Economic Research-Ekonomska Istraživanja

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/rero20

Assessment of Olympic performance in relation to economic, demographic, geographic, and social factors: quantile and Tobit approaches

Wang Shasha, Babar Nawaz Abbasi & Ali Sohail

To cite this article: Wang Shasha, Babar Nawaz Abbasi & Ali Sohail (2023) Assessment of Olympic performance in relation to economic, demographic, geographic, and social factors: quantile and Tobit approaches, Economic Research-Ekonomska Istraživanja, 36:1, 2080735, DOI: 10.1080/1331677X.2022.2080735

To link to this article: https://doi.org/10.1080/1331677X.2022.2080735

Published online: 28 May 2022.

ECONOMIC RESEARCH-EKONOMSKA ISTRAŽIVANJA
2023, VOL. 36, NO. 1, 2080735
https://doi.org/10.1080/1331677X.2022.2080735

Assessment of Olympic performance in relation to economic, demographic, geographic, and social factors: quantile and Tobit approaches

Wang Shasha[a], Babar Nawaz Abbasi[b] and Ali Sohail[c]

[a]Xi'an Physical Education University, Xian, China; [b]School of Education, Zhengzhou University, Henan, China; [c]School of Public Policy and Administration, Xian Jiaotong University, Xian, China

ABSTRACT
The sports industry's global gain is worth drawing attention to in the face of economics. Scholars argued that a country's success in sports is directly related to its economic, demographic, geographic, and social factors. Therefore, investigating the relationship between these factors and sports prizes could be a garment in formulating policies for promoting sports success. Thus, this study used cross-sectional data analysis to investigate such claims based on the recent Olympic Games of Rio 2016 from 207 countries with more than 11,000 athletes on 306 events in 28 different sports by employing Quantile Regression and Tobit Regression models. The findings revealed that inflation rate in moderate performed countries; economically active population in low, moderate, high, and very high performed countries; and countries' income classification in low, moderate, high, and very high performed countries are influencing the countries medal ranking performance in the Olympic Games. Furthermore, countries with high temperatures are not likely to do well in the games. However, the size of a country's GDP level, corruption ranking level, the number of athletes, and the topographical nature of a country have no impact on the countries medal ranking performance in the Olympic Games.

ARTICLE HISTORY
Received 20 August 2021
Accepted 17 May 2022

KEYWORDS
Olympic Games; sports performance; economic factors; demographic factors; social factors

JEL CLASSIFICATION CODE
Z20; Z23; Z28

1. Introduction

Initially, the sport began as a leisure time activity reserved only for a few declared amateurs. However, by the 20th century, it has become a social phenomenon for the masses where it has developed into its economic branch across the globe as the industry becomes a sector with significant influence to the extent that sport ranked among the top mainstream activities in the economy (Chappelet, 2005; Mendoza, 2017). Sport is an important sector of economic activity and spans medical treatment and rehabilitation, research and development, sports tourism, sales and trade of sports products,

CONTACT Wang Shasha 380039828@qq.com

construction and maintenance of sports venues, organization sports events, marketing, and advertising. It also creates various job and business opportunities (e.g., engineers and developers, coaches and sports doctors, sports journalists and commentators, retailers of sports goods and equipment). Sports are also part of the increasingly important leisure industry and have broader benefits such as boosting productivity, individual and community development and reducing the ill-health burden on society. Moreover, in addition to the significant impact of sport on a country's economy, the sport has assumed an ever-greater role within the globalization process and the regeneration of national identity (Nauright, 2004; Luiz & Fadal, 2011).

There are various types of sports events such as Olympics, Commonwealth Games, Cricket world cup, Football (soccer) world cup, and Rugby world cup. However, Olympic Games also called the Olympian Games, the Olympiad, the Games of the Olympiad, or commonly called the Olympics are considered the world's foremost sports event with over 200 participating nations. The Olympics are an international sports festival and the ultimate goals are to cultivate human beings, through sport, and contribute to world peace. In Olympics, the top three finishers in each event are famously awarded Olympic medals: gold for first place, silver for second place, and bronze for third place. Therefore, successful participation in a sporting event is usually measured by winning medals (Lozano et al., 2002; Li et al., 2015) or by its weighted difference (Wu, 2009). Typical Olympic Games can be depicted from the latest Olympic Games in Rio de Janeiro 2016, where there are 207 nations, 11,238 athletes, 306 events, and 28 different sports. However, the content of the Olympics can be traced since 1896 in Athens, Greece where the first modern Olympic Games were hosted in which athletics, cycling, fencing, gymnastics, shooting, swimming, tennis, weightlifting, and wrestling were all included, and they remain so today. Furthermore, previously the host country had extensive control over which sports were included but today, the International Olympic Committee (IOC) makes such a decision, which has several terms for competitions based on different sports. The game hosts summer and winter sports competitions, and it is attracting thousands of athletes from all over the world to compete in a variety of events. The games are usually held every four years, with the summer and winter Olympics alternated every two years (Overview of Olympic Games, 2008; International Olympic Committee, 2018).

The Olympic Games are regarded as a mega-event with global significance, where successful athletes' (sport) results can lead to higher incomes for national sports federations as well as increased public interest in a specific sport and athletes. The economic benefits of a winning team, according to Rosas and Flegl (2019), extend beyond the winning team or athlete. People have a tendency to publicly identify with successful sports teams (End et al., 2002). As a result, team success can have an economic impact by increasing consumer spending, because winning has a significant positive impact on real wage income per capita (Davis & End, 2010). Moreover, the Olympic Games' significance was reflected in the broadcasting contract, in which NBC Universal paid $4.38 billion for the rights to broadcast the Olympics through 2020 (with a $7.65 billion agreement extension from 2021 to 2032) (Rosas & Flegl, 2019). Therefore, investigating the factors that could trigger the Olympic Games performance of participating nations is of high importance as it provides crucial information to athletes, coaches, sports policymakers and sport science practitioners (Skarbalius et al., 2019).

Been the gain of the Olympic Games performance is rewarded with medals, why do some countries win medals while others do not, and thus, why some countries performed well while others are not? Previous studies have argued that a country's success in sports is directly related to the economic endowments and demographic features (Kiviaho & Makela, 1978; Andreff, 2001; Lozano et al., 2002; Bernard & Busse, 2004; Johnson & Ali, 2004; Churilow & Flitman, 2006; Luiz & Fadal, 2011). Furthermore, in addition to the main economic and demographic factors, other social factors such as a country's income classification, corruption perception index, and athlete population can all be important determinants of a country's sports performance (Churilow & Flitman, 2006; Flegl, 2014). However, the relationship between such factors of sporting success has been studied in numerous studies, including (Kiviaho & Makela, 1978; Baimbridge, 1998; Condon et al., 1999; Kuper & Sterken, 2001; Hoffmann et al., 2002, 2004; Tcha & Pershin, 2003; Bernard & Busse, 2004; Johnson & Ali, 2004, as well as Matros & Namoro, 2004; Luiz & Fadal, 2011; Rosas & Flegl, 2019) where the majority of the research used Olympic medal counts as the outcome variable to represent Olympic success, with socio-economic factors serving as explanatory variables.

Going from the existing studies, it is evidenced that most; all the existing studies ignored the economic classification of the participant nations, inflation, corruption, temperature, and topographical factors of the participant countries despite their potentiality in determining the success of the sports. Furthermore, no study ever covers economic, demographic, social, and geographical factors related to sports success. It also seems that the application of the Quantile Regression (QR) model is scarce despite its potential in such studies. Hence, the current study is expected to contribute to the theoretical literature by adding new determinants such as the economic classification of the participant nations, inflation, corruption, temperature, and topographical factors in conducting empirical studies. Second, it will add to the empirical literature by using a rare case methodology in investigating sports success which to the best of the study knowledge, this study is the first to assess the Olympic Games performance with respect to economic, demographic, social, and geographical factors in part or whole, using Quantile Regression (QR) model. For sports generally, the few that used it; frequently misinterpreted it, which according to Leeds (2014), there are many cases where sports researchers do not understand how to use QR or interpret it. However, out of the alternative models to Ordinary Least Squares (OLS) model; which is the benchmark for analyzing cross-sectional data, this study decides to adopt the QR model because of the following reasons: it has flexibility for modeling non-normal data, or heterogeneous conditional distributions and can take care with possible nonlinearity in the relationship (Geraci, 2014); again, its ability to enable the examination of various conditional quantiles of the dependent variable, revealing a range of heterogeneity in the analysis of Olympic Games performance differences; also, when the normality assumption was violated or outliers and long tails were present, more robust and complete estimates were obtained than with the mean regression (Huang et al., 2017); similarly, hitherto to all, it renders no assumptions about the distribution of the residuals, which appears to mean it is not bound by restrictive assumptions about the distribution of the error term and this paved it away to becoming mighty alternative method to the OLS method (Leeds, 2014, Flom, 2018). Third, in this study, for the sake of a robustness check to support the findings of the Quantile Regression model, the Tobit Regression (TR) model was employed.

Therefore, for the benefit of global managers and administrators in the sports sector for the Olympic Games as well as general readership on determinants of sports success, this study proposed to provide answers to the following questions: do a country's economic factors such as gross domestic product and inflation impact its performance in Olympic Games? Do a country's demographic factors, such as active population, influence its performance in Olympic Games? Do a country's social factors such as corruption level, countries' income classification, and the number of athletes affects its performance in Olympic Games? Are geographical factors such as temperature and topography impacting a country's performance in the Olympic Games? Finally, do the influences of the factors change with different quantiles of a country's Olympic Games performance?

