated and sent female athletes in 1980 or 1984 and participated in1988. The NOC's female medal share in 1988 is introduced in specifications 12.2 and 12.3. (Theoretically, it might be preferable to use the outcomes from the 1976 Montreal Games as the competitiveness proxy, since unlike the 1988 results, the 1976 results would not be affected by outcomes in 1980 and 1984, but the 1976 performances—and the availability of data—are affected by the African boycott). It is quite apparent from 12.2 and 12.3 that the 1980 and 1984 boycotts boosted the medal counts of the participants, particularly for the host countries. All things being equal, a 1 percent increase in the medal share in the "normal" 1988 Olympic Games would generate a 1.67 percent increase in the boycotted Games of 1980 and 1984.

DOPING

The first recorded use of a performance-enhancing drug was at the 1904 St. Louis Games when American marathoner Thomas Hicks consumed a concoction including strychnine before his race. By the 1950s, testosterone was used widely by weightlifters. Ampoules and used syringes were discovered at the 1952 Oslo Games during which several skaters became ill after excessive consumption of amphetamines. But a focusing event occurred at the 1960 Rome Games when Danish cyclist Knut Jensen died while suffering a heatstroke after effectively overdosing on Ronicol—a peripheral vasodilator known to enhance blood circulation—which had been administered to him by a team trainer. Since Jensen's death the Olympic movement has struggled with the medical, legal, and organizational complexities of dealing with doping.[16]

There are four ways to detect the use of PEDs: investigative journalism, government investigation, drug testing, and surveys; none are foolproof (Yesalis, Kopstein, and Bahrke 2001). Testing, in particular, may not be an accurate indicator of actual usage due to the strong incentive to avoid detection by athletes, coaches, and team officials, as well as the problematic history of testing in the Olympic context (Hunt

16. In 1963 the IOC defined doping as "an illegal procedure used by certain athletes in the form of drugs; physical means and exceptional measures which are used by small groups in a sporting community in order to alter positively or negatively the physical or psychological capacity of a living creature, man or animal in competitive sport" (Hunt 2011, 15).

2011). From a standpoint of estimation, it would appear that modeling doping is effectively impossible, with one important exception.

Beginning in the 1950s, East German sports authorities began exploring the potential impact of PEDs, building off scientific knowledge and practices developed during the Nazi period. Eventually, the authorities came to believe that the program had potentially significant effects and would need to be kept secret due to the nascent attempts to police doping by international sports bodies. By the 1960s, the doping program came under the direct control of the Stasi, or secret police. In 1974 East German sports authorities initiated a systematic program of administering PEDs to East German athletes in all sports except sailing and women's gymnastics. This effort, involving more than 10,000 athletes, has been documented from files that became available after the dissolution of East Germany, and the authenticity of these documents has been upheld in subsequent court cases (Franke and Berendonk 1997; Yesalis, Kopstein, and Bahrke 2001; Hunt 2011; and Ungerleider 2013).

The program began in earnest with female competitors prior to the 1968 Mexico City Games. Drugs were used to suppress menstruation, and then testosterone and other PEDs were given to the women, sometimes without their knowledge. From 1972 on, most East German medals were won by athletes who were on PEDs, including most gold medal winners in swimming events from the 1976 Montreal Games on, and all gold medals in throwing events at the 1988 Seoul Games. As Thomas Hunt (2011) observes, East German athletes medaled at a rate 15 times that of the United States on a per capita basis. The distortive effect of this program appears to have been especially large in women's competitions: As Werner W. Franke and Brigitte Berendonk (1997, 1262) write, "Special emphasis was placed on administering androgens to women and adolescent girls because this practice proved to be particularly effective for sports performance." The East German share of women's medals rose from 7 percent in Mexico City (1968) to 33 percent in Montreal (1976) to 39 percent in the boycott-marred Moscow Games (1980) (table 13). The latter two performances are the highest shares ever recorded, topping the shares achieved by those of either the Soviet Union or the United States in any other post-WWII Games. The disproportionate impact the East German doping program had on the women's competitions compared to the men's is quite evident in the far greater impact felt in the female medal counts.

It is effectively impossible to model comprehensively the impact of doping on outcomes—by its very nature the practice is concealed. But as a start, table 10 reports regressions on an NOC's share of female medals, with dummy variables added for East Germany (to capture any possible unique East German prowess, as well as the initial PED period, East Germany post-1972 once doping became ubiquitous, and the boycott years 1980 and 1984).[17] As can be seen in table 14, East Germany outperformed from

17. Out of necessity these regressions were estimated without a lagged dependent variable. East Germany competed as an independent entity only five times: 1968, 1972, 1976, 1980, and 1988; if a lagged dependent variable had been included, 1968 and 1988 (two of the five observations) would have been lost.

the start, either due to unique prowess or the early doping program, but the systematic doping campaign had an enormous impact, increasing the East German female medal share by more than 17 percent of the total won by women. To put some perspective on this, in 1972 the Soviet Union and the United States tied for 19th best female athletic performance of any NOC in a post-WWII Olympics, each taking home 17 percent of the medals. In other words, in the last Olympics before the era of large-scale doping, the East German doping effort had an impact on its female athletes' program roughly equivalent to the accomplishments of the entire Soviet or US female contingent.

AMATEURISM

Many of the founders of the modern Olympic movement subscribed to notions of amateurism derived from the 19th century British view of sport as the proper pastime of upper-class gentlemen, and for most of its history the Olympic movement and constituent athletic federations have struggled with the issue of professionalization (Senn 1999 and Guttmann 2002). At times enforcement of amateur rules was extremely rigorous, notably in the IOC's decision to strip American track athlete Jim Thorpe of his medals, ex post over his acceptance of money to play minor league baseball, and to ban Finland's Paavo Nurmi, then the world's premier distance runner, from the 1932 Los Angeles Games. But the clear if uneven trend was toward greater acceptance of financial reward for athletes, whether in the form of direct payments or indirectly through product endorsements.[18] The acceptance of professionalism gained momentum with the emergence of participation by Soviet and other Eastern bloc NOCs, which operated systems that made their competitors professionals in all but name only.

Tolerance of professionalism varied across sports federations, but by consensus the 1992 Barcelona Games were the first "professional" Games, particularly notable for the participation of the US "Dream Team" of professional men's basketball players from the National Basketball Association. Given the generally greater prominence of men's professional sports around the world, it is possible that this shift would have a bigger impact on men's results. Table 15 examines if the pattern of correlates of women's success shifts after the opening of competition to professional athletes. The correlation could go either way. One possibility is that if professionalism is tolerated openly, the impact of per capita income may be reduced since athletes from poor countries (who might face the biggest incentives to give up their amateur status) could openly turn professional. The other possibility is that sports require all sorts of training facilities, specialized coaching, and medical care, and if professionalism is permitted openly, then professionals from high-income countries are more likely to be able to access these performance-supporting inputs due to their far larger and more lucrative venues for professional competition.

18. When in 1968 IOC head Avery Brundage complained to French IOC member Marceau Crespin that half the French ski team failed to live up to the amateur rule, Crespin responded: "You have been misinformed, Monsieur. No one on the French ski team lives up to your definition" (Guttmann 2002, 128).

In specifications 15.1 and 15.2 we use tobit models to split the sample between the era of amateurism (1960–88) and that of professionalism (1992–2012) and estimate separately. Interestingly, the coefficient on logged GDP per capita drops considerably to 0.244 in the 1992–2012 sample, lending some evidence that wealth matters less. In specifications 15.3 and 15.4, however, we attempt a slightly more sophisticated operation in which we add a post-1990 interaction effect to the full sample, which records zero values until 1991, and the standard log GDP per capita values afterward. Here, however, we find no evidence for a statistically significant slope change in either direction from the 1992 Games onward.

If anything, the results reported in table 15 appear to support the former hypothesis: The shift to professionalism could have leveled the playing field for athletes from poorer countries who now have greater incentive to go pro. But the changes in results are not particularly dramatic or robust. The most likely explanation is that the pattern of change was more gradual than a simple regime change in Barcelona and may well not have affected results for women as much as those for men. It is also the case that Barcelona was the first competition following the demise of the Eastern bloc, and this also likely had an impact on the results (Bernard and Busse 2004).

SOCIETAL ATTITUDES

A final issue is whether societal attitudes regarding gender issues affect women's participation and performance in the Olympics. Intuitively, this notion is nearly self-evident, yet it is actually difficult to demonstrate rigorously. The Pew Research Center's Global Attitudes Project is one source of cross-national public opinion. In table 16, the percentage of the population agreeing that a marriage where both spouses work is more satisfying is added to core specifications on female participation and performance (specifications 16.1–16.4), while in specifications 16.5–16.8 the percent agreeing that university education is more important for a boy than a girl is included.[19] Unlike the previous regressions where it is difficult to argue that the Olympics affect income, labor force participation, or educational attainment at the national level, one could argue that the Olympics might possibly have ideational effects, such that these attitudinal measures might not be predetermined.

The Pew survey country samples are quite limited, severely constraining the sample size, and responses to the two questions on both spouses working and university education are collinear, generating

19. This table uses results from a 2010 Pew poll of 22 countries. The same results are assigned to country observations in both 2008 and 2012 under the assumption that opinions had not markedly changed within this four-year window. In addition to these questions, we ran results on the percent of respondents agreeing that women should have equal rights with men, women should be able to work outside the home, women have a better life than men, and men get more opportunities for jobs that pay well. These, however, were not as statistically robust as the chosen questions. The Pew report is available at http://www.pewglobal.org/2010/07/01/gender-equality/.

insignificant statistical results if added to the regressions jointly. Yet the regressors generally yielded estimates with coefficients statistically and significantly different from zero. A 1 percent increase in the population share believing in both spouses working is associated with a 0.25 percent increase in the share of total female athletes at the Games (equivalent to about 11 extra female athletes at London 2012) and a 0.7 percent increase in a delegation's female medals relative to total medals won. Conversely, for every additional 1 percent of country respondents believing higher education is more important for males, a delegation's female medals won relative to total medals drops about the same absolute value (–0.65 percent).

