Meeting expectations at the 2016 Rio Olympic Games: Country potential and competitiveness*

F. Javier Otamendi, Universidad Rey Juan Carlos

franciscojavier.otamendi@urjc.es

Luis Miguel Doncel, Universidad Rey Juan Carlos

luismiguel.doncel@urjc.es

Cecilia Martín-Gutiérrez, Junta de Castilla y León

cecilia.martin@jcyl.es

Abstract

Objective: Develop a framework to measure the extent to which countries meet their performance expectations at major sporting events using economists, experts and fan knowledge. *Method*: Long-term expectations based on socioeconomic potential are calculated using sport-agnostic econometric modelling. Short-term expectations based on performance and competitiveness are calculated using betting odds, which incorporate both expert-knowledge and "wisdom of the crowd". Robust statistics based on the chi-squared distribution are used to measure the extent to which countries' expectations are met. Our method is tested using real data from the Rio 2016 Summer Olympic Games. *Results*: Twelve countries over- or under-performed in terms of meeting either long- or short-term expectations. *Conclusions*: Uzbekistan was the only country that exceeded both long- and short-term expectations.

1. Introduction

When a sporting event takes place, athletes must be ready to perform and meet expectations. Athletes face huge pressure to succeed at major events such as the Olympic Games and the FIFA World Cup. These athletes spend years and invest massive amounts of time and money to prepare themselves physically and mentally to excel at the right place and time.

Similarly, governments develop sport policies that aim at improving athletes' performance at these major sporting events. Governments invest in state-of-the-art training facilities, hire the best coaches and even persuade athletes to change their sport nationality. Money is spent to host international competitions or help athletes participate in events abroad. Governments also seek private investment or reduce taxes for companies that invest in sport.

The primary goal of this study is to establish a knowledge-based framework that quantitatively calculates whether athletes meet the expectations of countries, governments and citizens, based on two types of publicly available data sources: on the one hand, socio-economic data of countries as indicators of country potential; on the other, betting odds based on expert or fans knowledge as indicators of individual performance and competitiveness at past or present events. We aim at analysing the results of major sporting events, not in terms of absolute success, but rather in terms of whether the results objectively meet expectations in relative terms.

For example, the USA and Great Britain are likely to win a large number of medals at the Summer Olympics. The question is whether they will obtain more or less than expected based on different knowledge-based predictions: socio-economic, sport experts and fans. The same is true of the 2018 FIFA World Cup. Germany is not necessarily favourite simply because Germany won the 2014 FIFA World Cup. The question is whether the only way to meet expectations in 2018 is to win or whether Germany and its citizens will be happy if the national team reaches the semi-finals. The fans were shocked by the premature elimination of Germany at the group stage. Unsurprisingly, therefore, any African country would be happy to reach the quarter-finals because this result would exceed expectations. Indeed, no African country reached the knock-out stage.

Our underlying research hypothesis is that performance at major sporting events is primarily the result of two axes that can be measured at the country level. The first axis relates to long-term socioeconomic indicators such as gross domestic product (GDP), population and expenditure on sport. Accordingly, athletes from countries with a strong economy are more likely to win than are athletes from poorer countries. Training facilities are state-of-the-art, coaches have a long history of success and athletes compete more regularly on the international stage. These richer and larger countries also have more resources to detect talent or recruit foreign athletes who are willing to change their sport nationality. The second axis relates to short-term performance, which refers to the ability of the athletes to compete as expected during the event. Some athletes perform well, whereas others choke under pressure. Being the favourite may actually hurt an athlete's chances of success. In both cases, the type of sport and the form of competition are crucial. Training and competing in a team sport such as basketball is quite unlike training and competing in individual sports. For example, in tennis, competitions are based on knockout draws. Another example, the triathlon, consists of a one-day event that comprises three disciplines.

In this study, we measure how well national teams, which are selected by National Olympic Committees (NOCs), meet expectations. To forecast expected long-term potential and preparation, we use econometric modelling. For short-term performance, we use betting odds. The novelty of this study lies in the fact that instead of seeking to correctly forecast the medal haul, we use the deviations in results from long- and short-term expectations to determine abnormal country performance in an attempt to stimulate the debate on sport policy. Meeting expectations is calculated as the relative squared deviation of expectations from actual results.

