CORE
Metadata, citation and similar papers at core.ac.uk
Provided by Research Papers in Economics

Thünen-Series of Applied Economic Theory

Thünen-Reihe Angewandter Volkswirtschaftstheorie

Working Paper No. 108

Country Performance at the International Mathematical Olympiad

by

Golo Henseke

Universität Rostock

Wirtschafts- und Sozialwissenschaftliche Fakultät
Institut für Volkswirtschaftslehre
2009

Country Performance at the International Mathematical Olympiad

Golo Henseke*

October 15, 2009

This study seeks to explain country differences in the performance at the International Mathematical Olympiad. Hypotheses on the relationship between, on one hand, performance at the Olympiads and, on the other, population size and dynamics, economic resources, human capital, schooling quantity and quality, and the political regime are tested with a panel dataset of 97 countries over the period 1993-2006. The analysis distinguishes between cross-country differences and intra-country differences. Results indicate that macro-conditions explain cross-country differences well but fail to predict changes in performance over time. Thus, long-term differences in country characteristics are associated with the average performance of Olympians.

Keywords: Science Olympiads, Talent in mathematics, Country panel, PISA
JEL-Classification: I21, H52, J24

1 Introduction

The International Mathematical Olympiad (IMO) is the oldest international annual science Olympiad, and 565 highly talented students from 104 countries participated in the 2009 competition. However, there has been large variation in the average performance of students, based on nationality. Team members

*University of Rostock, Institute of Economics, Email: golo.henseke@uni-rostock.de. I'm indebted to Thusnelda Tivig, Carsten Ochsen and discussants at the Rostock economic research seminar and at the RZ-HWWI meeting for their helpful suggestions.

from China, Russia and the USA have averaged more than 30 points in the last 5 years, while participants from Kuwait, Saudi Arabia, and Bolivia have scored only around a tenth of that during the same period. In this study, we investigate the macro-level determinants of Olympians' performance at the IMO in order to explain the national variations in average achievement.

Our approach draws on empirical research on country performance at international sporting competitions and tournaments like the Olympic Games and international soccer meets (e. g., Bernard and Busse, 2004; Johnson and Ali, 2002; Leeds and Leeds, 2009; Hoffmann et al., 2002b; Lui and Suen, 2008). The studies have generally found positive relationships between a country's performance and its economic resources, population size, and measures of institutional quality, although with varying magnitudes. All cited studies have an explicit macro-perspective, so they have concentrated on country-specific characteristics that systematically hinder or promote the performance of athletes, rather than on the development of athletic talent per se.

We apply this framework to contestants at the IMO and hence look at the mathematical talent of potential researchers and scientists of tomorrow. In contrast to previous studies, we include measures of the national human capital stock and the quality of secondary education in the analysis, and explicitly distinguish between inter-country and intra-country associations between the macro-conditions and the expected performance at the Olympiads.

Our results indicate that differences among countries at the IMO can be well explained by the size of the effectively participating population, general population dynamics, the market size for academic research, schooling inputs and outputs, and the political system. In particular, the average performance of contestants from countries with large effectively participating populations at the Olympiads is stronger than that of participants from a small effectively participating population. In contrast to studies on sporting events, there is no evidence that GDP per capita is associated with achievement. These findings receive strong support from an analysis of country performance at the International Physics Olympiads. We also find that variations in a country's macro-conditions are not accompanied by a corresponding shift in performance. Although, there is some evidence of a negative impact of long-term population growth on the outcome.

A problem for longitudinal analysis is the high level of persistence in a macro time series. Another problem for this longitudinal analysis in particular is variations in the exams at the Olympiads over the years. Both of these issues may hamper time series comparisons.

The remainder of the paper is structured as follows. In Section 2, we discuss the literature on talent formation and the findings from the Olympic studies.

In Section 3, we derive hypotheses based on several strands of research on the relationship between, on one hand, resources, the talent pool, and institutions, on the other, countries' achievement at the IMO. Section 4 contains an outline of the empirical methodology and a discussion of the data. In Section 5, the findings of the analysis are presented and discussed. Finally, Section 6 summarizes the main findings and concludes.

2 Literature Review

This study focuses on national variations among highly talented individuals in mathematics. Talent can generally be defined as, *"a performance which is distinctly above average in one or more fields of human performance."* (Gagné, 1985, pg. 108). Here, talent in mathematics will be indicated by the participation at the International Mathematical Olympiad.

Talented individuals are of importance to the economic performance of countries, particularly in terms of innovation and knowledge creation. Since Lotka's seminal work (Lotka, 1926), the concentration of scientific output has the regular focus of scholars in social and information sciences (e. g. Merton, 1968; Narin and Breitzman, 1995; Seglen, 1992). Scientific productivity is characterized by the convexity of returns to abilities and skills. A similarly skewed distribution of performance and productivity can be found for a variety of occupations, including physicians, managers, entrepreneurs, inventors, and politicians (Rosen, 1981). It is important for countries to sustain a steady influx of talented individuals, who may become key players in the economy, into the labor force (Murphy et al., 1991).

Individual talent formation is tightly connected to the development of cognitive (e. g., abilities and skills) and non-cognitive (e. g., commitment, motivation) factors and to the socio-cultural background (Heller, 2007). Cognitive and non-cognitive factors are formed during early childhood and stabilize at young ages. They are highly correlated to family background characteristics, like the educational level of parents. Once these factors have manifested themselves, the quality and resources of schools have only minor impact on the variations in endowment levels among children from different socio-economic backgrounds (Cunha and Heckman, 2007).

Talent formation is a time-intensive process, for which cognitive skills are a prerequisite, and individual variations in performance at the novice level are driven by unequal endowment with cognitive skills. However, cognitive skills, while necessary, are not sufficient for talent formation. With an increasing degree of expertise, non-cognitive skills like motivation and task commitment, as well as the approach to learning processes of domain-specific knowledge, grow in importance (Heller, 2007).

Follow-up studies on former participants in the International Mathematical Olympiads confirm the conclusions from the existing theory on talent formation that formation of outstanding talent is dependent on the family background and early access to intellectually stimulating resources (e. g., Campbell and Wu, 1996; Nokelainen et al., 2004; Shoho, 1996). The quality of schooling has a supportive impact on talent formation since proficient teachers, a challenging environment and peers with similar interests promote the development of mathematical talent (Subotnik et al., 1996). Still, the absence of these factors and even the presence of opposing factors—e. g., poor teachers and active hindrance at school—do not necessarily have an adverse impact on the individual Olympian, as long as there is a supportive home environment (Nokelainen, Tirri, Campbell and Walberg, 2007). Motivation, commitment, and non-cognitive factors in general are also major components of Olympiads' talent formation. Many former Olympians state that the level of effort has been besides cognitive skills an important attribute in their talent formation and a crucial reasons for success at IMO (Nokelainen, Tirri and Merenti-Valimaki, 2007; Shoho, 1996). However, a general limitation of many of these studies is the lack of a control group; findings often refer to former Olympians only.

As hypothesized from theory, socio-cultural factors like general attitudes towards mathematics and perception of talent in a society directly influence the direction of individual talent formation but also affect the resources dedicated to international academic competitions. The best-performing countries at the IMO are either in Asia (e. g., China, Iran, Japan, South Korea) or in Eastern Europe (e. g., Bulgaria, Hungary, Romania, Russia). The only exception is the USA, although approximately half of the US team consists of either foreign-born students or students whose parents immigrated to the USA from other top-ranked countries (Andreescu et al., 2008).

In contrast to most of the previous literature on participants at the International Mathematical Olympiad, we focus on country characteristics that contribute to the average performance of already highly talented adolescents. There is similar research on country rankings at the Olympic Games and international soccer performance that explain national performances in terms of various determinants. National teams from countries with a higher level of material wealth and larger populations win on average more medals at the Olympic Games than teams from smaller and poorer countries. Bernard and Busse (2004) found positive relationships between, on one hand, the population size, GDP per capita, host status, communist regimes and, on the other, the achieved national medal share at the Summer Olympics during the last 30 years. Their findings were unaffected by the inclusion of random country effects and held if the lagged dependent variable was added as additional regressor. Also climatic conditions contribute to national performance.

However, the influence of country characteristics depends on the sporting competitions and disciplines analyzed. For example, Johnson and Ali (2002) found differences in national performance between Summer and Winter Games: in contrast to the Summer Games, smaller countries can outperform countries with larger populations at the Winter Olympics.

Population size is generally interpreted as a measure of a country's talent pool. However, Krishna and Haglund (2008) argued that a distinction must be drawn between the overall population and the effectively participating population. Effective participation depends on the infrastructure, human capital and the availability of information. Krishna and Haglund's empirical results for data on the Summer Games 1992-2004 showed a sharp increase in the predictive power of the regression after the inclusion of participation measures, while their estimate for the coefficient of GDP per capita was not significantly different from zero.

The literature on countries average soccer performance points to other determinants like the role of culture and the popularity of the sport (Hoffmann et al., 2002b; Macmillan and Smith, 2007). For example, Latin countries get a bonus from the cultural affinity for soccer. In contrast to the results from the Olympic Games, these studies did not identify a linear influence of population size and showed that effect of GDP per capita is diminishing and might even become negative after reaching a certain threshold.