Hence, the specific objectives of this study are: To examine the impact of a country's economic factors such as gross domestic product and inflation on Olympic Games performance; to investigate the influence of a country's demographic factors such as active population on Olympic Games performance; to analyze the effect of a country's social factors such as corruption level, countries' income classification, and the number of athletes on Olympic Games performance; to explore the effect of a country's geographical factors such as temperature and topography on Olympic Games performance and to explore whether the influences of the factors change with different quantiles of a country's Olympic Games performance.

Therefore, this study tested the following null hypotheses: the economic factors of a country such as gross domestic product and inflation have no impact on its Olympic Games performance; the demographic factors of a country such as active population have no impact on its Olympic Games performance; the social factors of a country such as corruption level, countries' income classification, and the number of athletes do not influence its Olympic Games performance; the geographical factors of a country such as a temperature and topography do not affect its Olympic Games performance, and the influences of the factors do not change with different quantiles of a country's Olympic Games performance.

The remainder of the study is structured as follows: literature review, namely theoretical and empirical reviews; the methodology for achieving the study's objectives; results presentation, discussion of the results; and lastly is the concluding remarks of the study.

2. Literature review

Following the traditional GDP-population-based theory of Olympic success, sports performance is usually measured regarding GDP and population. These two factors are widely acknowledged as the most important economic and demographic factors Lozano et al. (2002) apart from social factors such as corruption level, countries' income classification, and the number of athletes; and geographical factors such as temperature and topography. Now, let's theoretically deliberate on how each of these factors/determinants impact sports performance; starting with economic, then demographic, and followed by social as well as geographical factors.

Higher GDP translates into better health and education outcomes, better sporting infrastructure, and more resources spent on sports. Luiz and Fadal (2011) have argued that a country's sporting success should be measured in relation to its economic resources. The size of the economy is strongly linked to a country's performance at the Olympics. More leisure time is related to increased money, allowing a larger percentage of the population to participate in competitive sports. Wealthy countries are better equipped to supply their top athletes with the sporting facilities, coaching skills, and sports science support services they need to compete internationally. Additionally, as a country grows richer, it usually produces better sportspersons. The critical factor is how public resources are spent (Debroy, 2011). In general, the larger a country's GDP is, the more medals it receives (Sen, 2021). Inflation is another economic factor to consider. Rising cotton prices are putting pressure on polyester and other man-made fibers in the outdoor and fitness apparel sector. Specifically, inflation and currency value can have a variety of effects on athletes. Currency exchange rates heavily influence the real worth of international players' playing contracts. A number of sports professionals sign contracts in currencies other than their home countries. As a result, their contracts are vulnerable to currency swings, particularly because they may be sending money home to their families or investing. Therefore, international athletes should be aware of currency patterns in the country where they are now based compared to currency trends in their home country. They should aim to exploit currency changes to their advantage rather than a disadvantage. In addition, inflation and the value of foreign currencies have a direct impact on the Money Smart Athlete's investment portfolio, both in terms of revenue earned, which has a lower real worth during inflationary periods, and in terms of investments in foreign currencies, foreign equities, bonds, and real estate. Money Smart Athletes can benefit from the guidance of a qualified financial advisor who can assist them navigates the complexity of currency values and inflation and avoid a detrimental influence on their total finances (Money Smart Athlete Blog, 2018).

A large population means a larger pool of fantastic talent to choose from for the game performance. Hence, argued that the active population could be used to determine sports performance. Nations with a high population have a better chance of winning international sports tournaments than countries with a smaller population. The variety in physical attributes within a population has an impact on a country's sporting achievement. The larger the population, the more genetically gifted athletes with physical attributes (such as height, limb measurements, body weight, cardiovascular ability, muscular strength, and muscle fiber qualities) adapted to improved sports performance. Foster et al. (2010) have it that the quantity of competitors also influences human sports performance. As a result, it is hypothesized that athletic performance will be affected by the size of the global human population.

Aside from economic and demographic characteristics; social characteristics such as corruption level, countries' income classification, and the number of athletes could also be used to determine Olympic Games. Ramírez and Sánchez (2013) and Transparency International (2016) expressed that corruption/corrupted environment are essential components to take into consideration while examining Olympic Games performance. According to Ibrahim (2016), integrity is widely seen as a vital issue in today's world, and the sports environment is no exception. Reasons for integrity behaviour in sport and its linked support systems can be attributed to a variety of variables, including corruption/doping and overemphasis. Athletes who do not achieve the levels of performance generally required in the sport in question to win the tournament and instead allow others to win are said to be involved in sports corruption, or Because they receive or expect pecuniary or non-pecuniary advantage1 for themselves (or, in case the person in question is acting as an agent: for a principal, e.g., acquaintances, relatives, and/or associated sports institutions) from sporting officials who consciously perform their assigned tasks in a manner at odds with the objectives and moral values of the relevant club, association, competitive sports in general, and/or society at large. Similarly, high-income countries like the United States, the United Kingdom, Australia, France, Germany, Brazil and Italy have an advantage in sporting competitions by earning more medals than countries with lower economic levels because of their economic endowments (Sen, 2021). In addition, the size of the Olympic team, which depends on the number of athletes, could be served as a predictor of Olympic success (Vagenas & Vlachokyriakou, 2012).

Furthermore, geographical factors such as temperature and topography could be key predictors of athletic achievement. According to Song and Zhang (2018), climate change will not only influence sports circumstances but will also interfere with athletes' physiological mechanisms and emotions, diminishing athletic performance. Sports and sports culture are inextricably linked to the physical environment. Human culture is influenced by the local geographical environment as well as national traits. Every ethnic culture emerges from and develops in the context of a specific natural geographical setting. At the same time, the geographical context will have a positive and bad impact on the growth and promotion of sports and sports culture. In fact, the climate is one of the most important aspects that influence regional sports growth. Climate conditions have an impact on not only the feasibility of sports but also the athletes' happiness and physiological function. People in different climates, such as Norwegians living in the Arctic Circle, have varied lifestyles and physical activities. Because the region is located in the subfrigid coniferous forest temperature zone adjacent to the Arctic Ocean and the Norwegian Sea, residents primarily engage in ice sports due to the chilly climate. The tropics include countries like Ecuador, which are located near the equator. Due to the hot heat and lengthy illumination time, sprint, long jump, surfing, and other appropriate activities for the local climate are popular. Topography also has a significant impact on the formation and development of sports. Many sporting events are inextricably linked to the topography of the area. People who live near steep cliffs, for example, are more likely to engage in rock climbing, mountain climbing, and abseiling; people who live near a large coastline are more likely to engage in swimming, surfing, and other sports; and people who live in plain terrain and plateau areas are more likely to participate in races, horseback riding, and other sports. Different geographical environments spawned various sports.

Varied sports events are formed by topography, mountains, rivers, and different geographical environments. People in Switzerland, Austria, and other countries near the Alps, for example, advocate skating; Australia, the United States, New Zealand, and other countries have abundant coastal resources, and people there advocate surfing and windsurfing; Africa has the world's largest plain—the Amazon plain—so Africans are good at long-distance running, and our country's terrain is very complex, so there are many different sports. Archery, javelin and other technological sports are popular among the Chinese minority Qiang, who dwell in towering mountains and steep hills with many rivers.

Nonetheless, the mechanism of effect is further discussed by concerning the institutional support of the countries. The countries' institutional support depicts a lack of support for individuals who demonstrate athletic talent, and sports have eluded some of the countries due to limited public investible resources. Misallocation, lack of transparency, poor asset management, and the lack of a framework for monitoring the impact of government expenditure exacerbate the problem. Despite the governments' best efforts, this is unlikely to change. There are, however, scholarships and endowments for athletes that provide a minimal minimum quality of life, but this system is riddled with red tape, political intervention, conflicts of interest, and corruption. Furthermore, sporting organizations are not immune to scandal. In the meantime, some people are blaming the Olympic sports selection process.

Empirically, a plethora of studies has been conducted between such factors and sports performance. For instance, Bosscher et al. (2008). They studied the concept of measuring nations' success in elite sport at the 2004 Olympic Games in Athens using the Ordinary Least Squares (OLS) method where population, gross domestic per capita, and communism were found to influence performance in the games. Gorse and Chadwick (2011) have examined corruption in sport and try to address the implications of corruption for sponsorship programs by coding and analyzing a data set gathered from cases of corruption in international sport as the first process of a five-stage mixed method approach where the results reveal that corruption in international sport is a very real problem that is jeopardizing the financial future of some sectors of the industry and subverting its integrity. Luiz and Fadal (2011) have conducted economic scrutiny of Olympic Games performance in African nations in Beijing 2008 using the Ordinary Least Squares (OLS) method in the form of multiple regression analysis were the findings revealed that besides GDP, sporting success appears to be influenced by population size and elite facilities while climate is not. Feizabadi et al. (2013) have investigated the relationship between a country's success at the Guangzhou 2010 Summer Asian Games and demo-economic factors by using the Kolmogorov–Smirnov, one-way Analysis of Variance (ANOVA), and Stepwise Multiple Regression (SMR) Analysis and the findings revealed a significant association between the success of nations at the Guangzhou 2010 Asian Games and all factors of demo-economic (population, GDP, health expense, growth rate, team size, Ex-host). Sharma (2015) has examined the effect of socio-economic status on the sports performance of junior national level weightlifters from India's rural and urban areas using means, standard deviations, and t-ratios where socio-economic status had a positive effect on junior national level male weightlifters' sports performance, while urban junior national level male weightlifters of high, mediocre, and low socio-economic status outperformed their counterparts. Rosas and Flegl (2019) have conducted research on the quantitative and qualitative impact of GDP, corruption, and other social factors on Olympic Games performance in Rio 2016 using the Ordinary Least Square (OLS) method and found that countries' income classification, with respect to gross national income, active economic population, and corruption characters would result in improved performance in Olympic Games, whereas inflation would not.