To calculate expectations based on country preparation, we forecast the number of medals that the country should win based on its socioeconomic potential on the underlying hypothesis that socioeconomic factors favour athletes from certain countries. Clarke (2000) stated a large home advantage or host country effect by the fatigue of the travel and local conditions acclimatization. Moreover, it could be a court advantage by influencing judges (Garicano, Palacios & Prendergast, 2001).

Kuper & Sterken (2001) focused on economic, geographic and demographic explanations, such as national sport culture or political systems, of Olympic medal success, whereas Tcha & Pershing (2003) applied a Tobit model to study the concept of "Revealed Comparative Advantage" to determine countries' specialization in the Olympic Games. They found that high-income countries win medals in a more diversified range of sports while low-income countries show less specialization.

Bernard & Busse (2004) constitute a reference in demonstrating the forecasting of Olympic performance based on countries' economic status and political regimes. They included socioeconomic variables such as gross domestic product, population, sporting tradition and weather into a Tobit model. Jiang & Xu (2005) applied the specifications of Bernard & Busse (2004) to the Chinese transition period from a planned to a market regime. They added education, inequality and government spending as explanatory variables finding a positive effect in explaining medal Olympic performance due to income level, population, home, and in the case of a market regime, the Government spending.

Forrest, Sanz & Tena (2010) included government spending on recreation in order to improve the forecasting performance. In the same vein, Blaiss-Morisset, Boucher & Fortin (2017) not only incorporated the impact of Government spending in sport, but also estimated the money needed, in millions of dollars, to obtain an additional medal. Andreff, Andreff & Poupaux (2008) studied the effects of introducing variables capturing cultural differences by region, some inertia in winning medals, the individual chance of participating into an Olympic final and to be a medal's winner. Estimations showed the importance of the political regime and the consequence of the cultural differences by region.

The gender effect was analysed by Leeds & Leeds (2012) by distinguishing the determinants of success for male and female athletes. They found that this distinction provides clear clues about what could determine success. Despite this fact, in general terms, most factors have similar effects on performance. In the same way, athlete's female share is considered into a theoretical model simulating a production process to win Olympic medals in conjunction with other variables such as team size, countries with no distinction between civil governments and religious affairs or persistence (Trivedi & Zimmer, 2014). All these factors exhibited high significance. However, a random component remained into the process. The gender role is also incorporated by

Lowen, Deaner & Schmitt (2016) by using the Gender Inequality Index. The greater the index, the higher the Olympic participation and success in winning medals. This relationship applies for both women and male athletes.

On individual sports, Forrest, McHale, Sanz & Tena (2017) employ a random effects Tobit model to conclude that income plays a very significant role in sports with substantial requirements for specific capital equipment and as mentioned above, the host variable has a significant role in sports strongly influenced by judge decisions.

In short, 'sport-agnostic' (Vaz de Melo, Almeida, Loureiro, & Faloutsos, 2012) econometric modelling is used to predict outcomes based on socioeconomic variables rather than sports expert's knowledge. Long-term expected outcomes are calculated using forecasting techniques based on publicly available data for each International Olympic Committee (IOC) member state. We use GDP and population because detailed knowledge about sport policies for all participating countries is not readily available.

To calculate expectations based on athletes' performance, we forecast the number of medals that the country should win based on the odds set by major betting houses. The relationship between betting and sport, or most specifically, the Olympic Games existed from the time of Greek Olympics (Kiernan & Daley, 1965). Nowadays, the link varies from gambling-related problems in athletes (Martin, Nelson & Gallucci, 2016), or anti-gambling rules for athletes (Kindt & Asmar, 2001), to sport betting games as an analytic tool for prediction market (Sauer, 1998).