An issue in the empirical analyses of Olympic Games or soccer performance is the strong persistence of country performance over time and the resulting correlation in the estimated residuals, which leads to biased estimates. The non-randomness in the residuals may reflect omitted variables that can be captured by country specific effects, the presence of serial correlation, or both. Furthermore, there is a high probability that explanatory variables and the non-random error component are correlated with each other. Limitations in the data of sporting competition often do not allow to overcome all statistical problems. Under the strong assumption that country effects are uncorrelated with the explanatory variables, Lui and Suen (2008) showed that the main findings concerning the influence of GDP per capita and population are similar to pooled-estimation approaches and are also robust to other treatments of the non-randomness in the errors (including the lagged dependent variable). However, Lui and Suen's fixed-effects estimation of medal results at the Pan American Games strongly indicated that country effects are correlated to the other explanatory variables. The magnitudes of Lui and Suen's estimates were clearly different from the random-effects estimation.

Overall, previous research on sporting competitions has identified important common determinants that hinder or contribute to the performance of athletes or teams, although the sign of the relationship and the magnitude depends on the discipline being investigated.

3 Hypotheses

Drawing on the literature of country performance at Olympic Games and on soccer results, we derive an empirical framework to explain country-level variations in the average performance per participant at the IMO. The major difference from related work on sporting events is the inclusion of measures of population age structure, human capital and schools' resource endowment and quality. The dependent variable in this study is the time and country-specific average score of the participating teams. The aim is to shed light on the relationship between aggregate measures like population size, income, and characteristics of the school system with the performance of mathematical talents at the IMO.

First, mathematical talent is assumed to be equally and randomly distributed among populations. Countries with bigger population can draw from a greater pool of potential outstanding talent, so they are expected to achieve more points on the average at the competition. However, not every potential talent participates in the selection procedure (Krishna and Haglund, 2008). Participation should be related to a) costs for the individual and b) cost per participant for schools. Urban environments reduce individual costs (e. g. transportation) and allow the fixed costs for selection and training, like special enrichment courses in mathematics at schools and the organization of local mathematical competitions, to be spread more broadly. Thus,

H1.1: Population size positively influence the performance at the IMO.

H1.2: The percentage of urban population increases the average achievement at the IMO.

Second, material resources should play an important role. Material resources and, more generally, the level of economic development are approximated by GDP per capita PPP (in constant 2005 international $) in this study, which represents the value of produced goods and services per capita and the resulting average income. Micro-econometric studies typically find a positive influence of parental income on the cognitive achievement of children (Fuchs and Wößmann, 2007). Studies on Olympic Games also usually report a significantly positive impact of GDP per capita on country performance, although often with diminishing returns (Hoffmann et al., 2002a). However, in contrast to Olympic Games, the relationship between performance and GDP per capita is not as obvious in the context of the IMO. The direct preparation for the IMO is less resource-intensive than it is for sporting competitions, but the process of talent formation requires resources over an extended period of time. Since most countries' level of development changes slowly better developed countries are able to dedicate more resources to children's cognitive development. Thus,

H2: We expect to find a positive inter-country correlation between GDP per capita PPP and the performance at the IMO.

Still, the influence of a change in GDP per capita is difficult to pin down; change in the average income must not have an immediate impact on scores but an increase in resources might go hand in hand with broadened possibilities and other ways to apply and challenge talent.

Third, countries' demographic structures should influence the available resources per child and per adolescent. There are potentially two effects of the age-structure. Micro-economic literature on fertility suggests that the number of children and children's human capital are substitutes (Becker and Lewis, 1973; Hanushek, 1992) so, as the number of births falls, the demand for children's' human capital rises, which demand implies an increase in the resources allocated to skill formation. We use the current crude birth rate to measure aggregate fertility and hypothesize that

H3.1: Lower birth rates benefit skill formation of children.

The other effect of the country's age structure is that relative cohort size may have an influence on individual performance. The *cohort crowding* hypothesis assumes that members of larger cohorts have to struggle more for resources because schooling systems and labor markets do not fully adjust to changes in cohort size (Easterlin, 1978); there is evidence that cohort size is negatively correlated with educational, specifically collegiate, achievement (Bound and Turner, 2007). On the other hand, a larger number of adolescents raises the number of potentially talented students, who can participate at the IMO. Thus, there is a trade-off between the number of potentially talented students and the actualization of talent. The number of adolescents will be measured by the percentage of population aged 15-19 years.

H3.2: Depending on the strength of the opposing effects, the number of people aged 15-19 years could have a positive, neutral, or even a negative influence.

Fourth, the size of the market for academic research and academic jobs should raise the incentive to invest in academic skills and spur the formation of mathematical talent and, consequently, on the average achievement at the IMO. The number of articles published in scientific and technical journals serves as proxy for the size of the prospective market for exceptional math talent. The measure refers to the number of published articles in the fields of physics, biology, chemistry, mathematics, clinical medicine, biomedical research, engineering and technology, and earth and space sciences (World Development Indicators 2008).

H4: A larger number of published articles is positively associated with national performance at the IMO.

Fifth, communist countries and former communist countries generally outperform other nations at many international sporting competitions (Bernard and Busse, 2004; Leeds and Leeds, 2009). Communist regimes traditionally build institutions that strongly support the preparation and training of talented individuals. Regimes in most former communist countries have changed with the fall of the Iron Curtain, but institutions change slowly. If institutions related to mathematical competitions are still intact in those countries, there will be a measurable conditional impact on IMO scores. Thus, we test whether there are influences of communist regimes and whether the legacy of former communist regimes is still evident. Following the literature of sporting competition, we expect that

H5: Communist regimes positively influence national performance at the IMO.

Finally, the reputation and the quality of national institutions related to the IMO are expected to influence the performance at the IMO. We argue that reputation and institutional quality can be adequately approximated by two variables: the number of times countries have participated and the number of times a country has hosted the IMO. The longer a country's history of participation at the IMO, the better developed the relevant national institution should be. Hence, contestants' performance should rise with the number of times countries have participated at the IMO. However, the learning effect is likely to level off, so increases in performance will diminish over time. The number of hosting activities also reflects the quality of national institutions, the available resources for the competitions, and the reputation of the IMO in the country, so the number of hosting activities is assumed to be positively correlated with the country's average results at the competition.

H6.1: National performance at the IMO should increases with the number of times a country has hosted the IMO.

H6.2: National performance at the IMO should raise with the number of times countries have participated at the IMO.

Despite the number of explanatory variables, we believe that country-specific unobserved heterogeneity that is due, for instance, to omitted variables like quality of the educational system and particularities of the national IMO preparation, still influences the results to a considerable extent. These country-specific factors are assumed to be constant over time, so they can be statistically captured by time-invariant country effects. It is also likely that the country effects are correlated with other explanatory variables, e. g., the level of GDP per capita and cohort size. Another unobserved influence that affects all countries is the varying difficulty of the examination problems. To control for these common period effects, the regression includes a set of year dummies.

Besides the macro-level conditions outlined here, talent formation itself should specifically depend on human capital of the parents and the quantity and quality of schooling. Therefore, in a finer-grained analysis, we look for influences by different input and output characteristics of the educational system and human capital on the outcome at the IMO. The assumption is that performance at the Math Olympiad is related to the average quality of talent formation and cognitive achievement in mathematics. Cognitive achievement can be interpreted as the joint outcome of family inputs and school resources (Todd and Wolpin, 2003). The parental level of human capital is a strong predictor of family input and, thus, cognitive achievement on the individual level (Fuchs and Wößmann, 2007). Human capital is the outcome of investments in knowledge and skills that increase future productivity and income. Human capital refers to the quantity and quality of education, as well as to individual health. Parents with higher levels of human capital and, thus, educational attainment are better able to provide their children with resources needed to foster cognitive development. We hypothesize that the effect holds on the macro-level as well. The national stock of human capital is typically measured by the distribution of levels of educational attainment in the population (Wößmann, 2003). So, the hypothesis is that

H7.1: The portion of the labor force that has completed tertiary education should positively correlated with the average performance at the IMO.

Input-based schooling indicators are the gross enrollment rate in secondary education and expenditures per student in secondary education measured as a percentage of GDP per capita. The gross enrollment rate in secondary education has a direct effect on the size of the effectively participating population since enrollment in an institute of secondary education is a prerequisite for participating at the IMO. While secondary education lays the basis of scientific knowledge and contributes to skill formation, a large relative number of adolescents in secondary education may also have an adverse impact if the level of resources is constant. Nevertheless, we expect that

H7.2: Higher enrollment rates lead to a better performance at the IMO by broadening the effectively participating population.

Still, resources dedicated to secondary education should influence school quality and, consequently, performance at the IMO. The level of resources is associated with the student-teacher ratio and the education and the experience of teachers. The indicator may also say something about the weight a society puts on education. So, we hypothesize that

H7.3: A higher level of national expenditures on secondary education is expected to spur performance at the IMO.

However, the link between resources and general schooling quality is controversial at best. Many empirical studies have failed to identify an association between input measures and schooling quality (Hanushek, 2003). Therefore, our analysis contributes to the ongoing debate and enhances understanding of the issue by investigating the role of schooling expenditures in explaining inter-country variation in the performance of IMO participants.