So, from the above foregoing, it is theoretically evidenced that medal accomplishments of the participant nations should be weighted in relation to a country's economic, demographic, social, and geographical factors. On the other hand, the empirical review evidenced that all the existing studies ignored the economic classes of the participant nations, inflation, corruption, temperature, and topographical factors of the participant countries despite their potentiality in determining the success of the sports. Moreover, no study ever covers economic, demographic, social, and geographical factors related to sports success. It also seems that the application of the Quantile Regression (QR) model is scarce despite its potential in such studies. This portrays that the current study will contribute to the theoretical literature by adding new determinants such as the economic classes of the participant nations, inflation, corruption, temperature, and topographical factors in conducting empirical studies. Second, it will add to the empirical literature by using a rare case methodology in investigating sports success.

3. Methodology

3.1. Data and variables

This study uses the recent Olympic Games of Rio 2016 data which is cross-sectional, consisting of 207 countries with more than 11,000 athletes on 306 events in 28 different sports. However, due to data availability, this study makes some adjustments before using the data. It excludes Independent Olympic Athletes and Refugee Olympic Athletes because they are not real countries and thus have no economic or demographic data. Furthermore, due to a lack of GDP or economic active population data, the study excludes the following 19 countries: Palestine, Andorra, Bermuda, British Virgin Islands, Cayman Islands, Chinese Taipei, Cook Islands, Dominica, Liechtenstein, Marshall Islands, Monaco, Nauru, American Samoa, Netherlands Antilles, Palau, Saint Kitts and Nevis, San Marino, Tuvalu, US Virgin Islands, and the Cayman Islands. Moreover, the data comprises weighted medal ranking (Y) based on gold, silver, and bronze. The first 8 places from each discipline were gold, silver, and bronze medals, as well as other higher positions, which are thought to be more valuable. The International Amateur Athletic Federation (IAAF) methodology assigned 8points to the gold medal, 7 points to silver, and so on until 1 point to 8th place in each discipline. This was used as the dependent variable to represent sports performance. Furthermore, the independent variables employed comprise economic, demographic, social, and geographical factors as independent variables. Gross domestic product in US dollars (GDP) and inflation in annual percentage (INF) were used for economic factors. The active economic population (people aged 15 to 64 as a proportion of the total population) (*EAP*) was used for demographic factors. For social factors, corruption level measured by the corruption perception index (*CPI*); countries' income classification (*CIC*) treated as a dichotomous variable: 1 – low-income economies, 2 – lower-middle-income economies, 3 – upper-middle-income economies, and 4 – high-income economies; and the number of athletes (*NOA*) was used. For geographical factors, the temperature in C^0 annual average (*TEMP*) and topography measured by average elevation in feet (*TOPOG*) were used. In total, there are 186 observations for each of the variables. However, *GDP*, *INF*, *EAP*, *CIC* and *TEMP* were sourced from the World Bank database (World Bank, 2019), whereas *CPI* was sourced from Transparency International (2016), and *TOPOG* was sourced from Wikipedia (2020) while *Y* and *NOA* were sourced from the Rio 2016 - athletics schedule and results. Since the duration of the preparation of the Olympic Games is usually based on 4 years period, the data used in the analysis covers 4-year-long cycles from 2011 to 2015 and then averaged based on the time period. On the other hand, only *GDP*, *EAP* and *TOPOG* were transformed into their respective logarithms due to their size, i.e., for tackling heteroscedasticity and better results; hence their coefficients will be interpreted in the form of elasticities.

3.2. Model specification

Model specification is a functional form stating a given relationship where a dependent variable is a function of the independent variable(s). In this study, the model specification can be specified as follows:

$$\begin{aligned} Y_i = \alpha_0 + \alpha_1 \ln GDP_{1i} + \alpha_2 INF_{2i} + \alpha_3 \ln EAP_{3i} + \alpha_4 CPI_{4i} + \alpha_5 CIC_{5i} + \alpha_6 NOA_{6i} \\ + \alpha_7 TEMP_{7i} + \alpha_8 TOPOG_{8i} + \mu_i \end{aligned} \tag{1}$$

where *Y* is the weighted medal ranking of the first 8 positions, *GDP* is the gross domestic product in US dollars, *INF* is the inflation rate, *EAP* is the active economic population, *CPI* is the corruption perception index, *CIC* is the countries' income classification, *NOA* is the number of athletes, *TEMP* is the temperature, *TOPOG* is the topography, μ is the error term of the model, and *ln* is the natural logarithm.

3.3. Estimation techniques

In regression analysis and estimation, regression analysis is widely used (Miller, 2006). It seeks to determine the relationship between one or more independent variables and one or more dependent variables. One such regression analysis is the Ordinary Least Squares (OLS) model, which is the benchmark for analyzing cross-sectional data. The OLS model is one of the most widely employed statistical techniques. If its underlying assumptions are true, it has favourable properties; however, if those assumptions are not true, it can produce misleading results; thus, ordinary least squares are said to be not robust to assumptions violations. As a result, it is predicated on assumptions that are frequently not met. In the violation of such assumptions, an OLS model's alternative method will be used depending on the nature of the violated assumption. Hence, as one of the alternative methods to the OLS model;

this study decides to adopt Quantile Regression (QR) because of its flexibility for modelling nonnormal data, or heterogeneous conditional distributions (Geraci, 2014), ability to take care with possible nonlinearity in the relationships, and its ability to enable the analysis of different conditional quantiles of the dependent variable, revealing a range of heterogeneity in the analysis of differences in Olympic Games performance when the normality assumption was violated or outliers and long tails were present, it provided more robust and complete estimates than mean regression. are present (Huang et al., 2017). Hitherto to all, it makes no assumptions about the distribution of the residuals, implying that it is not bound by restrictive assumptions about the distribution of the error term, and this paved the way to becoming a powerful alternative method to the OLS method (Leeds, 2014; Flom, 2018). So, the QR methodology provides a comprehensive and elaborated analysis for a better understanding of the contributions of independent variables by characterizing the dependent variable's entire conditional distribution (Zhang et al., 2020).

Therefore, the study will first estimate the descriptive statistics and correlation analysis to understand the statistical characteristics of the variables under the study. Then the OLS model within the framework of the Multiple Linear Regression (MLR) model will be employed. However, the OLS model will function as a baseline model. Next is the estimation of the Quantile Regression (QR) model as an advanced model. It is important to note that, as is common when error terms are not normally distributed, the standard errors might be difficult or impossible to compute; hence in such instances, bootstrapping should be used to unravel the difficulty (Leeds, 2014). Followed is the graphical assessment of the explanatory variables based on the QR model results in combination with the multiple regression results. What is more is the estimation of Tobit Regression (TR) model for the sake of robust check to support the findings of the QR model since the outcome variable has some few countries with zero score and thus, zero observations while in such instance following the work of Amore and Murtinu (2021); Tobit models are better suited for robustness check. Last, the study will make use of the Durbin-Wu-Hausman test (augmented regression test) for endogeneity to make sure that the explanatory variables are not correlated with the errors of the general quantile regression model, and because cross data are subject to significant heteroscedasticity, this study will make use of qreg2 wrapper that estimates quantile regression and reports standard errors and t-statistics that are asymptotically valid even under heteroskedasticity and misspecification (Machado & Silva, 2013).

3.3.1. Ordinary least squares (OLS) model

The ordinary least squares (OLS) model is a technique widely used to estimate the parameter of the linear regression model, with the goal of closely "fitting" a function to the data by minimizing the sum of squared errors (a difference between observed and predicted values). The OLS estimates are consistent when the regressors are exogenous, and the errors are homoscedastic and serially uncorrelated. The OLS specification is stated as follows:

$$y_i = \beta X_i + \varepsilon_i \tag{2}$$

where y_i denotes the dependent variable, β is a $p \times 1$ vector of unknown parameters, and X_i denotes a vector of independent variables. The error term ε_i represents the error term that is $iddN(0, \sigma^2)$ and uncorrelated with X_i. However, usually, the constant term is included in the set of regressors.