At this last point, a question arises. How good are these forecasting methods at predicting the outcome of sporting events? Can we trust these experts' forecasts or betting houses' odds, or are they just an 'illusion of control' (Cantinotti, Ladouceur, & Jacques, 2004; Khazaal et al., 2012). Scholars have studied the efficiency of sports betting markets by examining similarities between betting in these markets and trading in financial markets (Sauer, 1998, 2005; Williams, 1999, 2005). Market efficiency implies that all new information is immediately reflected in prices. Surowiecki (2004) demonstrated that in sport, an informed crowd outperforms an expert 90% of the time under the paradigm of the wisdom of crowds. Similarly, List (2012) showed that a crowd of diverse and differently informed individuals can outperform an expert in a specific field. Based on the economics of auctions (Klemperer, 1999, 2004), decentralised decision-making processes that use diverse information aggregate this information efficiently, doing so even better than bookmakers, who might fail to incorporate new information (Haan, Dijkstra, & Dijkstra, 2005). Seemann and Hungenberg (2008) obtained good market predictions by following the decisions of the crowd after analysing online bets on the 2006 FIFA World Cup. Similarly, Gramm and Owens (2005) found that when more horseracing bettors have all available public information, forecasts improve, and the market odds mimic the true odds. However, Forrest and Simmons (2008) and Levitt (2004) reported that bookmakers' forecasts differed from bettors' because bettors seek perfect predictions, whereas bookmakers calculate odds to meet consumer preferences or exploit fan sentiment. Smith, Paton, and Williams (2009) and Wise et al. (2010) provide a more thorough discussion of this topic.

Literature on sports betting has shown that odds can be used to predict the outcome of horse races (Ma, Tang, & McGroarty, 2016; Sung, McDonald, & Johnson, 2016), hockey (Cantinotti et al., 2004), tennis (McHale & Morton, 2011; Lisi & Zanella, 2017), the Ryder Cup (Maher, 2013), Olympics (Chen, Li & Zeng, 2011), and, of course, football (Andersson, Memmert, & Popowicz, 2009; Boulier & Stekler, 2003; Forrest, Goddard, & Simmons, 2005; Hvattum & Arntzen, 2010; Khazaal et al., 2012; Leitner, Zeileis, & Hornik, 2010; Štrumbelj & Šikonja, 2010). Constantly changing odds in all these sports are provided by bookmakers based on a combination of their expert knowledge and the wagers of fans, who place bets in the hope of beating the odds. The final quoted odds capture a myriad of available information, including subjective factors such as the reputation of participants (Herzog & Hertwig, 2011) and objective factors based on past results of similar games, the time and weather of the event and recent form. The bookmaker then combines expert and fan knowledge to set aggregate odds for each participant in a particular event. Thus, our hypothesis is that odds should be a good short-term predictor of the outcomes of sporting events and that betting predictions should be highly successful at picking likely winners.

In conclusion, on the one hand, the expert knowledge, and, on the other, the 'wisdom of crowds' (Munafo et al., 2015; Simmons, Nelson, Galak, & Frederick, 2011), is translated into bets, assuming that bettors and bookies seek economic rewards and indeed bet based on their knowledge of the form of the athletes in recent events. Athletes with the shortest odds are most likely to win.

Figure 1 summarises the proposed knowledge-based framework to analyse the behaviour of countries at major events. Figure 1 provides a graphical illustration of the statistical comparison between predictions and results to determine the failure to meet expectations. Figure 1 also depicts the validation stage, in which the long- and short-term expectations are statistically compared to demonstrate that both offer similar predictions, being equally important when comparing countries based on long- and short-term expectations. The results of the analysis can be shown in graphical format. Long-term deviations are shown in the abscissas, and short-term deviations are shown in the ordinates. Some countries lie outside the central square, whose size depends on the χ^2 distribution with 1 degree of freedom and the required confidence level using a minimum required expectation of 'u' predictions. These countries have over (under) prepared or over (under) performed and thus exceeded expectations or have failed to fulfil their socioeconomic potential or perform to the best of their supposed ability.

Insert Figure 1 about here

The research setting in this study is the Rio 2016 Summer Olympic Games, which took place in August 2016. This is the most comprehensive sporting event in the world, consisting of 41 sports that award medals in different disciplines. We will nevertheless discuss the applicability of the framework to other major events in the conclusions.

In Section 2, we describe the sports that formed part of Rio 2016 and propose a taxonomy of sports based on each sporting discipline. Section 3 presents the expectations based on long-term sport-agnostic econometric models. Section 4 presents the short-term odds-based expectations based on expert and fans knowledge. In Section 5, we test the forecasting precision of the long- and short-term expectations. In Section 6, we analyse the results to identify countries that failed to meet long- or short-term expectations. Section 7 concludes.