A final issue in this context is the impact of legislative or regulatory attempts to promote female sports, most notably the Title IX legislation adopted in the United States in 1972, which mandated gender equity in all educational programs receiving federal government support and resulted in a significant expansion in school athletic opportunities for girls. It was not possible to construct a cross-country panel of Title IX type measures, and attempts to model its impact on the United States were disappointing. Neither inclusion of a dummy variable for the United States starting in 1972, a lagged effect dummy, or a variable that cumulated in value starting in 1972 yielded robust results. It is surely the case that Title IX has had an impact on female sports participation and success in the United States, but it is difficult to demonstrate this proposition in the framework at hand.[20]

CONCLUSION

Previous research on performance at the Olympic Games has emphasized the roles of country size and wealth, host effects, political determinants, and has made headway in uncovering the more subtle socio-economic determinants of success at the Olympic games. In this paper we examine the determinants of women's participation and performance in the Olympics and perhaps not surprisingly find that these outcomes are generated by a complex process involving the socioeconomic status of women, and more weakly, societal attitudes on gender issues more broadly. Female labor force participation and educational attainment in particular are tightly correlated with participation and outcomes, even controlling for the level of per capita income. Female educational attainment is strongly correlated with the breadth of participation across sporting events and success in those events. Host countries and socialist states also are associated with unusually high levels of participation and medaling by female athletes.

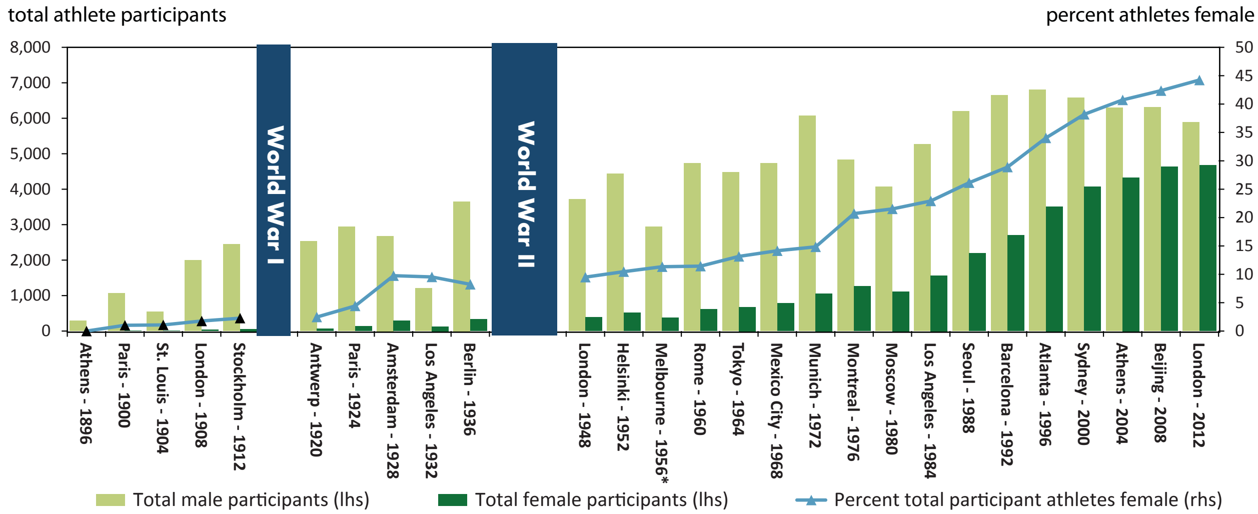
Regarding determinants of total female athletes sent or total medals won, all things being equal, increases in female education levels and representation in the labor force, while clearly positive and significant, do not yield particularly dramatic results. However, it is in the gender makeup of individual

20. Similarly, Lowen and colleagues (2014, forthcoming) examine the values of the US regression residuals to infer the impact of Title IX, but do not find a consistent pattern.

NOC delegations where we see the clearest gains for women specifically. Even for modestly sized delegations from small or poor countries, heightened female education and labor force participation can lead to both higher female athletic representation and a higher share of national glory relative to their male compatriots. For the spectators at home witnessing this transformation, the conveyed ideational effect of women as valuable members of society cannot be ignored.

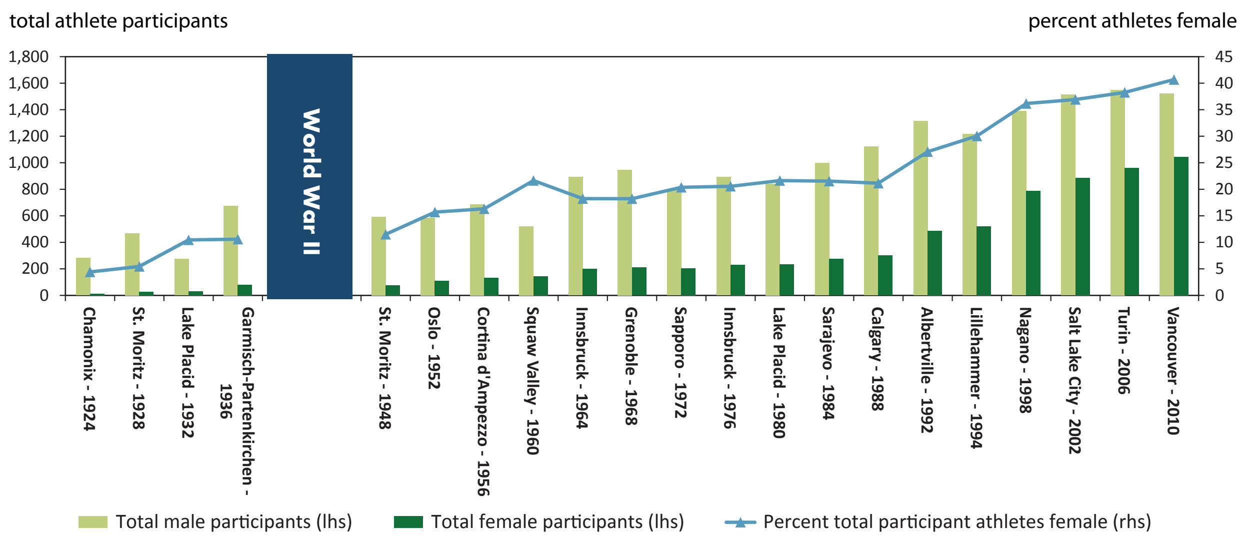
Like previous researchers, we find that medal performance is affected by large-scale boycotts, as occurred at the 1980 Moscow and 1984 Los Angeles Games. We also find that the opening of the competitions to openly professional athletes may have some impact in leveling the playing field for athletes from poorer countries. But the historical record for women's medal achievement is utterly distorted by the East German program of systematic administration of PEDs that was applied particularly intensely to female athletes. At its peak in the 1970s and 1980s, we estimate that the East German doping program was responsible for 17 percent of the total medals awarded to women, equivalent to the total medal hauls of the US and Soviet women's teams in 1972, the last Summer Games not marred by widespread doping.

Figure 1 Female athlete representation at the Summer Olympic Games, 1896–2012



* For Melbourne - 1956, information on the Equestrian Games not included.
Sources: International Olympic Committee, Guttman (2002), Senn (1999).

Figure 2 Female athlete representation at the Winter Olympic Games, 1924–2010



Sources: International Olympic Committee, Guttman (2002), Senn (1999).

Table 1 Participant National Olympic Committees (NOC) that do not send female athletes to Summer Olympic Games, selected years

	London, 1948	Mexico City, 1968		Seoul, 1988		Beijing, 2008
	Afghanistan*	Afghanistan	Libya	Afghanistan	San Marino	Aruba
	Burma (Myanmar)	Algeria	Liechtenstein	American Samoa	Saudi Arabia	Dominica
	Ceylon (Sri Lanka)	Bahamas	Madagascar	Andorra	Solomon Islands	Kiribati
	Colombia	Barbados	Malaysia*	Bahrain	Somalia	Kuwait
	China*	British Honduras (Belize)	Mali	Bangladesh	South Yemen	Liechtenstein
	Cuba*	Bermuda	Malta	Belize	Sudan	Nauru
	Egypt*	Bolivia	Monaco	Bhutan	Swaziland	Netherlands Antilles
	Guyana	Burma (Myanmar)	Morocco*	Botswana	Syria	Qatar*
	India*	Cameroon	Nicaragua	British Virgin Islands	Tanzania	Saudi Arabia
	Iran*	Central African Republic	Niger	Central African Republic	Togo	
	Iraq	Ceylon (Sri Lanka)	Pakistan	Chad	United Arab Emirates	
	Lebanon	Chad	Panama	Djibouti	Zambia*	
	Liechtenstein	Democratic Republic of the Congo	Paraguay	Guinea		
	Malta	Dominican Republic	San Marino	Iran*		
	Monaco	Egypt*	Senegal*	Iraq*		
	Pakistan*	Ethiopia	Sierra Leone	Kuwait*		
	Panama	Fiji	Singapore	Lesotho		
	Peru*	Greece*	Sudan	Libya		
	Philippines*	Guinea	Suriname	Malawi		
	Portugal*	Guyana	Syria	Maldives		
	Puerto Rico	Honduras	Tanzania	Mauritania		
	Singapore	Hong Kong	Thailand*	Niger		
	Spain*	India*	Trinidad and Tobago	North Yemen		
	Syria	Indonesia	Tunisia	Oman		
	Trinidad and Tobago	Iran	Turkey*	Pakistan*		
	Uruguay*	Iraq	Uganda	Panama		
	Venezuela	Ivory Coast (Côte d'Ivoire)	US Virgin Islands	Paraguay		
		Kuwait	Venezuela*	Qatar		
		Lebanon	Zambia	Samoa		
Total participating NOCs	**59**	**112**		**159**		**204**
NOCs that send no female athletes (percent)	**46**	**52**		**26**		**4**

* NOC sent 20 or more male participants

Source: International Olympic Committee.