3.3.2. Quantile regression (QR) model

Quantile regression (QR) is a method that is an alternative to the ordinary least squares (OLS) method. However, in contrast to the OLS, the QR does not make distributional assumptions, i.e., assumptions about residuals. Barnesa and Hughes (2002) have viewed that QR is a model for various conditional quantile functions, the median (0.5 quantile or 50th percentile) regression estimator was to estimate the conditional median function, the symmetrically weighted sum of absolute errors is minimized, and other conditional quantile functions are estimated by minimizing an asymmetrically weighted sum of absolute errors, with the weights varying according to the quantile of interest (Koenker & Hallock, 2001). The QR model specification is often specified as follows:

$$y_i = \beta_\theta x_i' + \varepsilon_{\theta i} \tag{3}$$

Or, alternatively in integral form:

$$\theta = \int_{-\infty}^{\beta_\theta x_i'} f_y(s|x_i)ds \tag{4}$$

where β_θ is an unknown $p \times 1$ vector of unknown parameters associated with the th percentile, xi is an independent variable vector of $p \times 1$, y_i is the outcome variable of interest, and $\varepsilon_{\theta i}$ is an unknown error term. The θ^{th} conditional quantile of y given x, on the other hand, is $Quant_\theta\ (y_i \mid x_i) = \beta_\theta x_i'$ where its estimate is given by $\hat{\beta_\theta}\, x_i'$ and as the θ increases, the conditional distribution of y given x is traced out. Although many empirical quantile regression studies assume that errors are distributed independently and identically (*i.i.d.*), the only necessary assumption concerning $\varepsilon_{\theta i}$ is $Quant_\theta\ (\varepsilon_{\theta i} \mid x_i) = 0$, which means that the conditional θ^{th} quantile of the error term is equal to zero (Barnesa & Hughes, 2002). As such, the quantile regression model accommodates parameter heterogeneity. Therefore, the quantile regression estimator can be found as the solution to the minimization problem as follows:

$$\hat{\beta}_\theta = \arg_\beta \min\left(\sum_{i:y_i > \beta x_i'} \beta|y_i - \beta x_i'| + \sum_{i:y_i > \beta x_i'} (1-\theta) \sum_{i:y_i > \beta x_i'} \beta|y_i - \beta x_i'|\right) \tag{5}$$

That is, by minimizing a weighted sum of the absolute errors, where the weights are symmetric for the median regression case ($\theta = 0.5$) and asymmetric otherwise. This minimization process can be expressed as linear programming (LM), a combination of the two, or a generalized method of moment (GMM). The LM implies that the method is computationally simple, whereas the GMM implies that $\sqrt{n}(\hat{\beta_\theta} - \beta_\theta) d\ N(0, \Psi_\theta)$, as a result, the test can be built with asymptotic justification using critical values from the normal distribution.

3.3.3. Tobit regression (TR) model

The tobit regression model (TR) belongs to a subset of econometric techniques known as censored regression models (Wooldridge, 2002). It was originally introduced to model non-negative continuous variables with several observations taking value zero (0) (Cunillera, 2014). Moreover, Tobit models have been used in management research to answer a variety of questions by reviewing existing practices and applications; Amore and Murtinu (2021) observed there are three issues to consider: assumptions about the nature of the data, the apparent interchangeability of censoring and selection bias, and potential violations of key assumptions in residual distribution. In the TR model, the outcome variable is called the latent variable and can be observed when it is positive though the observations would be censored in case of negative. Furthermore, the model allows for censoring from above instead of from below or for censoring both from below and above. Therefore, the TR model can be used to predict an outcome that is censored from above, from below or both. However, the Tobit regression model can be expressed as follows:

$$y^* = X'\beta + \varepsilon, \text{ with } \varepsilon|\ X \sim N(0; \sigma^2), \text{ with } y = y^* \text{ if } y^* > 0, \text{ and } y = 0 \text{ otherwise} \tag{6}$$

where y is the observed variable of interest and y^* is the latent variable where the equation, in general, reveals the expected impact of X on y^* is monotonic, the residuals have a normal distribution, and the dependent variable is left-censored. Typically, the TR model is estimated using the maximum likelihood method, and the TR model coefficients are interpreted similarly to OLS regression coefficients, except that the linear effect is on the uncensored latent variable rather than the observed outcome (Introduction to SAS, 2021).

4. Results presentation

This section presents the results of the study. It started with the estimation of the descriptive statistics and correlation analysis so as to understand the statistical characteristics of the variables under the study, next is the estimation of the multiple regression model, then is the estimation of the Quantile Regression (QR) model, and lastly is the estimation of the Tobit Regression (TR) model for checking the robustness of the QR model.

Table 1 shows the selected variables' average values, standard deviation, skewness, kurtosis, and Jarque-Bera. The table shows that the mean of *Y*, *lnGDP*, *INF*, *lnEAP*, *CPI*, *CIC*, *NOA*, *TEMP*, and *lnTOPOG* are 61.13514, 24.06522, 4.745712, 15.33302, 40.86127, 2.691892, 60.39459, 19.19157, and 7.016216, respectively. The variable with higher variability than all other variables is *Y*, with a standard deviation of 145.7367, while *CIC* is the variable with the least variability according to it is the standard deviation of 1.066888. The skewness revealed positive skewness in the series distribution of *Y*, *INF*, *CPI*, and *NOA*, while *lnGDP*, *lnEAP*, *CIC*, *TEMP*, and *lnTOPOG* are

Table 1. Descriptive statistics result.

	Y	*lnGDP*	*INF*	*lnEAP*	*CPI*	*CIC*	*NOA*	*TEMP*	*lnTOPOG*
Mean	61.13514	24.06522	4.745712	15.33302	40.86127	2.691892	60.39459	19.19157	7.016216
Median	6.000000	24.12260	2.960000	15.51484	37.25000	3.000000	14.00000	22.65000	7.160000
Maximum	1183.000	30.44800	65.21667	20.72626	91.00000	4.000000	554.0000	28.83000	9.180000
Minimum	0.000000	0.000000	−5.920000	11.00185	0.000000	1.000000	2.000000	−6.360000	3.060000
Std. Dev.	145.7367	3.407670	6.905609	1.912904	20.39680	1.066888	99.93199	8.183872	1.137387
Skewness	4.370842	−3.708876	4.914250	−0.213176	0.554175	−0.197451	2.643130	−0.792731	−0.990427
Kurtosis	26.63864	27.74583	37.59732	3.013414	2.878400	1.787457	10.00343	2.559855	4.394499
Jarque-Bera	4896.350	5144.381	9971.300	1.402579	9.583218	12.53536	593.4850	20.86969	45.23564
Probability	0.000000	0.000000	0.000000	0.495945	0.008299	0.001897	0.000000	0.000029	0.000000
Sum	11310.00	4452.065	877.9567	2836.608	7559.335	498.0000	11173.00	3550.440	1298.000
Sum Sq. Dev.	3908008.	2136.647	8774.489	673.2931	76549.43	209.4378	1837498.	12323.54	238.0316
Observations	187	187	187	187	187	187	187	187	187

Source: Authors' Computation.

Table 2. Correlations analysis.

	Y	lnGDP	INF	lnEAP	CPI	CIC	NOA	TEMP	lnTOPOG
Y	1.000000								
lnGDP	0.436865	1.000000							
INF	−0.027865	0.041609	1.000000						
lnEAP	0.448427	0.584207	0.200394	1.000000					
CPI	0.331788	0.404634	−0.273574	0.009459	1.000000				
CIC	0.320062	0.302680	−0.292462	−0.116815	0.631021	1.000000			
NOA	0.890139	0.515509	−0.021367	0.500987	0.417763	0.408336	1.000000		
TEMP	−0.367814	−0.262598	−0.059882	−0.125368	−0.397333	−0.381867	−0.353925	1.000000	
lnTOPOG	0.006983	−0.029757	−0.000620	0.081836	−0.142021	−0.171919	−0.029338	−0.068120	1.000000

Source: Authors' Computation.

negatively skewed. However, kurtosis and the Jarque-Bera show that all the series have volatile distribution except ln*EAP*, which has kurtosis close to 3 and an insignificant Jarque-Bera *p*-value. However, the observations of all the variables are balanced, hence; balanced cross-sectional data.

Table 2 presents the result of the correlation matrix between the variables under the study, namely *Y*, *lnGDP*, *INF*, *lnEAP*, *CPI*, *CIC*, *NOA*, *TEMP*, and *lnTOPOG*. Following the table, *INF* and *lnTOPOG* are low correlated with *Y*; ln*GDP*, ln*EAP*, *CPI*, and *TEMP* are moderately correlated with *Y*; but NOA has a high correlation with *Y*. Therefore, all the selected independent variables seem to have an association with the dependent variable *Y*. Furthermore, the correction result revealed that all the independent variables are positively correlated with the dependent variable except *INF* and *TEMP* which are negatively correlated with it.

Table 3 shows the estimate of the multiple regression model where the coefficients of the economic active population (*lnEAP*), countries' income classification (*CIC*), and temperature (*TEMP*) are statistically significant all at 1% level and positively related with the medal ranking (Y) except *TEMP* which is negatively related with it while that of gross domestic product (*lnGDP*), inflation (*INF*), corruption perception index (*CPI*), number of athletes (*NOA*), and topography (*TOPOG*) are not significant. In relation to the model's statistical healthiness, though the model is free from serial correlation as the test's *p*-value is not significant, the model has a problem of heteroscedasticity, nonnormal errors, and miss-specification problem as the tests' *p*-values are all significant. Notwithstanding, the study goes further to conduct other diagnostic tests, including the graphical normality test, outlier test, multicollinearity test, and linearity test. Furthermore, the study conducts a leverage effect test to know how far the explanatory variables deviate from their mean because high leverage points can greatly affect the regression coefficients.