2. The Rio 2016 Summer Olympic Games

2.1 Observed outcome

This study's first contribution is the choice of research setting. To the best of our knowledge, this is the first time the Olympic Games, the most multi-disciplinary sporting event in the world, has been used to study how well expectations are met. We considered the 2016 Rio Summer Olympic Games, which took place between 5 and 21 August 2016. The Rio 2016 Summer Olympic Games was the biggest sporting event in the world in terms of number of participating athletes, in-situ spectators and television audience. More than 11,500 athletes from 206 countries participated in Rio 2016 in 306 events across 41 sports.

The dataset comprises the actual results for each event at the Rio 2016 Summer Olympic Games, including the names and countries of participants, their performance in time, length or points, and the colour of the medal they achieved. These data are available on the IOC website for each of the 306 events.

In total, 974 medals, each of which corresponded to an individual medallist, were officially awarded. In most events, three medals (gold, silver and bronze) were awarded. In some sports (e.g. Boxing, Judo, Taekwondo and Wrestling), four medals were conferred because bronze medals were awarded to each semi-finalist. There were also ties in swimming (double gold in Women's 100m Freestyle and triple silver in Men's 100m Butterfly) and a tie in sprint canoeing (double bronze in Men's 200m Kayak Single). Table 1 shows the official country medal count at Rio 2016 by sport class for countries that won or were predicted to win at least five medals. As expected, based on the importance of long-term preparation, many of the world's most socioeconomically powerful countries (USA with 121 medals, China with 70 and Great Britain with 67) occupy the top positions. Only 87 of the 206 competing countries won at least one medal[i]. 83 medals were awarded to countries that won less than five medals[ii].

The indicator that we used in this study was the number of medals awarded, regardless of medal colour. This indicator measured the performance of the country as a whole because more than one athlete per country could compete in a particular event. Furthermore, to conduct the analysis at the sport level, a reasonable sample size was important for ensuring statistical robustness. Whereas some sports awarded more than 100 medals (Athletics awarded 141 and Swimming awarded 104), others awarded only 6. These low medal totals added complexity to the detailed statistical analysis of certain sports.

Insert Table 1 about here

2.2. Taxonomy of sports

This study's second contribution is to provide a taxonomy of the 41 sports that formed part of Rio 2016. The purpose of this exercise was to enable the analysis of preparation and performance by ensuring that the sample was large enough. Sports have been classified according to uncertainty (Robles et al, 2009; Valero and Gómez-Mármol 2016). This uncertainty may be based on the number of rivals, contact in the sport, the use of a ball, the sport's individual or team nature, the existence of an opponent, and so forth. In this study, we classified Olympic sports based on a mixture of these classifications and competition format. This classification yielded 10 classes. The minimum sample size was 42 medals for one class.

Team sports (42 medals awarded) formed one class. Participants were national teams (one per country) that competed in a group format prior to a knockout phase that determined the medals (three per event). Basketball (6 medals), Football (6), Handball (6), Hockey (6), Rugby Sevens (6), Volleyball (6) and Water Polo (6) formed this 'Team' class. The importance of subjective judges (Noland and Stahler, 2017) defined a special class of sports. Artistic Gymnastics (42 medals), Diving (24), Equestrian Dressage (6), Rhythmic Gymnastics (6), Synchronised Swimming (6) and Trampoline Gymnastics (6) awarded 90 medals in 'Judge' sports. All cycling disciplines, including triathlon, were grouped together. Cycling BMX (6 medals), Cycling Mountain Bike (6), Cycling Road (12), Cycling Track (30) and Triathlon (6) awarded 60 'Cycling' medals. Most events were one-day events. Many events had knockout formats, allowing entry of up to three participants per country. Archery (12 medals), Badminton (15), Beach Volleyball (6), Fencing (30), Table Tennis (12) and Tennis (15) formed the 'Elimination' class, which awarded 90 medals. The 'Bouts' class (212 medals) consisted of events with knockout formats. However, a special class was required for these events because four rather than three medals were awarded per event. All ‘Bouts’ sports were contact sports: Boxing (52 medals), Judo (56), Taekwondo (32), Wrestling Freestyle (24) and Wrestling Greco-Roman (48). Canoeing and Rowing both followed the same competition format, with six finalists advancing from the classification and semi-final stages. Canoe Slalom (12 medals), Canoe Sprint (37) and Rowing (42) awarded 91 ‘Canoe’ medals. Another class consisted of sports where the competition was based on rounds. Points were awarded in each round and were then added to give a total for each participant. Equestrian Eventing (6 medals), Equestrian Jumping (6), Golf (6), Modern Pentathlon (6), Sailing (30) and Shooting (45) formed the ‘Rounds’ class (99 medals). Athletics (141 medals awarded) and Swimming (104) had their own classes. These sports consisted of a group of events. Most events consisted of a classification stage, a semi-final and a final of eight participants. Finally, Weightlifting (45) was a unique competition format that required its own class.