Table 2 Bernard and Busse determinants of success revisited

	Full sample, 1960–2012		
	(2.1)	**(2.2)**	**(2.3)**
Dependent variable	**Total medal share**	**Male medal share**	**Female medal share**
Log GDP per capita (1990 G-K $I)	1.879***	0.286**	–0.00352
	(0.149)	(0.122)	(0.256)
Log population	1.380***	1.422***	2.625***
	(0.127)	(0.122)	(0.228)
Average years schooling (total 15 and older population)		0.833***	1.816***
		(0.0821)	(0.196)
Constant	–39.74***	–33.87***	–62.23***
	(3.361)	(2.678)	(5.514)
Sigma	2.879***	2.587***	4.650***
	(0.234)	(0.207)	(0.473)
Observations	1,656	1,443	1,170
Model type	**Tobit**	**Tobit**	**Tobit**
Additional controls	**Time dummy**	**Time dummy**	**Time dummy**

G-K $1 = Geary-Khamis international dollar

Notes: Robust standard errors in parentheses. *** and ** represent, respectively, $p<0.01$ and $p<0.05$. All dependent variables scaled from 0–100 percent.

Sources: International Olympic Committee (female/male Olympic medal shares), World Bank (population), Bolt and van Zanden (2013) (GDP per capita), Barro and Lee (2013) (average years female schooling), and authors' estimates.

Table 3 Correlates of female athlete inclusion and success at the Summer Olympic games, 1960–2012

Variable	**Female participation share (n=2056)** Total female participants sent by NOC / total number all female athletes at Games	**NOC participant gender share (n=2056)** Total female participants sent by NOC / total participants sent by NOC	**Female medal share (n=1632)** Total medals in female events won by NOC / total female event medals available at Games	**NOC medal gender share (n=764)** Total medals in female events won by NOC / total medals won by NOC in all events
Current host dummy	0.3696* *n*=2056	0.037 *n*=2056	0.2503* *n*=1632	0.037 *n*=764
Post-host dummy	0.1656* *n*=2056	0.025 *n*=2056	0.1143* *n*=1632	0.016 *n*=764
Communist bloc dummy	0.3550* *n*=2056	0.0598* *n*=2056	0.3707* *n*=1632	0.040 *n*=764
Country population	0.3014* *n*=1997	0.0948* *n*=1997	0.3090* *n*=1597	0.1004* *n*=754
Small state dummy	–0.1816* *n*=2056	0.0615* *n*=2056	–0.1163* *n*=1632	–0.067 *n*=764
GDP per capita (1990 G-K I$)	0.4053* *n*=1657	0.2131* *n*=1657	0.2216* *n*=1336	0.1613* *n*=733
Average years total female schooling	0.3747* *n*=1578	0.4023* *n*=1578	0.2518* *n*=1272	0.2601* *n*=690
Percent population Muslim	–0.243* *n*=1836	–0.1872* *n*=1836	–0.1247* *n*=1494	–0.0899* *n*=732
Adolescent fertility rate	–0.2919* *n*=1896	–0.2256* *n*=1896	–0.1647* *n*=1525	–0.1198* *n*=745
Percent population urbanized	0.2564* *n*=1983	0.0970* *n*=1983	0.1430* *n*=1582	0.1019* *n*=747
Ratio female to male labor force participation rate	0.1457* *n*= 888	0.2616* *n*= 888	0.1266* *n*=846	0.1788* *n*=391

G-K $1 = Geary-Khamis international dollar; NOC = National Olympic Committee

Notes: * represents coefficients significant at 95 percent or higher. All coefficients are pairwise correlates of the dependent variable at the top of each column for the entire sampling period 1960–2012, except for ratio female to male labor force participation, which contains values only for the period 1996–2012.

Sources: International Olympic Committee (female/male Olympic participation and medal shares, current host dummy, post-host dummy), World Bank (population, adolescent fertility rate, percent population urbanized, ratio female to male labor force participation), Bolt and van Zanden (2013) (GDP per capita), Barro and Lee (2013) (average years female schooling), ARDA (percent population Muslim), and authors' estimates.

Table 4 Binary choice models for female participation and medaling at the Summer Olympic Games

	Full sample, 1960–2012			Modern sample, 1996–2012		
	(4.1)	(4.2)	(4.3)	(4.4)	(4.5)	(4.6)
Dependent variable	**NOC sent at least one female athlete**	**NOC won at least one medal in female event**	**NOC won at least one gold medal in female event**	**NOC sent at least one female athlete**	**NOC won at least one medal in female event**	**NOC won at least one gold medal in female event**
Sampling frame	**1960–2008**	**1960–2012**	**1960–2012**	**1996–2008**	**1996–2012**	**1996–2012**
Log GDP per capita (1990 G-K $I)	0.172* (0.0964)	0.462*** (0.133)	0.316** (0.134)	0.0177 (0.181)	0.454*** (0.156)	0.264 (0.180)
Log population	0.326*** (0.0482)	0.476*** (0.0451)	0.522*** (0.0440)	0.433*** (0.118)	0.516*** (0.0592)	0.534*** (0.0605)
Small state dummy	–0.536** (0.266)	–1.154** (0.579)	n.a.	–0.170 (0.465)	–0.959 (0.600)	n.a.
Average years total female schooling	0.117*** (0.0348)	0.244*** (0.0357)	0.293*** (0.0390)	0.199*** (0.0692)	0.167*** (0.0430)	0.278*** (0.0511)
Ratio female to male labor force participation (1 = 1 percent)	n.a.	n.a.	n.a.	0.00880 (0.00621)	0.0214*** (0.00531)	0.0228*** (0.00703)
Percent population Muslim (0–100)	–0.0140*** (0.00155)	–0.00317 (0.00215)	–0.00348 (0.00281)	–0.00908** (0.00384)	2.43e–06 (0.00291)	–0.00106 (0.00354)
Communist bloc	1.106*** (0.344)	1.575*** (0.179)	1.340*** (0.201)	n.a.	0.926** (0.364)	1.116*** (0.414)
Adolescent fertility rate	0.00267* (0.00136)	–0.00151 (0.00242)	–0.000132 (0.00282)	0.00688* (0.00396)	–0.00523* (0.00309)	0.000197 (0.00405)
Percent population urban (0–100)	0.00438 (0.00353)	–0.00109 (0.00428)	–0.00297 (0.00389)	–0.0118** (0.00598)	0.00424 (0.00524)	0.00290 (0.00488)
Constant	–4.987*** (1.062)	–13.88*** (1.440)	–14.60*** (1.406)	–6.012** (2.612)	–15.55*** (1.884)	–16.27*** (2.300)
Observations	1,279	1,132	1,132	511	602	602
Model type	**Probit**	**Probit**	**Probit**	**Probit**	**Probit**	**Probit**
Time controls	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**
Country controls	**No**	**No**	**No**	**No**	**No**	**No**

G-K $1 = Geary-Khamis international dollar; n.a. = not applicable; NOC = National Olympic Committee

Notes: Robust standard errors in parentheses. ***, **, and * represent, respectively, p<0.01, p<0.05, and p<0.1. In models 4.1 and 4.4, the year 2012 predicts female participation perfectly and observations in these years are dropped. In model 4.4, Communist bloc dummy is not included because it predicts outcomes perfectly. In models 4.3 and 4.6, small state dummy is not included because it predicts outcome perfectly. Current host dummy excluded in all models because perfect predictor of success.

Sources: International Olympic Committee (female/male Olympic participation and medals), World Bank (population, adolescent fertility rate, percent population urbanized, ratio female to male labor force participation, small state dummy), Bolt and van Zanden (2013) (GDP per capita), Barro and Lee (2013) (average years female schooling), ARDA (percent population Muslim), and authors' estimates.

Table 5 NOC share of total female participant athletes at the Summer Olympics

	Full sample, 1960–2012				Modern sample, 1996–2012	
Dependent variable: Female participation share, 0–100 percent (total female athletes sent by NOC / total number of female athletes at Olympic games)	**(5.1)** **Simple pooled OLS**	**(5.2)** **Full instrument system GMM (lags 3+)**	**(5.3)** **Restricted instrument system GMM (lags 3-5 only)**	**(5.4)** **Collapsed instrument system GMM**	**(5.5)** **Full instrument system GMM (lags 2+)**	**(5.6)** **Collapsed instrument system GMM**
Sampling frame	**1960–2012**	**1960–2012**	**1960–2012**	**1960–2012**	**1996–2012**	**1996–2012**
Log GDP per capita (1990 G-K $I)	0.0130 (0.0197)	0.0214 (0.0214)	0.0192 (0.0195)	0.0377 (0.0250)	0.0908** (0.0386)	0.0771 (0.0487)
Log population	0.0459*** (0.0128)	0.0878*** (0.0278)	0.0774*** (0.0296)	0.111* (0.0599)	0.152*** (0.0510)	0.129* (0.0722)
Average years total female schooling	0.0215** (0.00851)	0.0404*** (0.0152)	0.0358** (0.0160)	0.0503 (0.0311)	0.0445** (0.0196)	0.0371 (0.0255)
Ratio female to male labor force participation (1 = 1 percent)	n.a.	n.a.	n.a.	n.a.	0.00370** (0.00165)	0.00296 (0.00227)
Current host dummy	3.338*** (0.472)	3.401*** (0.466)	3.357*** (0.472)	3.305*** (0.467)	2.640*** (0.530)	2.556*** (0.618)
Post-host dummy	–1.612*** (0.316)	–1.484*** (0.311)	–1.559*** (0.350)	–1.657*** (0.534)	–0.327 (0.452)	–0.523 (0.628)
Communist bloc	0.536*** (0.148)	0.641*** (0.196)	0.611*** (0.196)	0.699** (0.277)	0.0975 (0.211)	0.0923 (0.172)
LDV (t-1)	0.565*** (0.0542)	0.552*** (0.0660)	0.573*** (0.0740)	0.599*** (0.127)	0.388*** (0.131)	0.450** (0.194)
LDV (t-2)	0.286*** (0.0505)	0.226*** (0.0462)	0.225*** (0.0503)	0.129* (0.0717)	0.219*** (0.0621)	0.212*** (0.0644)
Constant	–0.353 (0.323)	–1.827*** (0.531)	–1.611*** (0.562)	–2.399* (1.222)	–3.653*** (1.220)	–3.073* (1.779)
Observations	1,136	1,136	1,136	1,136	603	603
Number individual NOCs	n.a.	136	136	136	130	130
Model type	**Pooled OLS**	**One–step system GMM**	**One–step system GMM**	**One–step system GMM**	**One–step system GMM**	**One–step system GMM**
Time controls	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**
Country controls	**No**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**
Number of instruments	n.a.	**138**	**72**	**35**	**71**	**28**
Arellano-Bond test for AR(2) in first differences (p-value)	n.a.	**0.143**	**0.164**	**0.543**	**0.226**	**0.214**
Hansen test of overriding restrictions (p-value)	n.a.	**0.749**	**0.015**	**0.522**	**0.255**	**0.758**

AR(2) = second-order autoregression; G-K $1 = Geary-Khamis international dollar; GMM = generalized methods of moments; LDV = lagged dependent variable; n.a. = not applicable; NOC = National Olympic Committee; OLS = ordinary least square

Notes: Robust standard errors in parentheses. ***, **, and * represent, respectively, p<0.01, p<0.05, and p<0.1. All system GMM models use T-statistics, orthogonal variations, and robust standard errors. Lagged dependent variables treated as endogenous and instrumented GMM-style; all other variables assumed strictly exogenous.