Figure 1 presents the graphical depiction of the normality test of the estimated model. By glancing at the figure, it can be deduced that the residuals of the model are not normally distributed.

Figure 2 reports the test for checking the presence of outliers in the model. From the figure, it can be observed that in all the graphs of the independent variables in relation to the dependent variable, there is evidence of outliers but vary in degree; thus, the observations of the independent variables have large residuals in predicting the dependent variable.

Table 3. Multiple regression (i.e., OLS) results.

Y	Coefficient	Std. Err.	*T*	$p > \vert t \vert$	[95% Conf. Interval]	
lnGDP	1.639969	3.702926	0.44	0.658	−5.667597	8.947536
INF	−.784013	1.388047	−0.56	0.573	−3.523265	1.955239
lnEAP	32.32492	6.055011	5.34	0.000*	20.37562	44.27423
CPI	.5749592	.5959673	0.96	0.336	−.6011567	1.751075
CIC	31.46064	11.40692	2.76	0.006*	8.949576	53.9717
NOA	.1846933	.3602071	0.51	0.609	−.5261599	.8955465
TEMP	−3.202626	1.23951	−2.58	0.011*	−5.648747	−.756505
lnTOPOG	2.5334	7.913957	0.32	0.749	−13.08446	18.15126
_Cons	−542.4079	114.7133	−4.73	0.000*	−768.7897	−316.026
Diagnostic tests						
.. 2.23939(0.6723)						
$\chi^2_{Heteroskedasticity}$ 236.17(0.0000)*						
$\chi^2_{Normality}$ 0.0371(0.0000)*						
χ^2_{RESET} 151.80(0.0000)*						

Source: Authors' Computation.
Notes: * indicates significance at 1% level, and the values in parentheses in the lower part of the table are *p*-values of the respective tests.

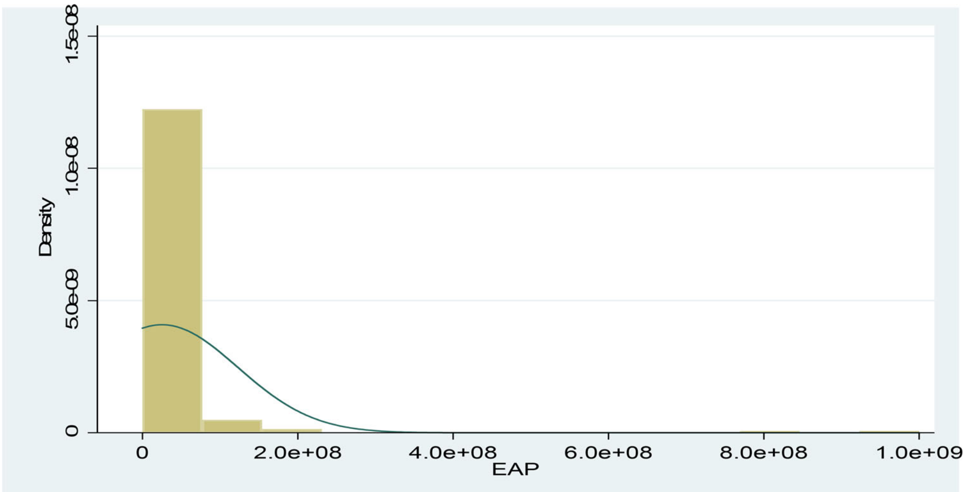

Figure 1. Graphical representation of the normality test.
Source: Authors' Depiction.

Table 4 displays the multicollinearity test of the model. The table tells that according to the variance inflation factor (*VIF*), none of the explanatory variables displayed evidence of multicollinearity as the VIF is less than 10 for each of the variables, and looking at the mean VIF, which is also less than 10, it can be stated that the model has no evidence of multicollinearity.

Figure 3 illustrates the test of linearity of the model. However, only the first row, where *y* is the dependent variable, is our concern. The graph shows that the relationship between the independent variables and the dependent variable is not on a straight line in each column of the first row except that of the countries' income classification. Therefore, the relationship between the dependent variable and independent variables of the model fails to meet linearity[1].

Figure 4 shows the leverage by the residual squared of the model. From the figure, the United States, Syria, Guam, and North Korea have observations with large

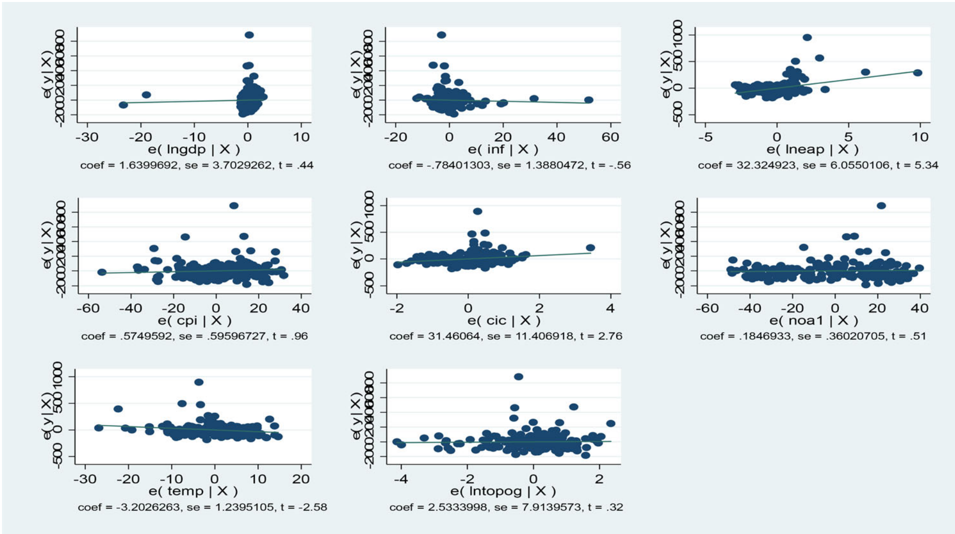

Figure 2. Outlier test.
Source: Authors' Depiction.

Table 4. Multicollinearity test.

Variable	VIF	1/VIF
lnGDP	2.15	0.465076
INF	2.02	0.495570
lnEAP	1.98	0.504005
CPI	1.84	0.542974
CIC	1.38	0.723822
NOA1	1.23	0.813840
TEMP	1.09	0.918555
lnTOPOG	1.05	0.953632
Mean VIF	1.59	

Source: Authors' Computation.

residual and large leverage; thus, they have observations that are potentially the most influential and outliers at the same time. Therefore, such countries' observations might greatly affect the integrity of the model coefficients.

So, the diagnostic tests conducted on the model, revealed that the model is not statistically healthy and therefore, unreliable. Hence, the analysis will further employ the planned alternative model to the multiple regression model i.e., Quantile Regression model which is not sensitive to the identified problems above that make the estimate of the multiple regression model unreliable.

Table 5 reports the result of the quantile regression model. From the table, the coefficient of the impacts of *lnGDP, CPI, and lnTOPOG* are not statistically significant throughout the quantiles, while that of *INF*, *lnEAP*, *CIC*, *NOA*, and *TEMP* are statistically significant.

The coefficient of the impact of *INF* on *Y* is only significant in 0.5 quantiles at 5% level and is positively related with *Y*, where the coefficient shows that in countries with medium medal ranking (medal ranking points up to 50% of the total medals won), a 1% increase in *INF* has an approximately 1-point impact on the medal ranking.

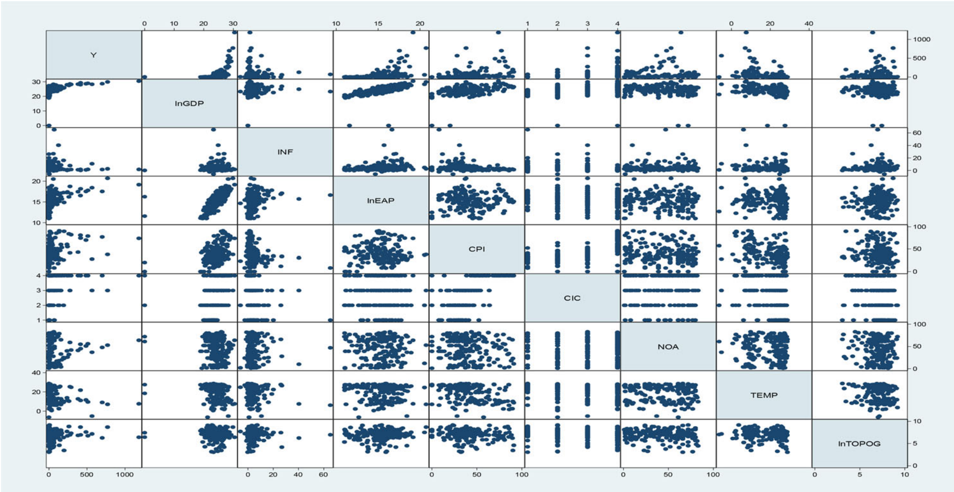

Figure 3. Linearity test.
Source: Authors' Depiction.