Our proposal therefore provides a taxonomy of 10 sport classes. Sports within each class resemble one another but differ from sports in the other classes. Other classifications are possible, but this classification is suitable for analysing expectations and provides a sufficient sample size of at least 42 medals per class.

Table 1 shows that certain countries dominate in certain sports. The USA excelled in swimming and athletics. The top four countries won the majority of the medals in ‘Judge’ sports. Jamaica won all of its medals in athletics. Kazakhstan, Uzbekistan, Cuba, Iran, Turkey and Georgia specialised in bouts. But did these countries perform as expected?

3: Expected socioeconomic potential: Sport-agnostic econometric modelling

In the era of the Internet of Things and Big Data, complex models based on artificial intelligence and machine learning tools (Schumaker, 2013) can be used to accurately predict the outcome of sporting events based on publicly available indicators. For example, scholars have used methods and models such as Monte Carlo simulation (Lahvicka, 2015) and Bayesian networks (Constantinou, Fenton, & Martin, 2013) in association football, time-series and clustering algorithms (Menéndez, Vázquez, & Camacho, 2015) in baseball, and point-process models (Baker & McHale, 2013) in the NFL.

One specific method that is used to forecast sporting event outcomes is econometric modelling. This forecasting method is built on the idea that the strength of the participating country’s economy is critical for predicting the winner of the event. Which country has more world champions: the USA, China or Kenya? Or perhaps the G20 member states?

The literature contains several references on this topic as previously stated. Some refer to the Winter Games (Kuper & Sterken, 2010; Andreff, 2013; Otamendi & Doncel, 2014a; Pfau, 2006), others to the Summer Games (Forrest, Sanz, & Tena, 2010; Otamendi & Doncel, 2014b) and others to both editions of the Games (Johnson & Ali, 2004). We adopted a model that predicts the outcome for each sport. Other models predict the overall outcome by country. The model that we used was the same model that was used to predict the results at the London 2012 Summer Olympic Games (Otamendi & Doncel, 2014b)[iii] and the results at the 2014 Sochi Winter Olympic Games (Otamendi & Doncel, 2018). The model is given by:

$$Medal_t = \beta_0 + \beta_1 GDPCAP + \beta_2 GDPCAP^2 + \beta_3 \sqrt{POP} + \beta_4 HOST + \beta_5 NEIGH + \beta_6 POL + \beta_7 FROST + \beta_8 MED + \beta_9 MedalHistory + \beta_{10} Medal_{t-1} + \varepsilon_{it}$$