Closely related to school resources is the extent to which performance at the IMO relates to measures of schooling quantity and quality. Quantity and quality of schooling refer to the ability of the educational institutions to endow as many students as possible with necessary competencies in reading, mathematical, and scientific literacy for later life and career. PISA is an international large scale attempt to assess differences in the quality of national schooling systems. The waves from 2003 and 2006 contain temporally comparable assessments of students' achievement in mathematics (OECD, 2007). If achievement at the IMO is rooted in the general quality of the underlying schooling system, we would expect a positive correlation of PISA scores in the field of mathematics with the performance at the IMO. A potentially more relevant measure is the average PISA score of students in the 95th percentile of the proficiency distribution, since it directly relates to achievement of the most talented students in mathematics.

H7.4: PISA scores should be positively correlated with the performance of participants at the IMO.

We use the percentage of population aged 15-24 years who have completed secondary education as a measure of a nation's schooling quantity. While mean PISA scores refer to school quality, percentage of secondary education provides general information about the overall coverage of general education. Higher coverage should positively influence the outcome at the IMO by broadening the potential talent pool.

H7.5: A higher percentage of population aged 15-24 years who have completed secondary education should positively influence the national performance at the IMO.

4 Methodology and Data

In this section, we outline the empirical methodology (1) and describe the dataset (2).

4.1 Methodology

The average score of Olympians, T_{it}, from country i in year t can be expressed as a linear function of the outlined explanatory variables, such that

$$T_{it} = \beta_0 + X_{it}\beta + Z_i\gamma + a_i + m_t + \varepsilon_{it}, \quad (1)$$

where the i subscript denotes countries, the t subscript denotes years, X_{it} and Z_i represents the covariates, a_i captures unobserved, time-invariant country effects, m_t denotes common period shocks (e. g., difficulty of examination problems), and e_{it} is an i. i. d. error.

The aim is to explain the influence of varying macro-level conditions on the average performance at the IMO. However, macro-conditions change slowly, so it may take some time before changes translate into a shift in the expected value of the explanatory variable. For instance, growth of GDP per capita will not immediately influence the resources dedicated to education, so any resulting change in the average IMO performance will occur after an even longer period. In fact, short- and long-run effects might even go in opposite directions. For example, a steep rise in the number of 15-19-year-olds could, in the short run, constrain resources and hamper individual performance at the IMO, while, in the long run, after adjustments in the schooling system haven taken place, a large number of 15-19-year-olds might increase the pool of available talents and thus the performance at the IMO. In other words, a large population might be beneficial to performance at the IMO, while rapid population growth may be detrimental. Therefore, short- and long-run effects should be analyzed separately.

Baltagi and Griffin (1984) and Pirotte (1999) showed that the between estimator provides a reasonably good estimator of long-run effects, and a fixed-effects specification provides an estimator for effects in the short run. Therefore, the long-run associations are estimated based on:

$$\bar{T}_{i.} = \beta_0^L + \bar{X}_{i.}\beta^L + Z_i\gamma + \bar{\varepsilon}_{i.} \tag{2}$$

where $\bar{T}_{i.}$, $\bar{X}_{i.}$, and $\bar{\varepsilon}_{i.}$ denote country-specific time averages over the observation period and are derived by a *between* transformation of the data. The coefficients β^L and γ give the expected difference in the average achievement between countries if they differ in $\bar{X}_{i.}$ or Z_i by 1 unit, and are interpreted as long run associations between the dependent variable and the regressors. Equation (2) can be estimated by OLS. The estimator will be consistent if the transformed explanatory variables do not correlate with the unobserved country effects a_i (Wooldridge, 2002).

In a second step, we analyze short-run effects by determining whether variations of the macro-variables contribute to changes in the average achievement *within* countries. Hence, equation (1) is differentiated, which yields,

$$T_{it} - T_{is} = (X_{it} - X_{is})\beta^S + m_t + (\varepsilon_{it} - \varepsilon_{is}) \tag{3}$$

The coefficient β^S gives the expected change in the average score of country i if X_{it} varies by 1 unit within this country, this change can be interpreted as a short-run impact. Differentiating eliminates the unobserved country effects and, thus, the potential bias from omitted variables. It also enables causal inference (if omitted variables are indeed time-invariant) and, since country-specific characteristics of the training and preparation procedure are captured, changes in expected achievement relate to variations in the underlying talent pool. Estimation of equation (3) is solely based on time series variation within panels, so only the effect of time-varying variables is estimable. In the case of the fixed-effects estimation, the subtrahends are replaced by the country means from equation (2), so fixed effects will regress deviations from the mean performance on deviations from the means of the explanatory variables.

Because the macro time-series change very slowly estimates may be imprecise. It may take several years before a shift is sufficiently large to reveal a measurable impact on the dependent variable. Therefore, we also look at the dynamics in average achievement over longer time periods by replacing t and s with sufficiently distant years (long-difference estimator). For more information and a comprehensive discussion of related issues, see Nichols (2007).

A simple way to test for differences between long- and short-run coefficients is to combine equations (2) and (3) and regress the yearly performance jointly on the means (long-run) and deviations from the means (short-run) of the explanatory variables. The regression can be estimated by random-effects. Generally, the random-effects estimator is a weighted average of the between- and fixed-effects results. The underlying assumption is the equality of short- and long-run effects, which assumption allows between and within variation to be combined. Like the between estimator, random effects will be inconsistent if unobserved country effects correlate with the explanatory variables. It seems plausible that country effects are correlated with the level of development (GDP per capita), however, the chosen specification helps to overcome both limitations. First, the means of the explanatory variables capture country-specific differences to a certain extent (Mundlak, 1978). Second, the specification takes possible differences between the short-run and long-run coefficients explicitly into account.

4.2 Data

The International Mathematical Olympiad (IMO) is the oldest science Olympiad. Since it began in Romania in 1959, it has been held annually (except in 1980) in a variety of host countries. Initially, invitations were restricted to socialistic European countries but, over the course of time, countries from all over the world joined the IMO. Currently, teams from more than 100 countries participate.

Since 1983, national teams have consisted of up to 6 contestants, who must be under the age of 20 and not registered at an institution of tertiary education, one team leader, one deputy leader and observers. Participants in the IMO are among the most talented young mathematics scholars in their country.

The competition takes place on two consecutive days and, on each day, the participants have 4.5 hours to solve three problems. Solutions are awarded 0 to 7 points, so the maximum score in the overall exam is 42 points. Gold-medals winners score, on average more than 33 points, although the threshold changes slightly with the overall result of the Olympiad. The average score is 14 points, or 33% of the maximum score.

The selection and preparation process of IMO contestants varies from country to country. Usually, selection takes place via different rounds of regional and national mathematical competitions to recruit a pool of possible candidates, which are weeded out until only the top six candidates are left. The final contestants undertake special training courses to prepare for the IMO, but duration and intensity of the training is highly country-specific.

The dependent variable in the analysis is the average score of the participants from country i at time t. Data on performance was collected from the official website of the IMO Advisory Board (www.imo-official.org). Figure 1 displays the 5 countries with the highest and lowest average scores between 1993 and 2006. The figure shows the high cross-sectional variation in the dependent variable and the high heterogeneity between countries with similar levels of performance.

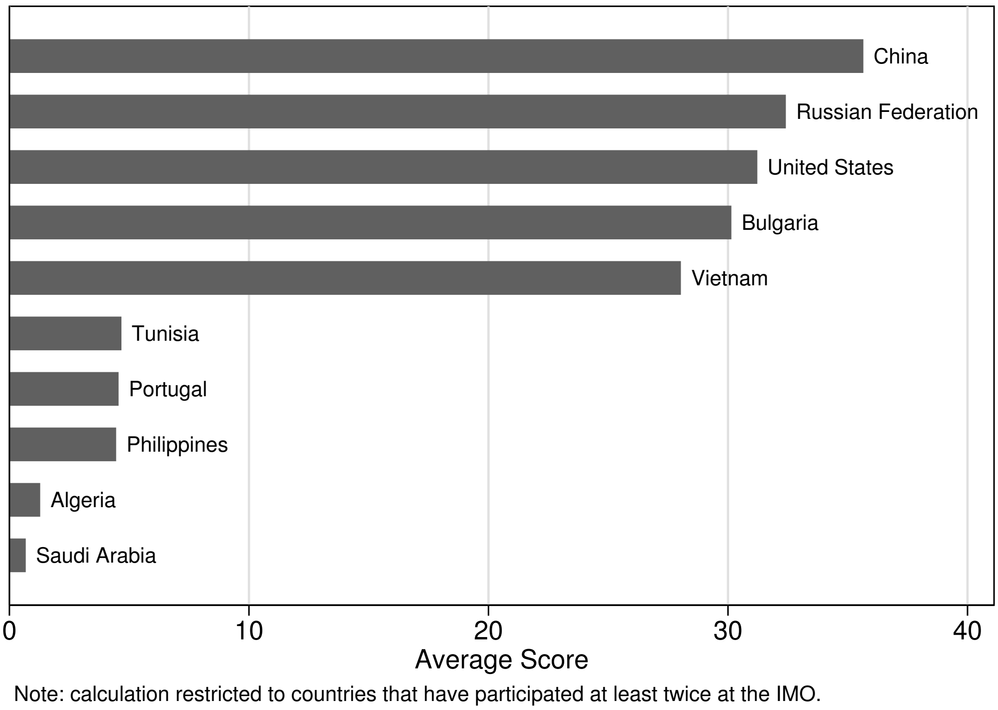

Figure 1: Average Score of Participants, Highest- and Lowest-scoring Countries, 1993-2006.