Source: International Olympic Committee (female Olympic participation shares, current host dummy, post-host dummy), World Bank (population, ratio female to male labor force participation), Bolt and van Zanden (2013) (GDP per capita), Barro and Lee (2013) (average years female schooling), and authors' estimates.

Table 6 Share of total female participant athletes within individual NOCs at Summer Olympic games

	Full sample, 1960–2012				Modern sample, 1996–2012	
Dependent variable: NOC participant gender share, 0–100 percent (total female athletes sent by NOC / total number athletes sent by NOC)	**(6.1)** **Simple pooled OLS**	**(6.2)** **Full instrument system GMM (lag 2+)**	**(6.3)** **Restricted instrument system GMM (lags 2–4 only)**	**(6.4)** **Collapsed instrument system GMM**	**(6.5)** **Full instrument system GMM**	**(6.6)** **Collapsed instrument system GMM**
Sampling frame	**1960–2012**	**1960–2012**	**1960–2012**	**1960–2012**	**1996–2012**	**1996–2012**
Log GDP per capita (1990 G-K $I)	–0.768 (0.532)	–0.614 (0.693)	–0.600 (0.689)	–0.762 (0.750)	0.0315 (1.032)	0.0287 (1.078)
Log population	0.946*** (0.254)	1.163*** (0.418)	1.160*** (0.418)	1.320*** (0.471)	1.529** (0.593)	1.652*** (0.609)
Average total years female schooling	0.771*** (0.196)	0.859*** (0.256)	0.844*** (0.254)	1.003*** (0.288)	0.363 (0.380)	0.386 (0.402)
Ratio female to male labor force participation (1 = 1 percent)	n.a.	n.a.	n.a.	n.a.	0.206*** (0.0536)	0.218*** (0.0554)
Current host dummy	3.466** (1.564)	3.234* (1.870)	3.180* (1.871)	3.466* (1.871)	1.819 (4.065)	1.775 (4.070)
Post-host dummy	0.540 (1.165)	0.540 (1.005)	0.525 (0.994)	1.060 (1.102)	1.231 (2.141)	1.087 (2.275)
Communist bloc	4.264*** (1.220)	5.734*** (1.953)	5.605*** (1.919)	5.772*** (1.991)	4.895* (2.888)	5.024 (3.058)
LDV (t-1)	0.465*** (0.0320)	0.321*** (0.0415)	0.323*** (0.0417)	0.218*** (0.0496)	0.169** (0.0756)	0.123* (0.0722)
Constant	–4.759 (5.069)	7.114 (7.218)	7.080 (7.169)	8.669 (8.208)	–9.539 (12.28)	–10.85 (12.73)
Observations	1,280	1,280	1,280	1,280	629	629
Number individual NOCs	n.a.	137	137	137	131	131
Model type	**Pooled OLS**	**One–step system GMM**	**One–step system GMM**	**One–step system GMM**	**One–step system GMM**	**One–step system GMM**
Time controls	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**
Country controls	**No**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**
Number of instruments	n.a.	**151**	**73**	**36**	**71**	**28**
Arellano-Bond test for AR(2) in first differences (p-value)	n.a.	**0.732**	**0.727**	**0.942**	**0.332**	**0.364**
Hansen test of overriding restrictions (p-value)	n.a.	**0.752**	**0.492**	**0.359**	**0.38**	**0.522**

AR(2) = second-order autoregression; G-K $1 = Geary-Khamis international dollar; GMM = generalized methods of moments; LDV = lagged dependent variable; n.a. = not applicable; NOC = National Olympic Committee; OLS = ordinary least square

Notes: Robust standard errors in parentheses. ***, **, and * represent, respectively, p<0.01, p<0.05, and p<0.1. All system GMM models use T-statistics, orthogonal variations, and robust standard errors. Lagged dependent variables treated as endogenous and instrumented GMM-style; all other variables assumed strictly exogenous.

Sources: International Olympic Committee (female/male participation gender share, current host dummy, post-host dummy), World Bank (population, ratio female to male labor force participation), Bolt and van Zanden (2013) (GDP per capita), Barro and Lee (2013) (average years female schooling), and authors' estimates.

Table 7 Share of medals in female event won by NOC at the Summer Olympic Games

	Full sample, 1960–2012			Modern sample, 1996–2012	
Dependent variable: Female medal share, 0–100 percent (total medals won in female event by NOC / total number female medals available at games)	**(7.1)** Tobit	**(7.2)** Random effects tobit	**(7.3)** Random effects tobit	**(7.4)** Tobit	**(7.5)** Tobit
Sampling frame	**1960–2012**	**1960–2012**	**1960–2012**	**1996–2012**	**1996–2012**
Log GDP per capita (1990 G-K $I)	0.818*** (0.306)	1.166*** (0.391)	0.615** (0.295)	0.543*** (0.150)	0.244*** (0.0754)
Log population	2.209*** (0.180)	2.005*** (0.262)	1.199*** (0.201)	1.403*** (0.117)	0.377*** (0.0507)
Average total years female schooling	1.359*** (0.147)	0.987*** (0.147)	0.628*** (0.114)	0.615*** (0.0823)	0.164*** (0.0341)
Ratio female to male labor force participation (1 = 1 percent)	n.a.	n.a.	n.a.	0.0379*** (0.00858)	0.0101*** (0.00343)
Current host dummy	5.016*** (1.538)	3.722*** (0.695)	2.732*** (0.538)	3.661*** (0.881)	1.027** (0.453)
Communist bloc	6.667*** (1.003)	4.302*** (0.674)	3.102*** (0.515)	2.229*** (0.666)	0.446 (0.284)
LDV (t–1)	n.a.	n.a.	0.478*** (0.0467)	n.a.	0.798*** (0.0312)
Constant	–58.21*** (5.129)	–53.93*** (5.526)	–32.16*** (4.210)	–36.69*** (2.922)	–10.99*** (1.175)
Sigma	4.173*** (0.408)	3.455*** (0.304)	2.445*** (0.272)	1.845*** (0.113)	0.955*** (0.0444)
Observations	1,170	1,170	978	614	572
Number individual NOCs	n.a.	137	134	n.a.	n.a.
Model type	**Tobit**	**Random effects tobit**	**Random effects tobit**	**Tobit**	**Tobit**
Time controls	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**
Country controls	**No**	**Yes**	**Yes**	**No**	**No**

G-K $1 = Geary-Khamis international dollar; LDV = lagged dependent variable; n.a. = not applicable; NOC = National Olympic Committee

Notes: Robust standard errors in parentheses. *** and ** represent, respectively, p<0.01 and p<0.05. Robust except in cases of random effects tobit. In the modern sample, a random effects tobit with lagged dependent variable was estimated, but yielded unstable results and is not reported.

Sources: International Olympic Committee (female Olympic medals shares, current host dummy), World Bank (population, ratio female to male labor force participation), Bolt and van Zanden (2013) (GDP per capita), Barro and Lee (2013) (average years female schooling), and authors' estimates.

Table 8 Female medals won as a share of NOC total medals won at Summer Games

	Full sample, 1960–2012				Modern sample, 1996–2012	
	(8.1)	**(8.2)**	**(8.3)**	**(8.4)**	**(8.5)**	**(8.6)**
Dependent variable: NOC medal gender share, 0–100 percent (total medals won in female event by NOC / total medals won by NOC)	**Tobit**	**Full instrument system GMM (lag 2+)**	**Restricted instrument system GMM (lags 2–4)**	**Collapsed instrument system GMM**	**Tobit**	**Collapsed instrument system GMM**
Sampling frame	**1960–2012**	**1960–2012**	**1960–2012**	**1960–2012**	**1996–2012**	**1996–2012**
Log GDP per capita (1990 G-K $I)	−2.879 (2.406)	−2.333 (2.226)	−2.492 (2.107)	−2.783 (2.436)	−1.977 (2.754)	−2.117 (3.011)
Log population	2.713*** (0.923)	1.483 (0.927)	1.440 (0.900)	1.823* (0.974)	3.076** (1.247)	2.718** (1.284)
Average total years female schooling	3.298*** (0.834)	1.809*** (0.585)	1.808*** (0.574)	2.176*** (0.652)	0.940 (1.117)	0.640 (1.006)
Ratio female to male labor force participation (1 = 1 percent)	n.a.	n.a.	n.a.	n.a.	0.371*** (0.142)	0.326** (0.147)
Current host dummy	10.39* (5.712)	5.490 (4.199)	5.317 (4.142)	5.083 (4.249)	−2.644 (5.745)	−1.120 (4.195)
Communist bloc	12.47*** (3.469)	7.680*** (2.866)	6.645** (2.605)	6.891** (2.900)	−0.255 (4.744)	1.637 (3.911)
LDV (t–1)	0.558*** (0.0696)	0.306*** (0.0919)	0.342*** (0.0970)	0.233** (0.0938)	0.529*** (0.0839)	0.227* (0.125)
Constant	−35.81 (23.25)	5.295 (26.01)	6.235 (24.88)	3.215 (27.69)	−52.87* (28.52)	−27.08 (32.86)
Sigma	25.77*** (1.458)	n.a.	n.a.	n.a.	25.47*** (1.752)	n.a.
Observations	532	532	532	532	276	276
Number individual NOCs	n.a.	80	80	80	n.a.	73
Model type	**Tobit**	**One−step system GMM**	**One−step system GMM**	**One−step system GMM**	**Tobit**	**One−step system GMM**
Time controls	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**
Country controls	**No**	**Yes**	**Yes**	**Yes**	**No**	**Yes**
Number of instruments	n.a.	**150**	**72**	**35**	n.a.	**27**
Arellano-Bond test for AR(2) in first differences (p-value)	n.a.	**0.4**	**0.364**	**0.48**	n.a.	**0.776**
Hansen test of overriding restrictions (p-value)	n.a.	**1**	**0.596**	**0.381**	n.a.	**0.351**