The coefficients of *EAP's* impact on *Y* in all the quantiles (0.25, 0.5, 0.75, and 0.9 quantiles) are significant at 5%, 1%, 5%, and 1% levels, respectively and all are positively related with *Y*; however, the coefficients continue to rise in each quantile from $\hat{\beta}_{0.25} = 2.905535$ to $\hat{\beta}_{0.5} = 7.525305$ to $\hat{\beta}_{0.75} = 13.19343$ and to$\hat{\beta}_{0.9} = 32.48525$. Therefore, in countries with low medal ranking (0.25 quantiles), a 1% increase in *EAP* has approximately 3 points impact on the medal ranking; in countries with moderate medal ranking (0.5 quantiles), a 1% increase in *EAP* has approximately 8 points impact on the medal ranking; in countries with high medal ranking (0.75 quantiles), a 1% increase in *EAP* has approximately 13 points impact on the medal ranking; and in countries with very high medal ranking (0.9 quantiles), a 1% increase in *EAP* has approximately 32 points impact on the medal ranking.

The coefficients of *CIC*'s impact on *Y* in all the quantiles (0.25, 0.5, 0.75, and 0.9 quantiles) are all significant at 5%, 1%, 1%, and 1% levels, respectively and all are positively related with *Y*; however, the coefficients continue to rise in each quantile from $\hat{\beta}_{0.25} = 4.26786$ to $\hat{\beta}_{0.5} = 10.56827$ to $\hat{\beta}_{0.75} = 13.19343$ and to$\hat{\beta}_{0.9} = 37.86277$. So, in countries with low medal ranking (0.25 quantiles), a unit increase in the index of *CIC* has approximately 4 points impact on the medal ranking; in countries with medium medal ranking (0.5 quantiles), a unit increase in the index of *CIC* has approximately 11 points impact on the medal ranking; in countries with high medal ranking (0.75 quantiles), a unit increase in the index of *CIC* has approximately 13 points impact on the medal ranking; and in countries with very high medal ranking (0.9 quantiles), a unit increase in the index of *CIC* has approximately 38 points impact on the medal ranking.

The coefficient of the impact of *NOA* on *Y* is only significant in 0.5 quantiles at 10% level and is negatively related with *Y*, where the coefficient shows that in countries with moderate medal ranking (0.5 quantiles), a 1% increase in *NOA* has an approximately 0-point impact on the medal ranking; thus, *NOA* has no impact on Y.

The coefficient of the impact of *TEMP* on *Y* is only significant in 0.5 quantiles at 5% level and is negatively related with *Y*, where the coefficient shows that in

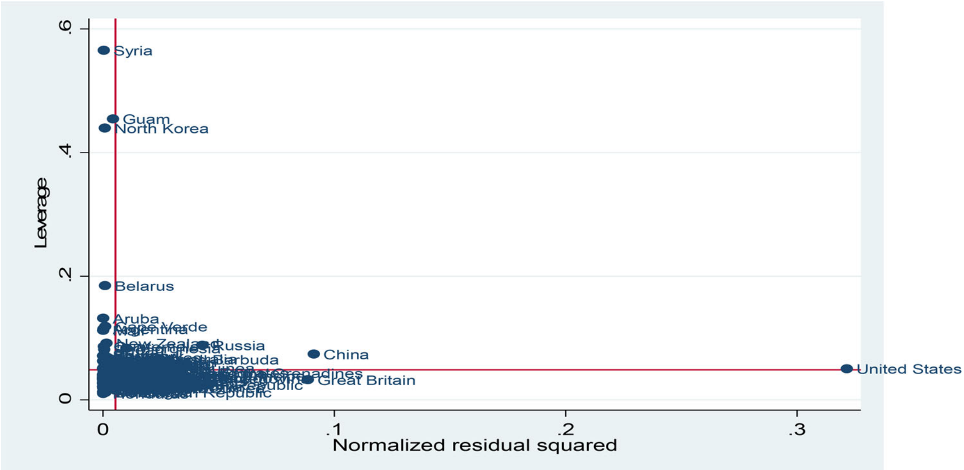

Figure 4. Leverage effect test.
Source: Authors' Depiction.

countries with moderate medal ranking, a 1% increase in *TEMP* has an approximately 1-point impact on the medal ranking.

The Durbin-Wu-Hausman test for endogeneity shows that the independent variables are not correlated with the errors of the overall quantile regression model since the *p*-value of the test statistic is not significant.

The analysis will now proceed with the estimation of the Tobit regression model for the purpose of checking the robustness of the results presented by the quantile regression model. However, the essence of choosing the Tobit regression model is that the dependent variable (*Y*) has some few countries with zero score and thus, zero observations and in such instance, following the work of Amore and Murtinu (2021); Tobit models can be used for robustness check of the estimated quantile regression model.

Table 6 reports the results of the Tobit regression model, and considering the *p*-value of the Likelihood Ratio (LR) Chi-Square test, which is significant at 1% level, it can be attested that the model is statistically healthy. However, from the table, it revealed that the coefficients of inflation rate (*INF*), the economic active population (*lnEAP*), countries' income classification (*CIC*), and temperature (*TEMP*) are all statistically significant at 5%, 1%, 1%, 1%, and 1% levels, respectively, and all are positively related with *Y* except *TEMP* which is negatively related with it. This is similar to what the study found when applied the quantile regression model. Therefore, the Tobit results confirmed the quantile regression results.

5. Discussion of the results

The findings of this study were investigated using a multiple regression model as a benchmark for cross-sectional data analysis, followed by a quantile regression model as the choosing model for the analysis, and then a Tobit regression model for checking the robustness of the choosing model. However, when comparing the results of

Table 5. Quantile regression results[2].

Quantile Regression Results at Q(0.25)						
Y	Coef.	Std. Err.	T	$p>$t	[95% Conf.	Interval]
lnGDP	.5248102	.7870128	0.67	0.506	−1.028326	2.077946
lnGDP	.5531322	1.462714	0.38	0.706	−2.333471	3.439735
lnGDP	1.546726	2.031634	0.76	0.447	−2.462617	5.556069
lnGDP	.0440504	.3166486	0.14	0.890	−.580842	.6689428
INF	.1005914	.3547055	0.28	0.777	−.5994046	.8005874
lnEAP	2.905535	1.325789	2.19	0.030**	.2891471	5.521924
CPI	.0881392	.0717933	1.23	0.221	−.0535418	.2298203
CIC	4.26786	1.828718	2.33	0.021**	.6589634	7.876756
NOA	−.0319075	.0548265	−0.58	0.561	−.1401051	.0762902
TEMP	−.243061	.1937099	−1.25	0.211	−.6253392	.1392171
lnTOPOG	−.7919844	1.302579	−0.61	0.544	−3.362567	1.778599
_Cons	−42.87139	21.5822	−1.99	0.049*	−85.46294	−.2798398
Quantile regression results at Q(0.5)						
INF	.6456179	.2733843	2.36	0.019**	.1061057	1.18513
lnEAP	7.525305	1.947107	3.86	0.000*	3.682773	11.36784
CPI	.0588453	.1907717	0.31	0.758	−.3176344	.435325
CIC	10.56827	3.187838	3.32	0.001*	4.277206	16.85933
NOA	−.2066727	.117059	−1.77	0.079***	−.4376837	.0243382
TEMP	−1.194007	.4805039	−2.48	0.014**	−2.142261	−.2457532
lnTOPOG	.5248102	.7870128	0.67	0.506	−1.028326	2.077946
_Cons	.6456179	.2733843	2.36	0.019**	.1061057	1.18513
Quantile regression results at Q(0.75)						
INF	−.0034454	.7834728	−0.00	0.996	−1.549595	1.542705
lnEAP	13.19343	5.151177	2.56	0.011**	3.027802	23.35906
CPI	.3316661	.6342813	0.52	0.602	−.9200609	1.583393
CIC	23.69873	7.950098	2.98	0.003*	8.009548	39.3879
NOA	.0310224	.3552632	0.09	0.931	−.6700743	.7321191
TEMP	.5531322	1.462714	0.38	0.706	−2.333471	3.439735
lnTOPOG	−.0034454	.7834728	−0.00	0.996	−1.549595	1.542705
_Cons	13.19343	5.151177	2.56	0.011**	3.027802	23.35906
Quantile regression results at Q(0.9)						
INF	−3.138702	2.188169	−1.43	0.153	−7.45696	1.179557
lnEAP	32.48525	7.805089	4.16	0.000*	17.08224	47.88826
CPI	.8150898	1.076632	0.76	0.450	−1.309597	2.939776
CIC	37.86277	11.3644	3.33	0.001*	15.4356	60.28993
NOA	.763364	.6209743	1.23	0.221	−.4621022	1.98883
TEMP	−8.901559	6.450398	−1.38	0.169	−21.63114	3.828025
lnTOPOG	.6433796	15.87251	0.04	0.968	−30.68035	31.9671
_Cons	−333.1312	314.0678	−1.06	0.290	−952.9305	286.6681
The Durbin-Wu-Hausman test for endogeneity:						
F(1, 176) = 7.5e + 14						
p-vaue = 0.4501						

Source: Authors' Computation.
Notes: *, **, and *** indicate significance at 1%, 5%, and 10% levels respectively and the values in parentheses are *p*-values.

the quantile regression model and the Tobit regression model, both provide the same results.