Data on real GDP per capita, the degree of urbanization, number of scientific articles, the gross enrollment rate in secondary education, the spread of tertiary education in the labor force, and expenditure per student in secondary education as a percentage of GDP per capita was taken from the World Development Indicators 2008. Information on population size, age structure and birth rates was taken from the International Database of the U. S. Census Bureau. The number of times countries have participated at the IMO and the number of times a country hosted the IMO were calculated from information at the website of the IMO Advisory Board. The classification of countries with communist regimes and former communist regimes follows Leeds and Leeds (2009). PISA scores in mathematics in 2003 and 2006 were taken from OECD (2007, 2004). Information about the distribution of educational attainment by age came from a publicly available database that is the joint work of the IIASA and VID; it consists of back and forward projections of educational attainment by age and sex for 120 countries (for more information see K.C. et al., 2008; Lutz et al., 2007).

The dataset consists of 1124 observations from 99 countries that have participated at least twice during the period 1993-2006. However, limited data availability regarding the explanatory variables restricts the actual number of usable observations. The average length of the observation period is 11.4 years. Information on GDP per capita PPP is in constant 2005 international $; total population, the crude birth rate, the percentage of adolescents aged 15-19, and the percentage of urban population is available for most of the countries over the whole period, with the exception of Taiwan, Lichtenstein, Cuba, and Turkmenistan. Time-series data on the number of scientific articles is available for most countries, but information on the schooling system and human capital endowment is generally sparse. Enrollment rates and expenditures per student are available only for a subset of countries from 1998 onwards, and information on the distribution of educational attainment is restricted to 1995, 2000, and 2005. PISA information is limited to 88 observations in the dataset, representing 52 countries in 2003 and 2006.

Therefore we have constructed two sets of explanatory variables. The *basic* set consists of GDP per capita PPP in constant 2005 international $, total population, the percentage of adolescents aged 15-19, the crude birth rate, the percentage of urban population, the number of scientific and technical journal articles, the first year of participation at the IMO, and the number of times a country has hosted the IMO. This basic variable set allows us to identify the relationships between IMO performance and fundamental indicators of the potential talent pool, that is, effective participation, available resources, political regimes, and the quality of IMO-related national institutions. Thus, the basic variable set represents fundamental macro-level conditions. Table 1 displays summary statistics of the basic variable set and the depend variable. The overall variation of the variables is decomposed into inter-country (*between*) and intra-country (*within*) variation.

The second, *extended* set of explanatory variables includes additional information on the prevalence of tertiary education in the adult population and details on the endowment, spread and quality of secondary education. More specifically, we use data on the gross enrollment rate in secondary education,

Table 1: Summary Statistics - Basic Variable Set

Variable		(1) Mean	(2) Std. Dev.	(3) Observations
Average score	overall	14.230	8.875	N = 1124
	between		8.234	n = 99
	within		3.905	T-bar = 11.354
ln Population	overall	16.398	1.639	N = 1116
	between		1.703	n = 99
	within		0.043	T-bar = 11.273
ln Real GDP p.c.	overall	9.246	0.997	N = 1071
	between		1.031	n = 94
	within		0.160	T-bar = 11.394
Percentage of Population 15-19	overall	8.193	1.808	N = 1116
	between		1.824	n = 99
	within		0.528	T-bar = 11.273
Crude Birth Rate	overall	15.137	6.190	N = 1084
	between		7.265	n = 98
	within		1.486	T-bar = 11.061
Degree of Urbanization	overall	65.836	19.590	N = 1098
	between		20.570	n = 96
	within		1.427	T-bar = 11.438
Articles (in 1000)	overall	8.741	25.041	N = 894
	between		22.837	n = 85
	within		1.800	T = 10.518
ln Participations	overall	2.434	0.833	N = 1124
	between		0.922	n = 99
	within		0.463	T-bar = 11.356
Hosting Activities	overall	0.535	1.021	N = 1124
	between		0.938	n = 99
	within		0.170	T-bar = 11.356
Communist	overall	0.036	0.185	N = 1124
	between		0.172	n = 99
	within		0.000	T-bar = 11.356
Former Communist	overall	0.020	0.139	N = 1124
	between		0.141	n = 99
	within		0.000	T-bar = 11.356

expenditure per student in secondary education as a percentage of GDP per capita, the percentage of persons in the labor force with tertiary education, the percentage of population aged 15-24 with completed secondary education, mean PISA scores in mathematics, and PISA scores from students in the 95th percentile. Compared to the basic variable set, the extended set's number of observations is much smaller. In particular PISA scores are available only for a subset of countries. Summary statistics are reported in Table 2.

Table 2: Summary Statistics - Extended Set

Variable		(1) Mean	(2) Std. Dev.	(3) Observations
Enrollment Secondary Education	overall	90.423	21.356	N = 575
	between		22.002	n = 87
	within		5.053	T-bar = 6.6092
Expenditures Secondary Education	overall	21.686	9.275	N = 353
	between		9.981	n = 75
	within		2.610	T-bar = 4.70667
Secondary Education 15-24	overall	72.820	17.629	N = 204
	between		18.062	n = 78
	within		3.344	T = 2.61538
Tertiary Education Labor Force	overall	23.326	11.511	N = 534
	between		11.184	n = 73
	within		5.134	T = 7.31507
Mean PISA score	overall	483.464	51.075	N = 88
	between		55.368	n = 52
	within		3.299	T-bar = 1.69231
PISA score 95th percentile	overall	632.195	49.870	N = 88
	between		54.568	n = 52
	within		4.296	T-bar = 1.69231

Both variable sets have relatively strong variations between countries compared to the variations within single countries over time. The dependent variable reveals the highest relative level of variation within countries. Persistence of time series can be measured by the ratio of between standard deviation to within standard deviation (b/w ratio). By contrast, the explanatory variables are characterized by relatively time-persistent patterns. Most of the variables describe institutions or other attributes, like the level of development and population size, that change only slowly over time.

The macro-level variables employed can be safely assumed to be exogenous. Reverse causation from the performance at the IMO on measures like population size, GDP per capita, degree of urbanization, youth percentage or number of articles is unlikely; only the number of hosting activities and the number of times countries have participated at the IMO might fail the assumption of strict exogeneity. A contemporaneous performance shock could lead to more (less) resources and more (less) involvement in the IMO in successive periods, so we will test whether this is an issue.

However, unobserved heterogeneity is likely to be a problem. If not addressed, it will lead to biased estimates because important but unobservable determinants may correlate with both the covariates and the dependent variable. The panel structure of the data set allows time-invariant unobserved heterogeneity in the cross-sectional units to be controlled. Unobserved heterogeneity is the result of omitted (unobservable) variables like the training of contestants in advance of the IMO and/or the overall quality of mathematical education in the nation's schools.

5 Results

This section presents the results from the various regressions. Since neither heteroscedasticity nor autocorrelation can be ruled out, inferences in all specifications are based on fully robust standard errors. The empirical analysis is structured as follows. First, we look at the general trends in average performance. Second, long- and short-run influences of basic covariates are estimated and compared. Since macro-conditions change slowly, the analysis also considers differences in the covariates between distant time periods. After exploring these dynamics, we use the extended variable set to examine the role of the national human capital stock and the schooling system in determining performance at the IMO, with a focus on cross-sectional variation. Finally, the findings from the International Mathematical Olympiad are cross-checked with estimates from a cross-section of country results at the International Physics Olympiads.

5.1 Dynamics of Average Scores

Consideration of the dynamics of the average performance of Olympians raises questions related to whether there are specific trends within countries in the average performance and whether there has been a general upward movement in the results as countries have grown richer and the population has increased. To answer these questions we use equation (3) and regress the average performance on a set of year dummies, with 2000 as reference year, conditional on the unobserved country effects. Fitted values from the regression are plotted in Figure 2.