AR(2) = second-order autoregression; G-K $1 = Geary-Khamis international dollar; GMM = generalized methods of moments; LDV = lagged dependent variable; n.a. = not applicable; NOC = National Olympic Committee

Notes: Robust standard errors in parentheses. ***, **, and * represent, respectively, p<0.01, p<0.05, and p<0.1. Robust except in cases of random effects tobit; (2) All system GMM models utilize T-statistics, orthogonal variations, and robust standard errors. Lagged dependent variables treated as endogenous and instrumented GMM-style; all other variables assumed strictly exogenous. In models 8.1–8.6, only NOCs that won one or more total medals are included, since a zero-denominator ratio has little meaning in this context.

Sources: International Olympic Committee (medal gender share, current host dummy), World Bank (population, ratio female to male labor force participation), Bolt and van Zanden (2013) (GDP per capita), Barro and Lee (2013) (average years female schooling), and authors' estimates.

Table 9 Herfindahl participation index rankings for 2012 participant NOCs

NOC	Herfindahl index for female participation	Number female participants	Herfindahl ratio (female/male)	Herfindahl index for female medaling	Number medals won in female events
		Top 15 NOCs			
France	0.083	143	1.1	0.200	15
Republic of Korea	0.084	111	1.2	0.306	7
Japan	0.084	156	0.9	0.170	17
Great Britain	0.086	260	1.1	0.130	20
Azerbaijan	0.092	14	0.5	1.000	2
Brazil	0.095	119	1.0	0.278	6
Chinese Taipei	0.101	25	0.7	0.500	2
Germany	0.101	172	1.2	0.195	13
China	0.103	213	1.1	0.114	50
Poland	0.107	86	0.9	0.200	5
Canada	0.109	156	0.9	0.160	9
Tunisia	0.110	20	0.8	1.000	1
United States	0.111	270	1.0	0.166	58
Russia	0.122	226	1.4	0.198	44
Netherlands	0.126	79	0.9	0.223	11
		Bottom 15 NOCs			
Mongolia	0.266	13	1.3	1.000	1
North Korea	0.285	38	1.3	0.556	3
Norway	0.294	28	2.2	1.000	1
Cuba	0.311	44	1.7	0.556	3
Moldova	0.313	8	1.7	1.000	1
Tajikistan	0.333	3	1.6	1.000	1
Malaysia	0.337	13	1.9	1.000	1
Ireland	0.341	28	2.2	1.000	1
Lithuania	0.346	23	2.2	0.500	2
Armenia	0.375	4	2.0	1.000	1
Bahrain	0.594	8	1.0	1.000	1
Kenya	0.815	20	1.0	1.000	4
Montenegro	0.876	15	1.6	1.000	1
Ethiopia	0.883	15	1.0	1.000	5
Jamaica	0.917	22	1.0	1.000	5

NOC = National Olympic Committee

Notes: Top/bottom 15 ranked countries ranked by Herfindahl index for female participation, from least concentrated (lowest index) to most concentrated (highest index). We include only NOCs that won at least one medal in a female event in 2012, of which there are 60 total.

Source: International Olympic Committee.

Table 10 NOC Herfindahl indexes on socioeconomic determinants, 2000–2012

	Herfindahl index for female participation			Herfindahl index for female medaling		
	(10.1)	(10.2)	(10.3)	(10.4)	(10.5)	(10.6)
	Tobit	Tobit with LDV	Random effects tobit	Tobit	Tobit with LDV	Random effects tobit
Sampling frame	All available NOCs (2000–2012) that sent at least one female athlete			All available NOCs (2000–2012) that won at least one medal in a female event		
Log GDP per capita (1990 G-K $I)	–0.0919*** (0.0151)	–0.0626*** (0.0148)	–0.0952*** (0.0239)	–0.0910** (0.0394)	–0.0925** (0.0392)	–0.131** (0.0571)
Log population	–0.0737*** (0.00844)	–0.0564*** (0.0100)	–0.0767*** (0.0127)	–0.185*** (0.0217)	–0.181*** (0.0231)	–0.195*** (0.0276)
Average total years female schooling	–0.0332*** (0.00621)	–0.0262*** (0.00629)	–0.0330*** (0.00930)	–0.120*** (0.0190)	–0.117*** (0.0197)	–0.117*** (0.0241)
Current host dummy	–0.0270 (0.0515)	–0.0231 (0.0460)	–0.0361 (0.0951)	–0.162* (0.0841)	–0.155* (0.0840)	–0.132 (0.154)
Communist bloc	–0.0338 (0.0622)	–0.0158 (0.0582)	–0.0452 (0.101)	–0.218** (0.105)	–0.211* (0.108)	–0.240 (0.181)
Labor ratio	–0.00104 (0.000694)	–0.001 (0.000661)	–0.00132 (0.000871)	–0.000997 (0.00228)	–0.00111 (0.00233)	–0.00199 (0.00310)
Lagged dependent variable	n.a.	0.289*** (0.0636)	n.a.	n.a.	0.0563 (0.0933)	n.a.
Constant	2.724*** (0.173)	2.020*** (0.247)	2.827*** (0.273)	5.946*** (0.518)	5.846*** (0.561)	6.555*** (0.738)
Sigma	0.237*** (0.0116)	0.226*** (0.0121)	0.172*** (0.0147)	0.334*** (0.0218)	0.333*** (0.0218)	0.203*** (0.0418)
Observations	501	499	501	199	199	199
Number individual NOCs	n.a.	n.a.	131	n.a.	n.a.	67
Time controls	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**
Country controls	**No**	**No**	**Yes**	**No**	**No**	**Yes**

G-K $1 = Geary-Khamis international dollar; LDV = lagged dependent variable; n.a. = not applicable; NOC = National Olympic Committee

Notes: Robust standard errors in parentheses. ***, **, and * represent, respectively, $p<0.01$, $p<0.05$, and $p<0.1$. Robust except in case of random effects tobit (10.3 and 10.6). In models 10.1–10.3, only NOCs that sent one or more female participants are included in regressions. In models 10.4–10.6 only NOCs that won one or more medals in a female event included in regressions.

Source: International Olympic Committee (Herfindahl index of participation and medaling, current host dummy), World Bank (population, ratio female to male labor force participation), Bolt and van Zanden (2013) (GDP per capita), Barro and Lee (2013) (average years female schooling), and authors' estimates.

Table 11 Substantive effect of selected independent variables on female Olympic participation and performance

	+1 SD shock	Female participation share	NOC gender participation share	Female medal share	NOC medal gender share	NOC participation Herfindahl index	NOC medaling Herfindahl index
Chosen model	n.a.	(5.4)	(6.4)	(7.3)	(8.4)	(10.3)	(10.6)
Average years of total female schooling (percent)	3.2	0.16	3	2	7.0	–10.6	–37.5
Absolute gain at 2012 Games		7.5	n.a.	8.3	n.a.	n.a.	n.a.
Current host (percent)	n.a.	3.3	3.5	2.7	5	–3.6	–13.2
Absolute gain at 2012 Games		154.5	n.a.	11.3	n.a.	n.a.	n.a.
Communist bloc (percent)	n.a.	0.7	5.8	3.1	6.9	–4.5	–24.0
Absolute gain at 2012 Games		32.7	n.a.	12.8	n.a.	n.a.	n.a.
Chosen model		(5.6)	(6.6)	(7.5)	(8.6)	(10.3)	(10.6)
Ratio female to male labor force participation (percent)	20.4	0.06	4	0.21	6.6	–2.7	–4.1
Absolute gain at 2012 Games		2.8	n.a.	0.9	n.a.	n.a.	n.a.

n.a. = not applicable; NOC = National Olympic Committee; SD = standard deviation

Notes: Ratio female to male labor force participation has observed values from 1996–2012 only and uses "modern" models to estimate. For NOC participation, Herfindahl index and NOC medaling Herfindahl index, Olympic periods only between 2000–2012 are estimated. Panels highlighted in light grey are not significant at 10 percent level; those in dark grey are not significant at 20 percent level. In 2012 Summer Games, there were 4,676 total female participants and 414 total female medals were available. Tobit model coefficient values from 7.2, 7.5, 10.3, and 10.6 are interpreted linearly and representative only of uncensored observations.

Source: Authors' estimates.