The results of the study revealed that out of all the variables analyzed, namely gross domestic product (*GDP*), inflation (*INF*), active economic population (*EAP*), corruption perception index (*CPI*), countries' income classification (*CIC*), number of athletes (*NOA*), temperature (*TEMP*), and topography (*TOPOG*); *lnGDP, CPI, and lnTOPOG* are not statistically significant throughout the quantiles, while *INF*, *lnEAP*, *CIC*, *NOA*, and *TEMP* are statistically significant, where inflation was found to be influencing the performance of the moderate performed countries in the Olympic Games. Though this is one of the few studies that used inflation, this finding

Table 6. Tobit regression results.

Y	Coefficient	Std. Err.	*t*	$p > \mid t \mid$	[95% Conf. Interval]	
lnGDP	.4707338	1.158417	0.41	0.685	−1.815264	2.756731
INF	1.086714	.4981098	2.18	0.030**	.1037535	2.069674
lnEAP	13.00175	2.10954	6.16	0.000*	8.838828	17.16468
CPI	.1249065	.1924898	0.65	0.517	−.2549492	.5047622
CIC	17.04983	3.920196	4.35	0.000*	9.31379	24.78587
NOA	−.1526814	.1096254	−1.39	0.165	−.3690141	.0636513
TEMP	−1.40681	.3751277	−3.75	0.000*	−2.14708	−.666540
lnTOPOG	−2.548608	2.391211	−1.07	0.288	−7.267378	2.170163
_Cons	−200.315	38.72213	−5.17	0.000*	−276.7286	−123.902
Log likelihood = −406.15727						
LR chi^2 (8) = 146.50 (*p*-value = 0.0000)*						

Source: Authors' Computation.
Notes: * and ** indicate significance at 1% and 5% levels, respectively.

contrasts with that of Rosas and Flegl (2019) in Rio 2016, who found it insignificant. The possible reason could be that of their use of OLS technique, which this study found it to be inferior. However, the finding is in line with the literature that since a number of sports professionals sign contracts in currencies other than their home countries, then athletes from countries with a considerable rate of inflation will receive more from the Olympic proceeds in their local currency. Furthermore, the study found the economically active population to be influencing the performance of the countries that performed low, moderate, high, and very high in the Olympic Games. This finding is in line with the traditional GDP-population based theory of Olympic success, also, the assertion that a large population means a more fabulous pool of talent to choose from, and to the game performance, additionally, it is in line with the findings of Bosscher et al. (2008) in Athens 2004, Luiz and Fadal (2011) in Beijing 2008, Feizabadi et al. (2013) in Guangzhou 2010 Summer Asian Games, Sharma (2015) in junior national level weightlifters from India's rural and urban areas, Rosas and Flegl (2019) in Rio 2016. Similarly, the study found that the countries' income classification is inducing the performance of the countries that performed low, moderate, high, and very high in the Olympic Games. Like in the case of incorporating inflation as part of the Olympic Games success determinants, though this study is one of the few that makes use of the countries' income classification, the finding is in line with that of Rosas and Flegl (2019) in Rio 2016, and in line with the literature that high-income countries have an advantage in sporting competitions by earning more medals than low-income countries. More so, the study found that temperature is inversely related to the performance of the countries that performed moderate, thus, the higher the temperature of a country the lower it is likely to perform better in the Olympic Games. This finding is in contrast with the finding of Luiz and Fadal (2011) in Beijing 2008 who found the climate to be insignificant. The possible reason could be their use of the OLS technique which this study found to be inferior, and the literature understanding that climate change will not only affect sports circumstances but will also interfere with athletes' physiological mechanisms and emotions.

Yet, when observing the results in the various quantiles, it can be deduced that the impact of the factors changed with different quantiles, which means that the results depend on the level of a country model ranking in the Olympic Games.

6. Conclusion and practical implication

Sport is an important sector of economic activity and spans medical treatment and rehabilitation, research and development, sports tourism, sales and trade of sports products, construction and maintenance of sports venues, organization of sports events, marketing and advertising. However, sports performance is being rewarded with medals, but why do some countries win medals while others do not or why do some countries win more medals than others? This makes scholars argue that a country's success in sports is directly related to its economic, demographic, geographical, and social factors. Therefore, this study used cross-sectional data to investigate such claims by specifically answering the following questions: do a country's economic factors such as gross domestic product and inflation impact its performance in Olympic Games? Do a country's demographic factors, such as active population, influence its performance in Olympic Games? Do a country's social factors such as corruption level, countries' income classification, and the number of athletes affects its performance in Olympic Games? Are geographical factors such as temperature and topography impacting a country's performance in the Olympic Games? Finally, do the influences of the factors change with different quantiles of a country's Olympic Games performance?

To conduct this study, recent Olympic Games of Rio 2016 from 207 countries with more than 11,000 athletes on 306 events in 28 different sports through the use of weighted medal ranking as the dependent variable, and the independent variables including gross domestic product, economic active population, corruption perception index, countries' income classification, rate of inflation, the number of athletes, temperature, and topography; were used. However, the estimation techniques employed include Quantile Regression model and Tobit Regression model where the findings revealed that inflation rate in moderate performed countries; economically active population in low, moderate, high, and very high performed countries; and countries' income classification in low, moderate, high, and very high performed countries are influencing the countries medal ranking performance in the Olympic Games. Furthermore, high temperature in moderate performed countries lowers the performance of the countries. However, the size of a country's GDP level, corruption ranking level, the number of athletes, and topographical nature of a country have no impact on the countries medal ranking performance in the Olympic Games. Therefore, growth in the economically active population, countries' income classification, and inflation would increase the participants' weighted medal ranking; nations with a considerable rate of inflation, high population, and high index of income classification have a better chance of winning more medals, while countries with high temperature are not likely to do well in the games. Consequently, the practical implication of the results is that participant countries should target growth in the economically active population, countries' income classification, and inflation rate for improving their performance in the Olympic Games. Also, analysts should be considering the gage of such factors in the various countries while forecasting/predicting the outcome of the Olympic Games. Additionally, being inflation is a contributing factor, international athletes should be aware of currency patterns in the country where they are now based and compared with the currency trends in their home country. Moreover, sports policymakers should place a high priority on grass-roots participation and assist athletic organizations to prioritize retention concerns that arise during adolescence. Finally, rigorous modelling could be used to replace traditional regression methods in investigating the determinants of Olympic Games performance, perhaps on a sub-national or micro-level in order to exploit more ways/information for better performance in the Olympic Games.

The limitation of this study is the lack of data for each of the countries on the money allocated by each country for sports; thus, the investment in the sports sector and the institutional support to the sports sector, hence unable to test for their effect on the Olympic success.

The study suggests further studies to study the performance based on different categories of the Olympic Games with respect to the medal ranking. More so, a study can also be executed based on the characteristics of the athletes in relation to the medal ranking. In addition, studies to come could add more determinants with the support of theoretical literature.

Notes

1. According to Dye (2020), quantile regression is an extension of linear regression that is used when the conditions of linear regression (such as linearity, homoscedasticity, independence, or normality) are not met.
2. All the model coefficients of 0.1 quantiles are zero (0); therefore, the study did not record the model estimate at 0.1 quantiles.

Acknowledgements

We thank the anonymous reviewers for their careful reading of our manuscript and their many insightful comments and suggestions.

Authors' contributions

Wang Shasha conceived of the presented idea, Babar Nawaz Abbasi designed the model and analyzed the data, Ali Sohail carried out the implementation and all authors drafted the manuscript.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Wang Shasha http://orcid.org/0000-0002-4243-3168
Babar Nawaz Abbasi http://orcid.org/0000-0002-2298-339X
Ali Sohail http://orcid.org/0000-0002-3718-067X

References

Amore, M. D., & Murtinu, S. (2021). Tobit models in strategy research: Critical issues and applications. *Global Strategy Journal*, *11*(3), 331–355. https://doi.org/10.1002/gsj.1363

Andreff, W. (2001). The correlation between economic underdevelopment and sport. *European Sport Management Quarterly*, *1*(4), 251–279. https://doi.org/10.1080/16184740108721902

Baimbridge, M. (1998). Outcome uncertainty in sporting competition: The Olympic Games 1896-1996. *Applied Economics Letters*, *5*(3), 161–174. https://doi.org/10.1080/758521374

Barnesa, M. L., & Hughes, A. W. (2002). *A quantile regression analysis of the cross-section of stock market returns*. Federal Reserve System.