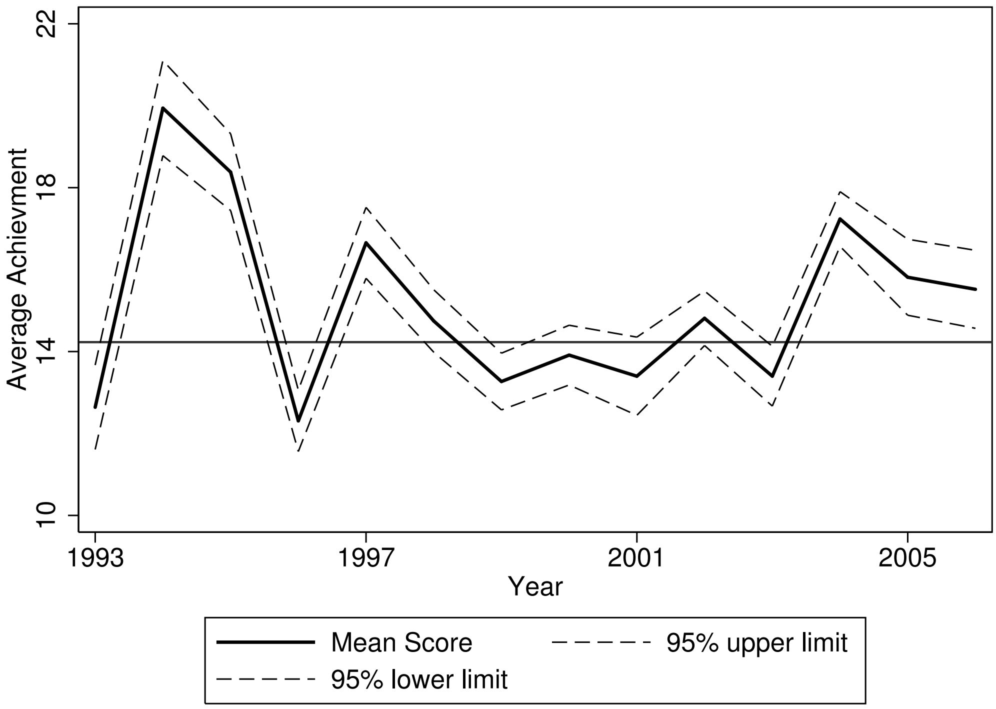

Figure 2: Country level Dynamics in the Average Performance at the IMO

Most noticeable in Figure 2 is the peak in achievement during the mid-1990s; in 1994 and 1995, performance was well above the sample average. In general, one can surmise that there are significant variations in the average achievement over the years, but these variations follow no specific trend. Either participants' performances do not systematically improve (or worsen) or the organizers of the IMO adapt to changes in skill and knowledge formation by developing exams that produce similar aggregate score distributions over the years in order to keep the mean score stable. In order to avoid training and preparation of solutions for certain problems and to represent a constant challenge, the IMO exams become in fact more difficult over time.

In the next section, we analyze to what extent macro-level differences between countries lead to shifts in expected average IMO performance and whether macro-level variations within countries contribute to changes in the performance of their Olympians and to departures from the general trend.

5.2 Performance at the IMO in the Long Run and the Short Run

In the following, we present results from the estimation of Equations (2) and (3) on the basic variable set. The variable names and their descriptions are found in Table 3. The analysis is restricted to countries with at least two complete observations.

Table 3: Variables and Description

Variable	Description
T_{it}	Average score of participants from country i in year t
$\ln POP_{it}$	ln population of country i in year t
$\ln GDPPC_{it}$	ln GDP per capita PPP in constant 2005 international $ of country i in year t
$SHARE1519_{it}$	Percentage of population ages 15 to 19 in country i in year t
CBR_{it}	Crude birth rate per 1000 population in country i in year t
$URBAN_{it}$	Percentage of urban population in country i in year t
$ARTICLE_{it}$	Number of scientific and technical journal articles (in 1000s) in the fields of physics, biology, chemistry, mathematics, clinical medicine, biomedical research, engineering and technology, and earth and space sciences
$\ln PART_{it}$	ln number of times of country i has participated at IMO till year t
$HOSTING_{it}$	Number of times country i has hosted the IMO up to year t
COM_i	Dummy for communist regime in country i
$FCOM_i$	Dummy for former communist regime i

To compare the results with findings in the previous literature, we also run estimations with $\ln POP_{it}$ and $\ln GDPPC_{it}$ as the only regressors. Estimation output is displayed in Table 4. Column (1) and (2) report inter-country findings based on equation (2). Similar to findings in the extant literature, participants from larger countries have performed better on the average than Olympians from smaller countries. A 1 % larger population implies a difference of 1.6 - 2.5 scores in expected performance depending on the included variable. Thus, hypothesis H1.1 is confirmed.

However, in contrast to studies on Olympic success (compare, e. g., Bernard and Busse, 2004; Hoffmann et al., 2002a) and in contradiction to hypothesis H2, there is no evidence that participants from wealthier countries achieve higher averages scores. Even in the long run, GDP per capita has no influence on the expected average achievement.

Adding the additional covariates gives a more complex picture and sharply increases the explanatory power of the regression. Around 70% of the cross-sectional variation in average performance is explained. The estimates reveal correlation patterns that confirm the hypotheses H1.1, H1.2, H3.1, H4, H5, and H6.1. Estimated coefficients of population size, crude birth rate, urban population, published articles, times hosting the IMO, and the influence of communist regimes have the expected signs such that larger populations, a lower birth rate, a higher degree of urbanization, larger academic markets (measured by published articles), the presence of communist regimes and the number of times a country has hosted the IMO are positively associated with participants' average performance at the IMO. There is also a positive association between a communist past and average performance, suggesting that beneficial institutions from the past still enhance the achievement of Olympians from those countries. Only the number of times countries have participated and the percentage of 15-19-year-olds in the population show no significant correlations with the IMO performance.

Therefore, overall, there is clear evidence that contestants from countries with larger effectively participating populations and beneficial macro-conditions, like bigger markets for academic research and lower fertility perform better on average at the IMO than do Olympiads from other environments.

Findings from the between-estimation shed light on determinants that explain inter-country differences in the IMO performance and suggest the long-run associations between the variables. Still, findings must be interpreted with care since it is likely that unobservable country characteristics influence the dependent variable and, at the same time, correlate with the covariates, a situation which would render the between-estimator inconsistent. Furthermore, the between-estimator cannot determine whether changes in macro-conditions within a country will have an effect on the expected performance of the country's Olympians.

In the next step, equation (3) is estimated by linear fixed-effects regression, with results reported in column (3) and (4) of Table 4. Generally, the F-test strongly rejects the null hypothesis of no country-specific effects. Because the influence of time-invariant variables is not estimable, the dummies for current or former communist regime drop out of the regression. Coefficients indicate the influence of intra-country changes and can be interpreted as short run effects.

Table 4: Regression Results - Equation (2) and (3)

	Long Run (between)		Short Run (within)		Both
COEFFICIENT	(1)	(2)	(3)	(4)	(5)

Avg. $\ln POP_i$	2.518***	1.590***			1.713***
	(0.49)	(0.53)			(0.52)
Avg. $\ln GDPPC_i$	0.916	-1.833			-1.329
	(0.78)	(1.11)			(1.14)
Avg. $SHARE1519_i$		1.166			1.272
		(1.16)			(1.09)
Avg. CBR_i		-0.747***			-0.751***
		(0.24)			(0.24)
Avg. $URBAN_i$		0.114***			0.105***
		(0.033)			(0.033)
Avg. $ARTICLE_i$		0.0610***			0.0638***
		(0.012)			(0.012)
Avg. $\ln PART_i$		1.304*			0.891
		(0.78)			(0.89)
Avg. $HOSTING_i$		2.525***			2.603***
		(0.54)			(0.55)
COM_i		14.01***			14.43***
		(2.40)			(2.53)
$FCOM_i$		3.328**			4.225***
		(1.58)			(1.49)
$\Delta \ln POP_{it}$			0.810	-6.890	-6.640
			(6.37)	(8.07)	(7.87)
$\Delta \ln GDPPC_{it}$			-1.046	-1.014	-1.048
			(1.94)	(1.91)	(1.92)
$\Delta SHARE1519_{it}$				-0.109	-0.106
				(0.53)	(0.53)
ΔCBR_{it}				-0.203	-0.222
				(0.21)	(0.21)
$\Delta URBAN_{it}$				0.346*	0.343**
				(0.18)	(0.17)
$\Delta ARTICLE_{it}$				0.0821	0.0780
				(0.064)	(0.064)
$\Delta \ln PART_{it}$				2.726***	2.638***
				(0.90)	(0.90)
$\Delta HOSTING_{it}$				1.262*	1.222*
				(0.68)	(0.68)
Constant	-35.69***	-6.157	10.25	110.2	-13.71
	(11.4)	(20.0)	(114)	(134)	(19.1)
Period Dummies			yes	yes	yes
Observations	74	74	954	833	833
R-squared	0.24	0.71	0.28	0.33	0.62
Number of Countries			74	74	74

Note: Heteroscedasticity and autocorrelation robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

In contrast to the between-estimation results, almost none of the coefficients are significantly different from zero. It appears that, while levels of aggregate measures correlate with the dependent variable, changes in the variables are not accompanied by a shift in the expected performance; in other words, it is not possible to identify a causal link between the regressors and the dependent variable. Only the log of the number of participations has a clearly significant impact on the dependent variable. Hence, after controlling for unobserved country characteristics and common period effects, we see an improvement in the average performance as the number of participations increases. The degree of urbanization and the number of hosting activities is still significant, but only at the 10% level.

A Wald test on the joint significance of the other covariates is rejected. Thus, we find no robust evidence that changes in macro-measures instantaneously influence the average performance of Olympians at the IMO. The formation of skills, knowledge and, more specifically, talent appears to be unaffected in the short run by changes in macro-conditions. The only exception is a positive but diminishing increase in the performance of Olympians with the number of a country's participations. Still, the variable may violate the assumption of strict exogeneity; for example, it is possible that a positive current performance shock has a feedback effect on future participation of countries at the IMO.

Wooldridge (2002) outlined a simple test of strict exogeneity wherein one-period-lead values ($t + 1$) of the suspicious variables are plugged into the fixed-effects regression. Under strict exogeneity, the coefficients of these lead values will be jointly equal zero. In our case, the one-period-lead value of the number of times countries have participated at the IMO is included. The estimates indicate that a violation of strict exogeneity cannot be decisively ruled out. The coefficient of the lead value is significant at the 10% level with a t-value of 1.87, so feedback effects of current performance shocks on future participation are likely. Generally, the estimator is inconsistent in this case.