Table 12 Major boycott effects at the Summer Olympic Games, 1980–84

Boycott sample only, 1980–84		
Dependent variable: Female medal share, 0–100 percent (total medals won in female event by NOC / total number female medals available at games)	**(12.1)** **Tobit (without 1988 medal share)**	**(12.2)** **Tobit (with 1988 medal share)**
Sampling frame	**1980 and 1984**	**1980 and 1984**
Log GDP per capita (1990 G-K $I)	2.779 (2.003)	1.683** (0.714)
Log population	1.472*** (0.503)	0.239 (0.246)
Average total years female schooling	1.109** (0.495)	0.112 (0.212)
Current host dummy	20.93*** (3.444)	6.818*** (1.694)
Communist bloc dummy	10.09** (4.592)	0.0855 (1.184)
Share of female medals in 1988 games	n.a.	1.665*** (0.0986)
Constant	-60.49** (24.89)	-21.20*** (7.497)
Sigma	6.053*** (1.909)	2.300*** (0.282)
Observations	121	114
Number individual NOCs	n.a.	n.a.
Model type	**Tobit**	**Tobit**
Time controls	**Yes**	**Yes**
Country controls	**No**	**No**

G-K $1 = Geary-Khamis international dollar; n.a. = not applicable; NOC = National Olympic Committee

Notes: Robust standard errors in parentheses. *** and ** represent, respectively, $p<0.01$ and $p<0.05$. Robust except in cases of random effects tobit.

Sources: International Olympic Committee (female medal share, current host dummy), World Bank (population), Bolt and van Zanden (2013) (GDP per capita), Barro and Lee (2013) (average years female schooling), and authors' estimates.

Table 13 Summer Games' top performers by female medal share, 1960–2012

Rank	Country/NOC	Year	Share of total female medals won (percent)	Share of total male medals won (percent)
1	East Germany	1980	39.2	14.1
2	East Germany	1976	33.3	8.7
3	Soviet Union	1960	32.2	21.0
4	United States	1984	28.2	18.2
5	Soviet Union	1980	31.4	30.9
6	United States	1968	28.2	18.2
7	Soviet Union	1964	25.3	18.4
8	Soviet Union	1976	24.5	20.5
9	United States	1964	24.2	16.3
10	East Germany	1988	23.2	10.1
11	East Germany	1972	22.5	8.2
12	Soviet Union	1968	20.5	17.1
13	Unified Team of Germany	1960	17.2	6.9
14	United States	1972	17.1	15.0
15	Soviet Union	1972	17.1	17.0

NOC = National Olympic Committee

Source: International Olympic Committee.

Table 14 Doping effects at the Summer Olympic Games

Full sample, 1960–2012		
	(14.1)	**(14.2)**
Dependent variable: Female medal share, 0–100 percent (total medals won in female event by NOC / total number female medals available at games)	**Tobit**	**Random effects tobit**
Sampling frame	**1960–2012**	**1960–2012**
Log GDP per capita (1990 G-K $I)	0.492** (0.231)	0.921*** (0.328)
Log population	2.138*** (0.183)	1.906*** (0.209)
Average total years female schooling	1.286*** (0.140)	0.910*** (0.124)
Current host dummy	5.123*** (1.634)	3.706*** (0.625)
Communist bloc dummy	5.005*** (0.647)	3.746*** (0.583)
East Germany dummy	9.119 (5.848)	10.13*** (3.148)
"Doping" dummy (East Germany 1976–88)	17.41** (6.809)	17.65*** (2.045)
Boycott year dummy (1 if year = 1980 or 1984)	3.107*** (0.856)	2.795*** (0.547)
Constant	–53.01*** (4.473)	–49.18*** (4.466)
Sigma	3.515*** (0.266)	2.653*** (0.239)
Observations	1170	1170
Number individual NOCs	n.a.	137
Model type	**Tobit**	**Random effects tobit**
Time controls	**Yes**	**Yes**
Country controls	**No**	**Yes**

G-K $1 = Geary-Khamis international dollar; n.a. = not applicable; NOC = National Olympic Committee

Notes: Robust standard errors in parentheses. ***, **, and * represent, respectively, p<0.01, p<0.05, and p<0.1. Robust except in cases of random effects tobit.

Sources: International Olympic Committee (female medal share, current host dummy), World Bank (population), Bolt and van Zanden (2013) (GDP per capita), Barro and Lee (2013) (average years female schooling), and authors' estimates.

Table 15 Amateurism at the Summer Olympic Games

Dependent variable: Female medal share, 0–100 percent (total medals won in female event by NOC / total number female medals available at games)	**(15.1)** Split sample tobit	**(15.2)** Split sample tobit	**(15.3)** Full sample tobit	**(15.4)** Full sample random effects tobit
Sampling frame	**1960–88**	**1992–2012**	**1960–2012**	**1960–2012**
Log GDP per capita (1990 G-K $I)	1.441** (0.711)	0.244*** (0.0859)	0.510 (0.323)	0.637 (0.417)
Post-1990 log GDP per capita	n.a.	n.a.	0.0415 (0.297)	–0.0217 (0.287)
Log population	0.943*** (0.354)	0.414*** (0.0622)	0.598*** (0.115)	1.198*** (0.201)
Average total years female schooling	0.447** (0.227)	0.214*** (0.0412)	0.353*** (0.0746)	0.626*** (0.117)
Current host dummy	5.184*** (0.707)	1.215 (0.820)	2.771*** (0.741)	2.734*** (0.538)
Communist bloc dummy	5.054*** (1.101)	0.849* (0.478)	2.793*** (0.610)	3.114*** (0.540)
LDV (t–1)	0.807*** (0.110)	0.806*** (0.0527)	0.825*** (0.0846)	0.478*** (0.0469)
Constant	–36.23*** (8.419)	–11.42*** (1.480)	–19.30*** (2.612)	–32.13*** (4.235)
Sigma	3.695*** (0.403)	1.076*** (0.0802)	2.319*** (0.227)	2.445*** (0.272)
Observations	326	652	978	978
Number country fixed effects	n.a.	n.a.	n.a.	134
Model type	**Tobit**	**Tobit**	**Tobit**	**Random effects tobit**
Time controls	**Yes**	**Yes**	**Yes**	**Yes**
Country controls	**No**	**No**	**No**	**Yes**

G-K $1 = Geary-Khamis international dollar; LDV = lagged dependent variable; n.a. = not applicable; NOC = National Olympic Committee

Notes: Robust standard errors in parentheses. ***, **, and * represent, respectively, p<0.01, p<0.05, and p 0.1. Robust except in cases of random effects tobit. Post-1990 log GDP per capita contains zero values prior to 1990, and log GDP per capita values from 1990–2012.

Sources: International Olympic Committee (female medal shares, current host dummy), World Bank (population), Bolt and van Zanden (2013) (GDP per capita), Barro and Lee (2013) (average years female schooling), and authors' estimates.

Table 16 Pew global attitudes and Olympic outcomes

	Percent agree married relationship where both work more satisfying				Percent agree university education more important for boy than girl			
	(16.1)	(16.2)	(16.3)	(16.4)	(16.5)	(16.6)	(16.7)	(16.8)
Dependent variable	**Female participation share**	**NOC participant gender share**	**Female medal share**	**NOC medal gender share**	**Female participation share**	**NOC participant gender share**	**Female medal share**	**NOC medal gender share**
Sampling frame	**2008 and 2012**	**2008 and 2012**	**2008 and 2012**	**2008 and 2012**	**2008 and 2012**	**2008 and 2012**	**2008 and 2012**	**2008 and 2012**
Log GDP per capita (1990 G-K $I)	–0.215 (0.232)	–1.388 (2.862)	–1.956** (0.907)	–3.702 (5.749)	–0.235 (0.226)	–1.699 (2.704)	–2.064** (0.852)	–5.368 (5.264)
Log population	0.702*** (0.0897)	–0.748 (1.860)	2.205*** (0.472)	1.486 (3.572)	0.717*** (0.0822)	–0.280 (1.861)	2.450*** (0.493)	2.856 (3.791)
Current host dummy	2.734*** (0.457)	0.653 (4.152)	2.984*** (0.548)	–2.242 (4.716)	2.661*** (0.436)	–0.477 (3.795)	2.491*** (0.496)	–8.288 (7.061)
Communist bloc dummy	0.858 (0.523)	12.85* (6.427)	5.671*** (1.149)	19.50* (10.94)	1.075** (0.475)	17.89** (6.898)	6.856*** (1.014)	37.82*** (13.10)
Average total years female schooling	0.658*** (0.0941)	1.957 (1.242)	1.830*** (0.396)	3.742 (2.517)	0.653*** (0.0911)	2.488* (1.250)	1.760*** (0.367)	3.655 (2.533)
Pew poll: percent agree married relationship where both work more satisfying	0.00505 (0.00580)	0.247** (0.116)	0.0292 (0.0193)	0.705** (0.325)				
Pew poll: percent agree university education more important for boy than girl					–0.00849 (0.00829)	–0.127 (0.123)	–0.0638** (0.0246)	–0.651** (0.280)
Constant	–15.10*** (1.246)	34.04 (35.36)	–38.98*** (8.389)	–40.43 (73.24)	–14.57*** (1.128)	43.79 (32.25)	–38.18*** (8.498)	17.83 (60.99)
Sigma	n.a.	n.a.	1.935*** (0.215)	16.44*** (2.127)	n.a.	n.a.	1.849*** (0.200)	16.30*** (2.165)
Observations	39	39	39	35	39	39	39	35
Model type	**Pooled OLS**	**Pooled OLS**	**Tobit**	**Tobit**	**Pooled OLS**	**Pooled OLS**	**Tobit**	**Tobit**
Time controls	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**
Country controls	**No**	**No**	**No**	**No**	**No**	**No**	**No**	**No**

G-K $1 = Geary-Khamis international dollar; n.a. = not applicable; NOC = National Olympic Committee; OLS = ordinary least square

Notes: Robust standard errors in parentheses. ***, **, and * represent, respectively, $p<0.01$, $p<0.05$, and $p<0.1$. Robust except in cases of random effects tobit. All dependent variables scaled 0–100 percent. In models 16.4 and 16.8, only NOCs that won one or more total medals are included, since a zero-denominator ratio has little meaning in this context.

Source: International Olympic Committee (female participation/medal/gender shares, current host dummy), World Bank (population), Bolt and van Zanden (2013) (GDP per capita), Barro and Lee (2013) (average years female schooling), Pew, and authors' estimates.