Bernard, A. B., & Busse, M. R. (2004). Who wins the Olympic Games: Economic resources and medal totals. *Review of Economics and Statistics*, *86*(1), 413–511. https://doi.org/10.1162/003465304774201824

Bosscher, V. D., Heyndels, B., Knop, P. D., Bottenburg, M. V., & Shibli, S. (2008). The paradox of measuring success of nations in elite sport. *Belgeo*, (2), 217–234. http://journals.openedition.org/belgeo/10303; https://doi.org/10.4000/belgeo.10303

Chappelet, J. L. (2005). *Sport and economic development* (Vol. 20). IDHEAP Swiss Graduate School of Public Administration. http://assets.sportanddev.org/downloads/sport__economic_development_in_english

Churilov, L., & Flitman, A. (2006). Towards fair ranking of olympic achievements: The case of Sydney 2000. *Computers & Operations Research*, *33*(7), 2057–2082. https://doi.org/10.1016/j.cor.2004.09.027

Condon, E., Golden, B. L., & Wasil, A. E. (1999). Predicting the success of nations at the summer Olympics using neural networks. *Computers & Operations Research*, *26*(13), 1243–1265. https://doi.org/10.1016/S0305-0548(99)00003-9

Cunillera, O. (2014). Tobit models. In Michalos A. C. (Ed.), *Encyclopedia of quality of life and well-being research*. Springer. https://doi.org/10.1007/978-94-007-0753-5_3025

Davis, M., & End, C. (2010). A winning proposition: The economic impact of successful National Football League franchises. *Economic Inquiry*, *48*(1), 39–50. https://doi.org/10.1111/j.1465-7295.2008.00124.x

Debroy, B. (2011). *Does GDP growth influence sporting performance?* https://economictimes.indiatimes.com/opinion/et-editorial/does-gdp-growth-influence-sportingperformance/articleshow/7851900.cms?utm_source=contentofinterest&utm_medium=text&utm_campaign=cppst

Dye, S. (2020). *Quantile regression*. https://towardsdatascience.com/quantile-regression-ff2343c4a03

End, C. M., Dietz-Uhler, B., Harrick, E. A., & Jacquemotte, L. (2002). Identifying with winners: A reexamination of sport fans' tendency to BIRG 1. *Journal of Applied Social Psychology*, *32*(5), 1017–1030.

Feizabadi, M. S., Khabiri, M., & Hamid, M. (2013). The relationship between the success of countries at the Guangzhou 2010 summer Asian games and demo-economic factors. *Procedia - Social and Behavioral Sciences*, *82*(2013), 369–374. https://doi.org/10.1016/j.sbspro.2013.06.277

Flegl, M. (2014). Performance analysis during the 2014 FIFA World Cup qualification. *The Open Sports Science Journal*, *7*, 183–197. http://dx.doi.org/10.2174/1875399X01407010183

Flom, P. (2018). An introduction to quantile regression-Towards data science. https://towardsdatascience.com/an-introduction-to-quantile-regression-eca5e3e2036a

Foster, L., James, D., & Haake, S. (2010). 8th Conference of the International Sports Engineering Association (ISEA). Understanding the influence of population size on athletic performance. *Procedia Engineering*, *2*(2), 3183–3189. https://doi.org/10.1016/j.proeng.2010.04.130

Geraci, M. (2014). Linear quantile mixed models: the lqmm package for Laplace quantile regression. *Journal of Statistical Software*, *57*, 1–29. https://doi.org/10.18637/jss.v057.i13

Gorse, S., & Chadwick, S. (2011). *The prevalence of corruption in international sport: A statistical analysis*. Coventry: Centre for the International Business of Sport. http://www.egba.eu/pdf/Report-FINAL.pdf

Hoffmann, R., Ging, L. C., & Ramasamy, B. (2002). The socio-economic determinants of international soccer performance. *Journal of Applied Economics*, *5*(2), 253–272. https://doi.org/10.1080/15140326.2002.12040579

Hoffmann, R., Ging, L. C., & Ramasamy, B. (2004). Olympic success and ASEAN countries: Economic analysis and policy implications. *Journal of Sports Economics*, *5*(3), 262–276. https://doi.org/10.1177/1527002503261826

Huang, Q., Zhang, H., Chen, J., & He, M. (2017). Quantile regression models and their applications: A review. *Journal of Biometrics & Biostatistics*, *08*(03), 1000354. https://doi.org/10.4172/2155-6180.1000354

Ibrahim, L. Y. (2016). Integrity issues in competitive sports. *Journal of Sports and Physical Education*, *3*(5), 67–72. https://doi.org/10.9790/6737-03056772

Introduction to SAS. *UCLA: Statistical Consulting Group*. https://stats.idre.ucla.edu/stata/output/tobitregression/#:~:text=To%20generate%20a%20tobit%20model,specified%20in%20parentheses%20after%20ul.

Johnson, D. K., & Ali, A. (2004). A tale of two seasons: Participation and medal counts at the summer and winter Olympic Games. *Social Science Quarterly*, *85*(4), 974–993. https://doi.org/10.1111/j.0038-4941.2004.00254.x

Kiviaho, P., & Makela, P. (1978). Olympic success: A sum of non-material and material factors. *International Review of Sport Sociology*, *13*(2), 5–22. https://doi.org/10.1177/101269027801300201

Koenker, R., & Hallock, K. (2001). Quantile regression an introduction. *Journal of Economic Perspectives*, *15*(4), 143–156. https://doi.org/10.1257/jep.15.4.143

Kuper, G. H., & Sterken, E. (2001). *Olympic participation and performance since 1896*. Department of Economics.

Leeds, M. A. (2014). Quantile regression for sport economics. *International Journal of Sport Finance*, *9*(4), 346–359.

Li, Y., Xiyang, L., Qianzhi, D., & Liang, L. (2015). Performance evaluation of participating nations at the 2012 London Summer Olympics by a two-stage data envelopment analysis. *European Journal of Operational Research*, *243*(3), 964–973. https://doi.org/10.1016/j.ejor.2014.12.032

Lozano, S., Villa, G., Guerrero, F., & Cortés, P. (2002). Measuring the performance of nations at the Summer Olympics using data envelopment analysis. *Journal of the Operational Research Society*, *53*(5), 501–511. https://doi.org/10.1057/palgrave/jors/2601327

Luiz, J. M., & Fadal, R. (2011). An economic analysis of sports performance in Africa. *International Journal of Social Economics*, *38*(10), 869–883. https://doi.org/10.1108/03068291111170415

Machado, J. A. F., & Silva, J. (2013). Quantile regression and heteroskedasticity. https://www.semanticscholar.org/paper/Quantile-regression-and-heteroskedasticity-Machado-Santos/98b316c77fd8732ea9249514c9c60a106ad798ed

Matros, A., & Namoro, S. D. (2004). *Economic incentives of the Olympic Games*. University of Pittsburgh.

Mendoza, R. G. (2017). *Football and economy relations at the international level.* [Ph.D. thesis]. Universitat De Barcelona.

Miller, S. J. (2006). Methods of least squares. *Statistics Theory, Cornell University, USA*, *3*, 1–2.

Money Smart Athlete Blog. (2018). *Economic environment*. https://moneysmartathlete.com/2018/08/08/unemployment-inflation-and-currency-values-and-their-effect-on-the-money-smart-athlete/

Nauright, J. (2004). Global games: Culture, political economy and sport in the globalised world of the 21st Century. *Third World Quarterly*, *25*(7), 1325–1336. https://doi.org/10.1080/0143659042000281302

Ramírez, L., & Sánchez, I. (2013). Crecimiento económico, corrupción e instituciones en México. *Nóesis. Revista de Ciencias Sociales y Humanidades*, *22*(43-1), 104–133. http://dx.doi.org/10.20983/noesis.2013.1.4

Rio. (2016). Rio 2016 – athletics schedule and results. https://www.rio2016.com/en/athletics-schedule-and-results-download

Rosas, L. A. A., & Flegl, M. (2019). Quantitative and qualitative impact of GDP on sport performance and its relation with cocrruption and other social factors. *Nóesis. Revista de Ciencias Sociales y Humanidades*, *28*(1), 15–37. https://doi.org/10.20983/noesis.2019.1.2

Sen, S. (2021). *Does economy determine a country's performance at Olympics?* https://www.thehindu.com/data/data-does-economy-determine-a-countrys-performance-at-olympics/article35899178.ece

Sharma, R. (2015). Effect of socio-economic status on sport performance of national level junior weightlifters. *International Journal of Applied Research*, *1*(5), 212–214.

Skarbalius, A., Vidunaite, G., Kniubaite, A., Reklaitiene, D., & Simanavicius, A. (2019). Importance of sport performance monitoring for sports organization. *Transformations in Business & Economics*, *18*(2), 279–303.

Song, M. Y., & Zhang, Y. (2018). Research on the relationship between geographical factors, sports and culture. *Advances in Physical Education*, *08*(01), 66–70. https://doi.org/10.4236/ape.2018.81008

Tcha, M., & Pershin, V. (2003). Reconsidering performance at the summer Olympics and revealed comparative advantage. *Journal of Sports Economics*, *4*(3), 216–239. https://doi.org/10.1177/1527002503251636

Transparency International. (2016). *Global corruption report: Sport*. Routledge. http://www.transparency.org/news/feature/sport_integrity

Vagenas, G., & Vlachokyriakou, E. (2012). Olympic medals and demo-economic factors: Novel predictors, the ex-host effect, the exact role of team size, and the "population-GDP" model revisited. *Sport Management Review*, *15*(2), 211–217. https://doi.org/10.1016/j.smr.2011.07.001

Wooldridge, J. M. (2002). *Econometric analysis of cross section and panel data*. MIT Press.

World Bank. (2019). *World Bank open data*. World Bank. http://data.worldbank.org/

Wu, J., Liang, L., & Feng, Y. (2009). Achievement and benchmarking of countries at the Summer Olympics using cross efficiency evaluation method. *European Journal of Operational Research*, *197*(2), 722–730. https://doi.org/10.1016/j.ejor.2008.06.030

Zhang, S., Gomez, M. A., Yi, Q., Dong, R., Leicht, A., & Lorenzo, A. (2020). Modelling the relationship between match outcome and match performances during the 2019 FIBA Basketball World Cup: A quantile regression analysis. *International Journal of Environmental Research and Public Health*, *17*(16), 5722. https://doi.org/10.3390/ijerph17165722