Since the regression includes a set of year dummies, identification of the influence of the number of participations comes from countries with gaps in the participation history, but the gaps are not randomly distributed across nations and may just indicate a lack of funding and/ or a lack of institutions able to organize the national selection and training process. This gap may explain the feedback effects from current performance shocks on future participation. Once year dummies are dropped from the regression, the number of participations has no significant influence anymore. Thus, the positive coefficient of the log number of participations in the fixed-effects regression reflects a correlation between performance and participation that stems from irregularly participating countries.

So far, the determinants have been able to explain the between-variation in the dependent variable very well, although they fail to explain changes in the average performance over time. However, there is still no clear evidence of whether the difference in predictive power also indicates a difference in the long- and short-run effects. Column (5) of Table 4 summarizes the previous findings; it displays the estimates from a joint random-effects regression of between and within influences, which allows us to test for differences between the coefficients. A Wald test on the joint similarity of the within and between coefficients of the time and country varying variables provides information about the validity of random effects.

A similar test procedure was proposed by Mundlak (1978) and Wooldridge (2002) as a substitute for the classical Hausman test. Mundlak replaced the time-demeaned terms with the untransformed values, while Wooldridge suggested substituting the mean terms with the untransformed values. As before, inter-country differences in the explanatory variables are strongly associated with differences in the expected average achievement. The only exemptions are $\ln GDPPC_i$, $SHARE1519_i$, and $\ln PART_i$. However, again, there is almost no evidence that an intra-country change alters the performance at the IMO. Only changes in the degree of urbanization have a significant positive impact, while the number of participations correlates with performance because irregular participation and average performance are mutually dependent.

Do between and within coefficients differ? A joint test rejects the null hypothesis of similarity only at the 10% level of significance (p-value=0.052), and comparing only the corresponding pairs of coefficients reveals no significant differences. Therefore, from a statistical point of view, there is no clear evidence of differences between the long run and the short run. The problem is that the time series of the explanatory variables are highly persistent, which leads to imprecise estimates from the fixed-effects method. One could argue that variations in macro-level determinants will not instantaneously affect individual outcomes, but that they take time to build up a measurable impact. To test this, we consider the influence of changes over longer time periods.

To learn more about the dynamics of how changes in macro-variables lead to a change in the average performance at the IMO, we use a first difference approach but restrict estimation to data from differently distant years, i. e., the beginning and end of the observation period with and without intermediate periods. As outlined in Section 4.1, this long difference method, like the fixed-effects method, eliminates unobserved heterogeneity. Results are displayed in Table 5. Columns (1) to (4) compare findings from regressions on long run differences in the variables. Column (5) presents additionally estimates from regressions over shorter time periods.

The results can be interpreted as follows. In contrast to the fixed-effects method, the long difference approach results in some of the covariates' showing a significant effect. Based on the results, population dynamics could play a role in that a reduction in fertility may lead to an improved performance at the IMO. In all long difference estimations, the effect is statistically significant at least at the 10% level. Furthermore, changes in income show a significant negative effect in two cases, and a change in the number of published articles has a positive impact in one case.

Further interpretation of the results reveals that, findings are not robust to the choice of investigation periods. With 2005 as the end point, the regression model works fairly well, with several coefficients significantly different from zero. The picture changes if 2004 is chosen instead: in this case, only changes in fertility have a weak statistical influence. Finally, the inclusion of additional intermediate periods does not improve the results. As was the case for the fixed-effect approach, none of the included macro-variables appears to be associated with changes in the average achievement at the IMO. In short, there are some indications that changes over a longer period can have an effect on the outcome variable, but it seems that the measured effects are sensitive to the included countries and investigated periods.

Overall, the results provide mixed support for the hypotheses. Cross-sectional analysis confirms most of the previous considerations: hypotheses H1.1, H1.2, H3.1, H4, H5, and H6.1 are supported and coefficients of the effectively participating population, the birth rate, the academic job markets, communist (current and past) regimes, and measures of national institutions (hosting history) have the expected signs and are highly significant. Cohort size of adolescents and the number of times countries have participated at the IMO are found to have no influence on average national performance since the coefficient of the number of 15-19 year olds in the population and the log number of participation are insignificant.

In contrast to the literature on the Olympic Games, the level of GDP per capita is not associated with the outcome at the IMO. In other words, it seems that economic activities and average income have no beneficial influence on the average achievement of Math Olympians. Talent formation in mathematics could depend more on human capital and the socio-cultural background.

However, the hypotheses fail to explain changes in the average intra-country performance. Conditioned on country effects, there is no evidence that short-run

Table 5: Long Difference Results - Equation (3)

COEFFICIENT	(1) 1993/ 2005	(2) 1993/ 2004	(3) 1994/ 2005	(4) 1994/ 2004	(5) 1993/ 1999/ 2005
$\Delta \ln POP_{it}$	-27.15**	-5.730	-32.93*	-8.245	-11.35
	(10.2)	(11.3)	(17.3)	(14.5)	(11.7)
$\Delta \ln GDPPC_{it}$	-9.312**	-3.958	-11.18**	-2.479	-1.505
	(3.99)	(4.36)	(4.91)	(5.35)	(2.55)
$\Delta SHARE1519_{it}$	-1.295*	-0.343	-1.416*	-0.970	0.322
	(0.68)	(0.74)	(0.72)	(0.78)	(0.73)
ΔCBR_{it}	-1.079***	-0.753*	-1.210**	-0.878*	0.0483
	(0.37)	(0.38)	(0.50)	(0.51)	(0.26)
$\Delta URBAN_{it}$	0.144	0.173	0.0542	0.0974	0.169
	(0.27)	(0.30)	(0.36)	(0.27)	(0.31)
$\Delta ARTICLE_{it}$	0.152	-0.116	0.379***	0.0898	0.162
	(0.16)	(0.20)	(0.11)	(0.11)	(0.15)
$\Delta \ln PART_{it}$	2.371*	2.183	5.379***	4.002**	1.436
	(1.24)	(1.33)	(1.88)	(1.80)	(1.31)
$\Delta HOSTING_{it}$	4.208*	4.098	-0.328	-0.560	2.883*
	(2.26)	(2.46)	(2.87)	(2.34)	(1.64)
Constant	0.308	0.714	-7.881***	-7.674***	-0.435
	(2.36)	(2.45)	(2.03)	(1.77)	(1.58)
Period Dummies					yes
Observations	53	53	55	55	118
R-squared	0.36	0.21	0.35	0.21	0.08

changes in macro conditions directly affect the dependent variable. The measured influence of urbanization depends on the methodology and the subsample (cf. findings in Table 4 and 5). Only the number of participations is clearly associated with the outcome at the Olympiad, but the estimate is potentially biased because of feedback effects and the selectivity in the distribution of the variable.

However, there is some weak evidence that shifts that occur over longer time periods affect performance at the IMO. Specifically, a continued reduction of the crude birth rate seems to spur performance, in accordance with hypothesis H3.1. Along the lines of the quantity-quality trade off, a decrease in birth rates might be accompanied by an increasing demand for child quality and lead to increased spending on skill and knowledge formation. Other within findings are highly sensitive to the analyzed subsample. Therefore, while there are highly significant cross-sectional associations between the covariates and the outcome variable in support of the hypotheses, we find no robust statistical evidence that changes in macro-conditions have a causal effect on the average achievement of Olympians.

Still, the within analysis is hampered by the high persistence of macro time-series; this persistence leads to imprecise estimates and the risk of accepting wrong null hypotheses. The between-estimator, in turn, is biased if regressors and unobserved heterogeneity are correlated.

One way to reduce the omitted variable bias in the between regression is to include additional country characteristics. Talent formation should also depend on available human capital, and the endowment and quality of the schooling system, which are captured by the indicators of the extended variable set. Therefor, in the next section, we will analyze the role of these measures in more detail.

5.3 Human Capital, the School System and the Performance at the IMO

It is not a priori clear from the data whether top-ranked countries are, on the average, more successful in talent formation or whether they rely on better preparation processes for IMO contestant, so the question arises: Do higher average scores result from talent formation in general or from talent selection and preparation? The aim of this section is to shed light on the relationship between, on one hand, indicators of human capital and schooling inputs and outputs and, on the other, the average performance at the IMO. If the results are rooted in the schooling system and related to the general level of human capital, there will be a statistically measurable influence.

Because of slow changes in educational institutions, the analysis is focused on inter-country differences. Inferences are based on the regression function (2) and data is restricted to the years after 1997. The extended variable set is describes in Table 2.

A problem for the regression arises from the high level of collinearity between the mean PISA score and the PISA score of the 95th percentile; the correlation coefficient varies between 0.98 (2003) and 0.97 (2006). Therefore, the mean PISA score was dropped from the regressions. Table 6 summarizes the findings of the regressions, with output restricted to the variables of interest.

Table 6: Role of schooling input and output indicators, 1998-2006

COEFFICIENT	(1) Input	(2) Output	(3) Jointly	(4) Full Input	(5) Full Output
Avg. $ENROLL_i$	-0.0304		-0.124**	-0.0968**	
	(0.040)		(0.052)	(0.044)	
Avg. $SECXPD_i$	-0.0769		-0.295**	-0.172*	
	(0.075)		(0.14)	(0.092)	
Avg. $TERT\_TLF_i$	0.159**		0.140	0.156*	
	(0.063)		(0.099)	(0.092)	
Avg. $SEC1524_i$		0.0697	0.104		0.0382
		(0.061)	(0.069)		(0.057)
Avg. $PISA95th_i$		0.0738***	0.0762***		0.0226
		(0.025)	(0.020)		(0.034)
Basic Covariates	yes	yes	yes		
Full Covariates				yes	yes
Observations	61	46	43	52	43
R-squared	0.36	0.43	0.47	0.66	0.71

Note: Basic Covariates are avg. $\ln POP_i$ and avg. $\ln GDPPC_i$. Full Covariates also contain avg. CBR_i, avg. $URBAN_i$, avg. $ARTICLE_i$, Avg. $HOSTING_i$, COM_i, and $FCOM_i$. Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1.