APPENDIX A: DATA

Our study uses panel data for over 200 individual National Olympic Committees (NOCs) that sent participant athletes to the Summer Olympic Games between the years 1960 to 2012. Basic information recorded on every participant NOC between these years, a total of 2,056 observations, includes country name, 3-letter NOC code, International Olympic Committee (IOC)-designated continental association,[21] total number participant athletes sent (total, female), total medals won (total, male, and female),[22] and total gold medals won (total, female). We have included an additional 200 observations for participant NOCs in the Summer Games between 1948–56 to minimize information loss when employing lagged values in regressions; observations during these years contain information only on participation, medal counts (total, female), and whether the NOC was the host of the Summer Olympic Games.

The IOC Research Center generously supplied all data regarding participation and medal counts at the Summer Games, 1960–2012. The center also supplied participation information for Olympic years between 1948 and 1956, although we use Sports-Reference.com's *Medal Finder* database to fill in medal count values for these early years.[23]

DEPENDENT VARIABLES

The four primary female athlete–specific dependent variables used throughout this paper include the following:

Female participation share. Percentage share (bounded 0 to 100) of total female participants sent by an individual NOC divided by the summation of total female participants from all NOCs at an individual Olympic game. There are no missing values.

Female medal share. Percentage share (bounded 0 to 100) of total medals won by an NOC in female events at an Olympic Game over the total number of female-event medals available that year. NOCs that did not report sending female athletes take a missing value.

NOC participant gender share. Percentage share (bounded 0 to 100) of female athletes sent, divided by total number athletes sent by an individual NOC. There are no missing values.

---

21. There are five continental associations: Asia, Europe, Africa, the Americas, and Oceania.

22. In addition to male-specific and female-specific events, the Summer Games have a smaller proportion of "mixed" events that both genders can technically enter. Medals in mixed events are counted in total medals, but are not counted in the female- or male-specific medaling variables.

23. The Sports-Reference.com *Medal Finder* is available at http://www.sports-reference.com/olympics/friv/medal_finder.cgi.

NOC medal gender share. Percentage share (bounded 0 to 100) of medals won in female events, divided by total medals won by an individual NOC. NOCs that did not report sending female athletes take a missing value. NOCs that report winning zero total medals also take missing values.

In addition, this paper utilizes binary-choice dummy variables for probit regressions: NOC sent at least one female athlete, NOC won at least one medal in female event, and NOC won at least one gold medal in female event.

INDEPENDENT VARIABLES

Our data also include a number of independent variables to control for country/year specific economic, political, and social-environment effects.

GDP per capita (1990 G-K I$). Country GDP per capita denoted in 1990 Geary-Khamis international dollars (G-K I$) to reflect relative purchasing power parities. Individual country data from 1960 to 2010 are sourced from the *Maddison Project Database* (Bolt and van Zanden 2013),[24] and values for 2012 are extrapolated using World Bank estimates for GDP per capita growth.[25] In a small number of cases (see "Notes on Historic Communist Bloc Countries" below), GDP values are supplied by Maddison's *Historical Statistics of the World Economy: 1 – 2003 AD*[26] and Maddison (1995). We use the natural log of these values in our regressions.

Population. Annual total country population between 1960 and 2012 provided by the World Bank.[27] In a small number of cases (see "Notes on Historic Communist Bloc Countries" below), population is supplied by Maddison's *Historical Statistics of the World Economy: 1 – 2003 AD*. We use the natural log of these values in our regressions.

Average years total/female education. National average of total years of education (total population aged 15 and older) and female-specific education (female population aged 15 and older) is reported in five-year increments between 1960 and 2010 from Barro and Lee (2013). Annual values between 1960 and 2010 are linearly interpolated; 2012 values are linearly extrapolated from 2010.

---

24. Maddison Project Database, www.ggdc.net/maddison/maddison-project/data/mpd_2013-01.xlsx (accessed on March 25, 2014).

25. World Bank, *World Development Indicators* (*WDI*), "GDP per capita growth (annual %)," http://data.worldbank.org/indicator/NY.GDP.PCAP.KD.ZG/countries/1W-XQ-EG-SY-MA-IR-SA?display=graph (accessed on March 25, 2014).

26. Angus Maddison's *Historical Statistics of the World Economy: 1 – 2003 AD* dataset, downloaded from http://www.ggdc.net/maddison/maddison-project/data/mpd_2013-01.xlsx (accessed on March 26, 2014).

27. World Bank, *WDI*, "Population, total," http://data.worldbank.org/indicator/ SP.POP.TOTL?page=6 (accessed on September 4, 2013).

Percent population urbanized. Percent of total population (bounded 0 to 100) urbanized is reported annually between 1960 and 2012 from the World Bank.[28]

Adolescent fertility rate. Births per 1,000 females in population aged 15 to 19 are reported annually between 1960 and 2011 from the World Bank; values for 2012 are linearly extrapolated.[29]

Percent population Muslim. Percentage of total country population (bounded 0 to 100) identified as adherents of Islam is reported in five-year increments between 1960 and 2010 from the Association of Religion Data Archives (ARDA) *World Religion Dataset*.[30] Annual values between 1960 and 2010 are linearly interpolated; values for 2012 are linearly extrapolated. In cases where country reports values above 100 percent due to linear extrapolation, value is bounded at 100.

Labor force ratio. The modeled International Labor Organization (ILO) estimate of ratio of population aged 15 and older female to male labor force participation (1 = 1 percent) between 1996–2012 is taken from the World Bank.[31]

Small state dummy. Countries designated as "small states" (Commonwealth Secretariat/World Bank 2000, 4). In general, countries with populations of fewer than 1.5 million are included. These values are treated as time-invariant fixed effects in the data.

Gender-related survey questions. Country-level results for selected gender-related survey questions are taken from the 2010 Pew Research Center's Global Attitudes Project.[32] Specifically, we use two questions from the survey: "What kind of marriage do you think is a more satisfying way of life: one where the husband provides for the family and the wife takes care of the house and the children, or one where both have jobs and both take care of the house and the children?" and "Is a university education more important for a boy than for a girl?" Authors record values for percent agree. Twenty-one countries report values for each survey question. In the authors' models, values from the 2010 Pew Survey are applied to both 2008 and 2012 Summer Olympic observations.

---

28. World Bank, *WDI*, "Urban population (% of total)," http://databank.worldbank.org/ data/views/reports/tableview.aspx (accessed on March 25, 2014).

29. World Bank, *WDI*, "Adolescent fertility rate (births per 1,000 women ages 15-19)," http://data.worldbank.org/indicator/SP.ADO.TFRT (accessed on March 25, 2014).

30. ARDA, *World Religion Dataset: National Religion Dataset*, www.thearda.com/Archive/Files/ Downloads/WRDNATL_DL2.asp (accessed on September 30, 2013; variable used named "ISGENPCT").

31. World Bank, "Ratio of female to male labor force participation rate (%) (modeled ILO estimate)," http://data.worldbank.org/indicator/SL.TLF.CACT.FM.ZS (accessed on May 20, 2014).

32. Pew Research Center, "Gender Equality Universally Embraced, But Inequalities Acknowledged," July 1, 2010, www.pewglobal.org/2010/07/01/gender-equality/ (accessed on March 25, 2014).

Additional dummy variables are included to designate whether a country was the current or post-Olympic host, and whether the country is currently governed by a Communist-style centrally planned economic regime (table A.1).

NOTES ON HISTORIC COMMUNIST BLOC COUNTRIES

For some country values, rough approximations were made to include observations in the regressions. Specifically, a handful of historic NOCs—those for the Soviet Union, Czechoslovakia, Yugoslavia, East Germany, and West Germany—required the authors to take some care in recreating comparable values for the necessary control variables.

All countries in the "Communist bloc" dummy and years in which the dummy is applied between 1960–2012 are shown in table A.1. While the World Bank and other sources generally report data for currently existing countries with Communist legacies (Bulgaria, Hungary, Poland, etc.) as far back as 1960, we report data sourcing and approximation assumptions below for each of our five Soviet bloc–era nations not reported in standard databases. Basically, values are filled in where historical data are available in a comparable form from another source; where data are not available, "core country" values[33] are substituted.

Soviet Union. GDP per capita data from 1960–88 are taken directly from the *Maddison Project Database* (Bolt and van Zanden 2013), which specifically reports the Soviet Union. The population of the Soviet Union is provided by Maddison's *Historical Statistics of the World Economy: 1 – 2003 AD*. Total average years of female education uses values for the Soviet Union from Barro and Lee (2000). The adolescent fertility rate and percent population urban for the Soviet Union for 1960–88 are approximated by using data entries listed for the "core" country, now the Russian Federation. Similarly, the authors' approximate percent of population Muslim uses values listed for the Russian Federation in years 1960–88.

East Germany. GDP per capita data from 1968 to 1988 are provided directly by Maddison (1995, 132), which specifically reports values for East Germany. The population of East Germany is provided directly from Maddison's *Historical Statistics of the World Economy: 1 – 2003 AD* and multiplied by 1,000. Total average years of female education uses values listed for East Germany from Barro and Lee (2000). Percent population Muslim is taken from values listed for the "German Democratic Republic" in the ARDA *World Religion Dataset*. In all other cases, values are approximated using "core country" data listed for Germany from 1960 to 1988.

33. Values from a currently existing country where the previous capital of the Soviet era nation was located. For example, since Prague was the historic capital of Czechoslovakia, values from the Czech Republic are used as proxies for Czechoslovakia.

West Germany. West German GDP per capita in 1990 G-K I$ is not directly reported from any sources that we are aware of. Therefore, we have estimated West Germany's GDP per capita values for 1968–88 through a simple algebraic rearrangement of existing values provided by various Maddison data explained in detail in the footnote.[34] The population of West Germany is provided directly from Maddison's *Historical Statistics of the World Economy: 1 – 2003 AD* and multiplied by 1,000. Total average years of female education uses values listed for West Germany from Barro and Lee (2000). The percent population Muslim is taken from values listed for the Federal Republic of Germany in ARDA *World Religion Dataset*. In all other cases, values are approximated using data listed for Germany from 1960 to 1988.