where GDPCAP is the GDP per capita two years prior to the Games and POP is the population two years prior to the Games. High income can provide better infrastructure and more funding whereas larger population provides a larger pool of talent (Hoffman Ging & Ramasamy, 2002); HOST is a dummy variable that indicates the host country (Brazil = 1) (Clarke, 2000; Garicano, Palacios & Prendergast, 2001); NEIGH is a dummy variable that indicates the host country's neighbouring countries (Clarke, 2000); POL is a dummy variable that captures certain countries' efforts at the Games as a means of boosting the country brand and it is highly influenced by the political regime (Kuper & Sterken, 2001; Tcha & Pershing, 2003, Andreff, Andreff & Poupaux, 2008); FROST indicates the share of land covered by frost for at least five days per winter month. Temperature is included since climate may influence the popularity of sports as well as the country's conditions for training, (Tcha & Pershing, 2003); MED reflects the number of available medals; MedalHistory reflects the medals previously awarded to a country in one sport as a measure of tradition. The idea behind this variable is the existence of an inertial effect in some countries that creates an Olympic cult to succeed in (Andreff, 2013; Celik and Gius, 2014). In addition, tradition is not only sport culture but also the fact that countries might spend their budgets in developing their expertise by building infrastructures and attracting coaches and training methods to build and acquire skills (Williams & Ford, 2009). Investing in technological background also supports excelling in some sports, (Brownlie & Kyle, 2009; Haaker, 2009). This persistence effect is different but complementary to Medal in previous editions of the Games; and $MEDAL_{t-1}$ accounts for the success of the country in the sport at the previous edition of the Games. It corresponds to short term success in just the previous games. Apart from investment, there are two known possibilities that might account for short term success: "talent pipeline" (Bale & Sang, 1996), where the government may grant and allow their young talents to train abroad in the best centers to improve their ability and competitiveness, and the nationalization of star athletes, (Shachar, 2011). Both factors increases the possibility of winning a medal and the performance of their own athletes in competition. The dependent variable, Medal, captures the share of medals that a given country was predicted to win in a given sport. The model did not differentiate between types of medal. Theoretically, one country could obtain any number of medals up to the total number of medals (MED) that were awarded. Therefore, for each sport, the number of predicted medals had to coincide with those that were actually awarded. The number of forecast medals had to be a discrete number, so the model was further adjusted with a routine to discretise the results (Otamendi & Doncel, 2014a).

Table 1 shows the results of the prediction exercise for the 974 awarded medals after calibrating the model 10 times, once for each sport class. The forecast was therefore performed after the Games, once the total number of awarded medals per sport had been determined.[iv, v] To estimate the coefficients for each model, we used data corresponding to the last five Summer Olympic Games (Atlanta 1996, Sydney 2000, Athens 2004, Beijing 2008 and London 2012).

As expected, the economic powerhouses led the medal table. This finding supports the hypothesis of a relationship between a country's socioeconomic potential and performance. This finding also indicates that the sport-agnostic model is a reasonable proxy for long-term preparation. Econometric modelling predicted that 83 countries would obtain at least one medal, but only 27 would obtain more than 10 medals. 'TAIL< 5' corresponds to a super-country that comprises all the countries that were expected to win fewer than five medals.

4. Expected performance and competitiveness: Expert knowledge and the wisdom of crowds

Next, we assessed short-term performance and competitiveness. To do so, we addressed the following question: What type of data is made publicly available and can be used to calculate expected behaviour? Over the last two decades, online sports betting has become phenomenally popular. Fans enjoy predicting the results of sporting events, especially if winnings are paid out to those whose predictions hold true. Correctly analysing and forecasting competitive sports is the key to successful sports betting, which is influenced by past performance and current form.

Whenever a sporting event touches the hearts of the masses, people bet. When picking winners of sporting events, experts' forecasts, betting houses' odds, and ratings and rankings of competitors' ability are valuable, as are modelling processes that predict the outcome of these events. In this study, we used betting odds to study athletes' short-term performance.

We collected data from two popular online betting houses: bwin and Betfair. The odds were taken directly from these betting house official websites. In 1997, bwin began operating online under the name Betandwin until

2006. We picked this betting house as an example of expert knowledge, since in this case, the bookmakers typically set odds according to objective analysis by in-house experts and the odds are shifted only if the book becomes very strongly unbalanced.

Betfair was created in 1999. We picked this betting house as an example of betting exchange, that is, the odds are not set by the bookies, who rather provide a platform in which the bettors exchange bets. Odds can fluctuate in this case largely since the crowd, mostly fans, is the one that sets the value.

4.1. Expert knowledge: bwin

We downloaded all available information manually and created a database for data processing and analysis. For bwin, each record comprised the name of the athlete, the country code and the odds. To provide an example, Table 2 shows the odds for the Men's 800 Metres.