The regressions also include either $\ln POP_i$ and $\ln GDPPC_i$, column (1) – (3), or the full set of basic covariates, column (4) and (5), in order to take general correlations between the independent variables into account. To save degrees of freedom, the average number of participations and the average percentage of 15-19-year-olds in populations have been dropped.

The average PISA scores of top students in mathematics are positively associated with the IMO performance. A difference of one standard deviation in PISA scores between countries is associated with an expected difference at the IMO of around 4 points (approximately 10% in scoring). The percentage of the population aged 15-24 that has completed secondary education is found to have no influence. The inclusion of measures of schooling input and human capital does not alter the results. Again, PISA scores are positively correlated with the performance of Olympians, while the coefficient of percentage of population ages 15-24 who has completed secondary education is not significantly different from zero. The results support hypothesis H7.4, while hypothesis H7.5 is not supported. However, if IMO performance is regressed on schooling output and the full set of basic covariates, neither measure shows a significant influence.

Therefore, there is some evidence that school quality matters, while school quantity is not associated with the average achievement at the IMO. But with the inclusion of further determinants, the association between school quality and results at the IMO vanish, and it seems that the relationship is mediated by other variables in the regression function.

Of the measures of human capital and schooling input, only the incidence of tertiary education in the labor force correlates with the average performance at the IMO in the basic specification. Olympians from countries with better trained labor forces perform better on the average, which result supports hypothesis H7.1. Similar to inter-country studies of school quality, there is no evidence that expenditures per student in secondary education matter in terms of schooling output, nor does the coverage of secondary education show significant association with the performance at the IMO.

However, the picture changes if input and output measures are jointly estimated. The enrollment rate and expenditure per student in secondary education have a negative influence and the incidence of tertiary education is insignificant. In other words, holding quality constant, students from countries with a higher coverage of secondary education and higher spending per student perform worse at the IMO. The negative influence of spending on school output is not a unique finding in the literature. For example, Wößmann (2001) reported similar results for students' mathematical and scientific achievement based on TIMSS data. From the figures, one could conclude that countries with more efficient schooling systems (less input for the given output) and that focus on gifted students (the given quality but for fewer students) have advantages at the IMO.

When the average IMO performance is regressed on the inputs measures and the full set of basic covariates, the gross enrollment rate and the expenditures per student in secondary education, expressed as percentage of GDP per capita are again negatively associated with the average performance at the IMO, although the influence of spending is only significant at the 10% level. Therefore, there is evidence that, after taking the other covariates into account, Olympians from countries with relatively poorly endowed secondary schooling systems perform better at the International Mathematical Olympiad. Thus, hypotheses on schooling inputs, H7.2 and H7.3, are not supported. Instead of positive relationships, we get negative estimates

Summing up the findings of this section, one can derive two conclusions. First, after controlling for general determinants of performance at the IMO, schooling inputs influence inter-country differences in IMO performance, although not as hypothesized since the direction of the effect is negative. However, as expected, the incidence of tertiary education in the labor force has a positive relationship with the IMO performance. The significance of the finding is sensitive to the included set of covariates, but the point estimators are similar in all specifications.

The second conclusion from this section is that mathematical achievement of top students, as measured by mean PISA scores of students in the 95th percentile, correlates with IMO performance. However, once further determinants are taken into account, the estimated influence disappears. Since the quantity of secondary education generally plays no role, hypothesis H7.5 receives no support from the data. Overall, there is evidence that performance at the IMO is not solely the result of IMO-related training and preparation, but is rooted in the level of education of the parental population and schooling inputs. The influence of schooling quality is ambiguous; in the basic model, there is a statistical influence, but it is not reproduced in the enriched specification.

5.4 Determinants of Success at the International Physics Olympiads

In order to ascertain the validity of the results we obtained, we investigate related data. The *International Physics Olympiad* (IPhO), another well established annual science Olympiad for students in secondary education, was originally founded by eastern European countries. Like the IMO participants, participants in the IPhO must be younger than 20 and not be enrolled in an institute of tertiary education. Olympians are selected through rounds of regional and national competitions, and the final team consists of, at most, 5 contestants. Cross-sectional data on national performance is available from the website of the IPhO (http://www.jyu.fi/ipho/). Performance data is available for two aggregated time periods (1967-1996 and 1997-2006). We concentrate on performance during the latter period. Unfortunately, there is no information about average scores of participants, so the dependent variable is a weighted index of medal counts, divided by the possible number of participants

$$(G + 0.75 \cdot S + 0.5 \cdot B + 0.25 \cdot H)\,/V$$

where G, S, and B are the number of gold, silver, and bronze medals, respectively; H is the number of honorable mentions; and V represents the possible number of participants over the aggregated time period. The resulting measure, called “efficiency of participation,” closely resembles the dependent variable in the previous analysis of results at the IMO.

Performance at the IMO and IPhO are highly correlated; the simple correlation coefficient is 0.85. China, Russia, and the United States are, again, among the best-performing countries, while Kuwait, Lichtenstein and Bolivia line the bottom of the performance ranking.

The basis of the following analysis is regression function (2). We focus on macro-conditions and schooling measures, so we drop participation and hosting history from the regression. Participation at the IPhO is not as common as participation at the IMO, so observations are available from only 67 countries, and the number declines further depending on the set of covariates. Regression results are displayed in Table 7. Column (1) contains finding from the basic specification, the influence of macro conditions is displayed in column (2), and the relationship between schooling measures and the performance is summarized in the final column.

All together, the estimates support the previous findings related to the IMO, but also show some distinct patterns. First, countries with bigger populations perform better at the IPhO, while the level of GDP per capita has no beneficial influence. Second, contestants from countries with a current communist regime, a large number of published scientific articles and low birth rates are relatively more successful. In contrast to findings from the IMO, cohort size matters, since a larger percentage of 15-19-year-olds in the population is associated with higher performance at the IPhO. Thus, while the period birth rate is negatively related to available resources for skill and talent formation of adolescents, countries with a large percentage of 15-19-year-olds in the population were able to endow the cohort with the necessary resources and enlarge the pool of talent. Furthermore, the degree of urbanization and the legacy of communist regimes are uncorrelated with national performance at the IPhO.

A third finding is that schooling inputs and outputs matter, as higher average PISA scores of top students (still in mathematics) are positively associated with performance at the IPhO. A standard deviation difference among countries translates into an expected difference in the performance of 27.6. This time, the quantity of secondary education also exhibits a significant positive influence. This may be explained by the possibility that achievement in physics is more closely connected to teaching in secondary education than mathematics is. Similar to the IMO findings, spending and enrollment rates negatively influence performance at the IPhO if schooling outputs are held constant. However, the coefficient of the gross enrollment rate is significant only at the 10% level. Therefore, there is evidence from IPhO results that participants from countries with more “efficient” schooling systems perform better at science Olympiads.

Table 7: Estimation Results of the International Physics Olympiad, 1997-2006

COEFFICIENT	(1) Basic	(2) Macro	(3) Schooling
Avg. $\ln POP_i$	8.956***	9.188***	9.463***
	(1.55)	(2.04)	(1.97)
Avg. $\ln GDPPC_i$	1.933	1.283	-16.73***
	(2.12)	(3.97)	(5.74)
Avg. $SHARE1519_i$		6.019**	
		(2.43)	
Avg. CBR_i		-2.873***	
		(0.59)	
Avg. $URBAN_{it}$		0.0388	
		(0.21)	
Avg. $ARTICLE_i$		0.119**	
		(0.059)	
COM_i		17.55**	
		(8.40)	
$FCOM_i$		8.036	
		(6.69)	
Avg. $ENROLL_i$			-0.442*
			(0.22)
Avg. $SECXPD_i$			-1.552**
			(0.69)
Avg. $TERT\_TLF_i$			0.149
			(0.34)
Avg. $SEC1524_i$			0.582***
			(0.19)
Avg. $PISA95th_i$			0.485***
			(0.070)
Constant	-135.1***	-144.7**	-236.7***
	(31.0)	(63.5)	(44.0)
Observations	62	55	35
R-squared	0.36	0.57	0.72

Note: Heteroscedasticity and autocorrelation robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

6 Conclusion

The aim of this study was to analyze country differences at the International Mathematical Olympiad. Drawing on the literature on country performance at Olympic Games and international soccer rankings, we derived hypotheses and an empirical framework to explain country-level variations in the average performance per participant at the IMO.

The major difference from related work on sporting events is the inclusion of measures of human capital, schools' resource endowment and school-quality indicators. The analysis also differentiates between inter-country associations and the effect of intra-country changes in the covariates on IMO performance. Cross-sectional analyses provide information about country characteristics that systematically contribute or hamper performance of talents at the Olympiads; however, inferences about the underlying talent pool requires information about the country-specific selection and training procedures. Given that these procedures are fairly time-invariant, a fixed-effects or first difference approach can overcome this limitation.