Yugoslavia. GDP per capita data from 1960 to 1988 are taken directly from the *Maddison Project Database* (Bolt and van Zanden 2013), which specifically report Yugoslavia. The population of Yugoslavia is provided from Maddison's *Historical Statistics of the World Economy: 1 – 2003 AD* and multiplied by 1,000. Total average years of female education uses values listed for Yugoslavia from Barro and Lee (2000). The percent population Muslim is taken from values listed for Yugoslavia in ARDA *World Religion Dataset*. World Bank data on the adolescent fertility rate are approximated using "core" data from Serbia in 1960–88; urban population percent from 1960 to 1988 uses values for Bosnia and Herzegovina because Serbia data were not available.

Czechoslovakia. GDP per capita data from 1960 to 1988 are taken directly from the *Maddison Project Database* (Bolt and van Zanden 2013), which specifically reports Czechoslovakia. The population of Czechoslovakia is provided from Maddison's *Historical Statistics of the World Economy: 1 – 2003 AD* and multiplied by 1,000. Total average years of female education uses values listed for Czechoslovakia from Barro and Lee (2000). The percent population Muslim is taken from values listed for Czechoslovakia in the ARDA *World Religion Dataset*. In all other cases, values for Czechoslovakia between 1960 and 1988 are approximated using listings for the Czech Republic.

---

34. The *Maddison Project Database* (Bolt and van Zanden 2013) and Maddison's *Historical Statistics of the World Economy: 1–2003 AD* provide aggregated GDP per capita and population data for a unified Germany from 1968 to 1988. Additionally, Maddison (1995) gives us data for East German GDP per capita for 1968–88 specifically. Assuming that

$$GDPcapita_{UNIFIED\ GERMANY} = (GDPcapita_{EAST\ GERMANY}\ x\ UnifiedGermanPopulationShare_{EAST\ GERMANY}) + (GDPcapita_{WEST\ GERMANY}\ x\ UnifiedGermanPopulationShare_{WEST\ GERMANY}),$$

simple algebraic reordering will yield the equation for our value in question:

$$GDPcapita_{WEST\ GERMANY} = (GDPcapita_{UNIFIED\ GERMANY} - (\ GDPcapita_{EAST\ GERMANY}\ x\ UnifiedGermanPopulationShare_{EAST\ GERMANY}))\ /\ Unified\ GermanPopulationShare_{WEST\ GERMANY}).$$

Using the above equation, GDP per capita values for West Germany were derived.

NOTES ON MISSING DATA POINTS

Reflecting the reality of the Summer Games, our panel is highly unbalanced. Only 12 percent of total NOCs participated in all 14 Summer Games between 1960 and 2012. One leading reason is that the Games have become far more inclusive than their immediate post-WWII predecessors: 204 NOCs participated in 2012, up from only 83 in 1960. Moreover, many NOCs may not have participated in select years due to boycotts, in-country instability, or war, or, as in the case of the Soviet Union, the state ceased to exist. Therefore, we cannot conclude that the unbalanced nature of the panel stems from completely random causes.

Another significant reason for missing observations is the lack of control-variable data for some countries. For example, there are noncountries and countries that no longer exist in the data: In 1992, 2000, and 2012 NOCs known as "Independent Olympic Participants" or "athletes" entered the Olympics but technically were not sponsored by a specific country. In other cases, countries that no longer exist, such as the Republic of Tanganyika, contain no usable data from any of our primary sources.

The vast majority of observation loss, however, comes from lack of data from one or more control variables in the regressions. A table of missing observations reproduced in table A.2 shows the relative data loss that each specific variable produces. Clearly, educational data from Barro and Lee (2000) appear to be a significant limitation, excluding almost a quarter of our possible 2,056 observations. Moreover, this variable appears to exclude data from historically poor countries and small island nations in particular.[35] Therefore, we cannot say that data are missing purely at random, which may be leading to biases in the estimations toward larger, richer NOCs.

35. A complete list of countries and territories totally excluded from regression due to lack of educational data: American Samoa, Andorra, Angola, Antigua and Barbuda, Aruba, Azerbaijan, The Bahamas, Belarus, Bermuda, Bhutan, Bosnia and Herzegovina, Burkina Faso, Cape Verde, Cayman Islands, Chad, Comoros, Cook Islands, North Korea, Djibouti, Dominica, Equatorial Guinea, Eritrea, Ethiopia, Georgia, Grenada, Guam, Guinea, Guinea-Bissau, Kiribati, Lebanon, Liechtenstein, Madagascar, Marshall Islands, Micronesia, Monaco, Montenegro, Nauru, Netherlands Antilles, Nigeria, Oman, Palau, Palestine, Puerto Rico, St. Kitts and Nevis, Saint Lucia, Samoa, San Marino, São Tomé and Príncipe, Seychelles, Solomon Islands, Somalia, St. Vincent and the Grenadines, Suriname, Macedonia, Turkmenistan, Tuvalu, Uzbekistan, Vanuatu, British Virgin Islands, United States Virgin Islands.

Table A.1 Participant Communist bloc countries in Summer Olympic Games, 1960–2012

NOC	Participating Olympic years
Albania	1972
Bulgaria	1960–88
Cuba	1964–2012
Czechoslovakia	1960–88
North Korea	1972–2012
East Germany	1968–88
Hungary	1960–88
Laos	1980–2012
China	1984–2012
Poland	1960–88
Romania	1960–88
Soviet Union	1960–88
Yugoslavia	1968–88
Vietnam	1980–2012

Source: Authors.

Table A.2 Independent variables missing (or not missing) in Summer Olympic Games data, 1960–2012

Variable	Missing observations	Nonmissing observations	Percent missing
Population	59	1,997	3
GDP per capita (1990 G-K I$)	399	1,657	19
Average years total female schooling	478	1,578	23
Adolescent fertility	160	1,896	8
Percent population urbanized	73	1,983	4
Percent population Muslim	220	1,836	11
Gender labor ratio[a]	119	888	12

G-K $1 = Geary-Khamis international dollar

a. Gender labor ratio contains values only from 1996 to 2012.

Sources: World Bank (population, adolescent fertility, percent population urbanized, gender labor ratio), Bolt and van Zanden (2013) (GDP per capita), Barro and Lee (2013) (Average years female schooling), ARDA (percent population Muslim).

REFERENCES

Andreff, Wladimir. 2013. Economic development as major determinant of Olympic medal wins: predicting the performance of Russian and Chinese teams at Sochi Games. *International Journal of Economic Policy in Developing Countries* 6, no. 4: 314–40.

Barro, Robert, and Jong-Wha Lee. 2000. *International Data on Educational Attainment: Updates and Implications.* CID Working Paper 42. Cambridge, MA: Center for International Development at Harvard University.

Barro, Robert, and Jong-Wha Lee. 2013. A new data set of educational attainment in the world, 1950–2010. *Journal of Development Economics* 104: 184–98.

Bernard, Andrew B., and Meghan R. Busse. 2004. Who Wins the Olympic Games: Economic Resources and Medal Totals. *Review of Economics and Statistics* 86, no. 1: 413–17.

Commonwealth Secretariat/World Bank Joint Task Force on Small States. 2000. Small States: Meeting Challenges in the Global Economy. April 17. Washington. Available at http://www.cpahq.org/cpahq/cpadocs/meetingchallengeing lobaleconomyl.pdf (accessed on September 24, 2014).

Bolt, Jutta, and Jan Luiten van Zanden. 2013. The First Update of the Maddison Project: Re-Estimating Growth Before 1820. Maddison Project Working Paper WP-4. January. University of Groningen, The Netherlands: Maddison Project. Available at http://www.ggdc.net/maddison/maddison-project/publications/wp4.pdf (accessed on September 24, 2014).

Franke, Werner W., and Brigitte Berendonk. 1997. Hormonal Doping and Androgenization of Athletes: a Secret Program of the German Democratic Republic Government. *Clinical Chemistry* 43, no. 7: 1262–79.

Guttmann, Allen. 2002. *The Olympics: the History of the Modern Games*, 2nd ed. Champaign, IL: University of Illinois Press.

Hunt, Thomas. 2011. *Drug Games: the International Olympic Committee and the Politics of Doping, 1960–2008*. Austin: University of Texas Press.

Johnson, Daniel K. N., and Ayfer Ali. 2004. A Tale of Two Seasons: Participation and Medal Counts at the Summer and Winter Olympic Games. *Social Science Quarterly* 85, no. 4: 974–93.

Klein, Michael. 2004. Work and Play: International Evidence of Gender Equity in Employment and Sports. *Journal of Sports Economics* 5, no. 3: 227–42.

Lowen, Aaron, Robert O. Deaner, and Erika Schmitt. 2014 (forthcoming). Guys and Gals Going for Gold: the Role of Women's Empowerment in Olympic Success. *Journal of Sports Economics.*

Lui, Hon-Kwong, and Wing Suen. 2008. Men, Money, and Medals: An Econometric Analysis of the Olympic Games. *Pacific Economic Review* 13, no. 1: 1–16.

Maddison, Angus. 1995. *Monitoring the World Economy, 1820–1992.* Paris: Organization for Economic Cooperation and Development.

Roodman, David. 2009. How to do xtabond2: An introduction to difference and system GMM in Stata. *Stata Journal* 9, no. 1: 86–136.

Senn, Alfred E. 1999. *Power, Politics and the Olympic Games.* Champaign, IL: Human Kinetics.

Ungerleider, Steven. 2013. *Faust's Gold: Inside the East German Doping Machine,* updated edition. CreateSpace Independent Publishing Platform.

Yesalis III, Charles E., Andrea N. Kopstein, and Michael S. Bahrke. 2001. Difficulties in Estimating the Prevalence of Drug Use Among Athletes. In *Doping in Elite Sport: The Politics of Drugs in the Olympic Movement*, ed. Wayne Wilson and Edward Derse. Champaign, IL: Human Kinetics.