Our results show a mixed picture. On one hand, cross-sectional findings support the hypotheses: there are clear relationships between the level of explanatory variables and country-level IMO performance. We find a strong positive influence of a large population, specifically, a large effectively participating population. However, fertility levels and, thus, population growth rates have a negative influence on average performance. In addition the size of the academic market is influential and has the expected positive sign.

We also find evidence that IMO performance is related to a country's schooling system: Mean PISA scores of top students positively correlate with country results at the IMO, but the relationship vanishes once the other covariates are included in the regressions. Perhaps surprisingly, holding school quality constant, spending per student and coverage of secondary education has a negative influence on IMO performance. In contrast to previous studies on sporting events, there is no evidence that higher levels of GDP per capita correspond to better average performance.

These results are largely confirmed by findings from a dataset of country performance at the International Physics Olympiad, another well established international science Olympiad for students in secondary education.

On the other hand, the hypotheses fail to predict changes in performance over time. The advantage of analyzing within-variation is the elimination of unobserved country effects. Consequently, results allow to draw inferences about changes in the underlying talent pool. However, findings from a fixed-effects estimation suggest that only the number of participations is related to changes in performance, but this variable is unlikely to be strictly exogenous, and identification of the influences comes from a non-random subsample of countries.

Only intra-country estimations over longer time periods reveal certain effects. In particular, population growth is found to have a negative impact, which corresponds to the cross-sectional results. Other effects are sensitive to the investigated time periods and the country sample.

The dataset imposes two important limitations. First, cross-sectional comparisons are possible, but the lack of information about country-specific training and selection procedures may bias the results. Second, exams differ from Olympiad to Olympiad; it is not a priori clear whether problems are replaced by similar ones to represent a constant level of achievement over time, or whether they are adapted to reflect changes in knowledge and skills to represent a constant level of challenge. The first limitation can be overcome by either collecting and including adequate information or eliminating the unobserved heterogeneity. The latter limitation may be tackled by the included period dummies since countries with macro-conditions that develop relatively faster would still experience a relative growth in performance.

An issue not addressed in this study is whether former math Olympians fulfill their potentials later in life. There is some evidence that many of them are successful and excel specifically in academia and sciences (Engel et al., 2009). In fact, four of the last ten winners of the Fields Medal were at least once gold medalists at the IMO. Some western societies (e. g., the United States) seem to offer incentives that promote the academic productivity of Olympians (Shoho, 1996). The moderating role of environmental characteristics on the transition of young talent from education to career could be a promising direction for future research.

Overall, our results confirm the advantage of large populations and, more specifically, large effectively participating populations. Thus, it is not only the sheer number of heads that matters; the level of human capital and infrastructure is also key to success at the competition.

References

Andreescu, T., Gallian, J. A., Kane, J. M. and Mertz, J. E. (2008). Cross-cultural analysis of students with exceptional talent in mathematical problem solving, *Notices of the American Mathematical Society* **55**(10): 1248–1260.

Baltagi, B. H. and Griffin, J. M. (1984). Short and long run effects in pooled models, *International Economic Review* **25**(3): 631–645.

Becker, G. S. and Lewis, H. G. (1973). On the interaction between the quantity and quality of children, *The Journal of Political Economy* **81**(2, Part 2): S279–S288.

Bernard, A. B. and Busse, M. R. (2004). Who wins the Olympic Games: Economic resources and medal totals, *Review of Economics and Statistics* **86**(1): 413–417.

Bound, J. and Turner, S. (2007). Cohort crowding: How resources affect collegiate attainment, *Journal of Public Economics* **91**: 877–899.

Campbell, J. R. and Wu, W. (1996). Development of exceptional academic talent: International research studies, *International Journal of Educational Research* **25**(6): 479–484.

Cunha, F. and Heckman, J. (2007). The technology of skill formation, *American Economic Review* **97**(2): 31–47.

Easterlin, R. A. (1978). What will 1984 be like? Socioeconomic implications of recent twists in age structure, *Demography* **15**(4): 397–432.

Engel, W., Gronau, H.-D., Langmann, H.-H. and Sewerin, H. (2009). *The German Teams at the International Mathematical Olympiads 1959-2008*, K. H. Bock, Bad Honnef.

Fuchs, T. and Wößmann, L. (2007). What accounts for international differences in student performance? A re-examination using PISA data, *Empirical Economics* **32**(2): 433–464.

Gagné, F. (1985). Giftedness and talent: Reexamining a reexamination of the definitions, *Gifted Child Quarterly* **29**(3): 103–112.

Hanushek, E. A. (1992). The trade-off between child quantity and quality, *The Journal of Political Economy* **100**(1): 84–117.

Hanushek, E. A. (2003). The failure of input-based schooling policies, *The Economic Journal* **113**(485): F64–F98.

Heller, K. A. (2007). Scientific ability and creativity, *High Ability Studies* **18**(2): 209–234.

Hoffmann, R., Ging, L. C. and Ramasamy, B. (2002a). Public policy and Olympic success, *Applied Economics Letters* **9**(8): 545–548.

Hoffmann, R., Ging, L. C. and Ramasamy, B. (2002b). The socio-economic determinants of international soccer performance, *Journal of Applied Economics* **V**(2): 252–272.

Johnson, D. K. and Ali, A. (2002). A tale of two seasons: Participation and medal counts at the summer and winter Olympic games, *Wellesley College Working Paper* **2002-02**.

K.C., S., Barakat, B., Goujon, A., Skirbekk, V. and Lutz, W. (2008). Projection of populations by level of educational attainment, age and sex for 120 countries for 2005-2050, *IIASA Interim Report* **08-038**.

Krishna, A. and Haglund, E. (2008). Why do some countries win more Olympic medals? Lessons for social mobility and poverty reduction, *Economic & Political Weekly* **43**(28): 143–151.

Leeds, M. A. and Leeds, E. M. (2009). International soccer success and national institutions, *Journal of Sports Economics* **00**(00): 1–22.

Lotka, A. J. (1926). The frequency distribution of scientific productivity, *Journal of the Washington Academy of Sciences* **16**(12): 317–323.

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

Lutz, W., Goujon, A., K.C., S. and Sanderson, W. (2007). Reconstruction of populations by age, sex and level of educational attainment for 120 countries for 1970-2000, *Vienna Yearbook of Population Research* **2007**: 193–235.

Macmillan, P. and Smith, I. (2007). Explaining international soccer rankings, *Journal of Sports Economics* **8**(2): 202–213.

Merton, R. K. (1968). The Matthew Effect in science, *Science* **159**(3810): 56–63.

Mundlak, Y. (1978). On the pooling of time series and cross section data, *Econometrica* **46**(1): 69–85.

Murphy, K. M., Shleifer, A. and Vishny, R. W. (1991). The allocation of talent: Implications for growth, *The Quarterly Journal of Economics* **106**(2): 503–530.

Narin, F. and Breitzman, A. (1995). Inventive productivity, *Research Policy* **24**: 507–519.

Nichols, A. (2007). Causal inference with observational data, *The Stata Journal* **7**(4): 507–541.

Nokelainen, P., Tirri, K. and Campbell, J. R. (2004). Cross-cultural predictors of mathematical talent and academic productivity, *High Ability Studies* **15**(2): 229.

Nokelainen, P., Tirri, K., Campbell, J. R. and Walberg, H. (2007). Factors that contribute to or hinder academic productivity: Comparing two groups of most and least successful Olympians, *Educational Research and Evaluation* **13**(6): 482–500.

Nokelainen, P., Tirri, K. and Merenti-Valimaki, H.-L. (2007). Investigating the influence of attribution styles on the development of mathematical talent, *Gifted Child Quarterly* **51**(1): 64–81.

OECD (2004). *Learning for Tomorrow's World - First Results from PISA 2003*, OECD Publishing.

OECD (2007). *PISA 2006: Science Competencies for Tomorrow's World*, Vol. 1, OECD Publishing.

Pirotte, A. (1999). Convergence of the static estimation toward the long run effects of dynamic panel data models, *Economics Letters* **63**: 151–158.

Rosen, S. (1981). The economics of superstars, *The American Economic Review* **71**(5): 845–858.

Seglen, P. O. (1992). The skewness of science, *Journal of the American Society for Information Science* **43**(9): 628–638.

Shoho, A. R. (1996). A cross-cultural analysis of similarities and differences among Math Olympiads in China, Taiwan, and the United States, *International Journal of Educational Research* **25**(6): 575–582.

Subotnik, R. F., Miserandino, A. D. and Olszewski-Kubilius, P. (1996). Implications of the Olympiad studies for the development of mathematical talent in schools, *International Journal of Educational Research* **25**(6): 563–573.

Todd, P. E. and Wolpin, K. I. (2003). On the specification and estimation of the production function for cognitive achievement, *The Economic Journal* **113**(485): F3–F33.

Wooldridge, J. (2002). *Econometric Analysis of Cross Section and Panel Data*, The MIT Press, Cambridge, London.

Wößmann, L. (2001). Why students in some countries do better: International evidence on the importance of education policy, *Education Matters* **1**(2): 67–74.

Wößmann, L. (2003). Specifying human capital, *Journal of Economic Surveys* **17**(3): 239–270.