Insert Table 2 about here

Assigning medals to countries to build the predicted medal table according to odds was straightforward if no odds were tied. This was not always the case, however. For example, Table 2 shows that Rudisha 'wins' one medal for Kenya, and Amos another for Botswana, but Rotich and Kszczot tie for third. We assigned three medals if three were awarded during the Olympic Games by splitting the bronze medal between Kenya and Poland. In the case of an event awarding 4 medals, we proceeded similarly, assigning 4 medals to those 4 athletes with the higher odds. We broke ties by splitting the medals so that the total number of medals per event was 4. Summing the medals for all events provided the country medal count. The total sum of predicted medals based on odds therefore coincides with the total sum of awarded medals.

We were unable to collect data on every sport and event because bwin did not offer odds on some events. We nonetheless gathered odds data for 141 out of 306 events (25 out of 41 sports). Table 1 shows bwin's predictions for each sport class as of 30 July 2016. The total number of medals was 431. The number of countries with a low total was proportionally higher for the long-term prediction. Star athletes from poor countries or newcomers may be predicted to win medals according to this short-term prediction[vi].

4.2. Betfair: Wisdom of crowds

Medals were assigned to countries in the same way based on Betfair's odds. The athletes' names and the odds were downloaded from the website on 27 July 2016. The name of the country was missing. Time was spent assigning athletes to countries. Ties were shared evenly amongst athletes who had the same odds of winning a particular medal. As with bwin, it was impossible to collect data on every event and sport, but Betfair offered odds on more events than bwin did, covering 214 out of 306 events (35 out of 41 sports). Table 1 shows the results for Betfair by sport class and country. The total number of medals was 695[vii].

5. Validation: Congruence amongst expectations

We compared subjective betting predictions, based on the expert knowledge and wisdom of crowds, against forecasts, based on the objective 'sport-agnostic' method of econometric modelling. We performed econometric predictions for each sport so that we could perform robust comparisons with betting houses' odds, which were at the event level.

We performed an overall comparison, comparisons by sport and comparisons by sport class. Consequently, this paper contributes to the literature by comparing the predictive power of two sporting event prediction methods: betting odds and econometric modelling. All forecasting methods should offer similar predictions for the framework to be valid and based on the three expert sources of knowledge at once. Also, for the results to lead to robust conclusions about deviations from short- and long-term expectations.

A priori, it is unclear which method should offer the best predictions. Although the literature provides some comparisons between models and betting odds, neither consistently outperforms the other. Models perform best in some cases such as football (Constantinou et al., 2013) and UFC (Wise et al., 2010), whereas bookmakers' predictions perform best in other scenarios (Boulier & Stekler, 2003; Scheibehenne & Broeder, 2007). In some cases, there is no statistically significant difference (Easton & Uylangco, 2010). In our setting, the two prediction methods should have a similar predictive power to guarantee that long- and short-term predictions are equally important.

The success ratio provided the indicator to compare the different forecasting methods. The success ratio equated to the proportion of successful predictions or successes in the sample. The success ratio was calculated as the number of successful predictions divided by the total number of predictions. The number of successful predictions was defined for a given country as follows: The minimum of the number of awarded medals or the number of predicted medals. The number of successful predictions was thus calculated by comparing the predictions for a given country against the medals that were actually awarded to that country. If more medals were predicted than awarded, the number of successful predictions equated to the number of medals that were awarded. If fewer medals were predicted than awarded, the number of successful predictions equated to the number of medals that were predicted. The number of medals was the total medal count, regardless of type of medal (gold, silver or bronze).

The comparison method consisted of hypothesis testing based on the difference in proportions. We used this method because success ratios are proportions. The proportion π is the known parameter of the binomial distribution $\xi_j \rightarrow B(1, \pi_j)$ that represents a given set of predictions j. A test of equality of proportions can be used to compare the success ratios of different predictions. The hypotheses are as follows:

H0: $\pi_1 = \pi_2$

H1: $\pi_1 \neq \pi_2$

The z-value was used to test for equal proportions based on sample success ratios, p_j:

p_j = correct picks / medals awarded for any prediction j

$$z = \frac{p_1 - p_2}{\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}} \rightarrow N(0,1)$$

The null hypothesis was rejected (i.e. differences were considered significant) if the absolute value of the z-value was greater than 1.96 (0.05 significance). This test allowed us to establish statistical pairwise differences between the success ratios of the three sets of predictions (econometric, bwin and Betfair). Table 3 shows the results that we used for the validation exercise by country and by sport class[viii].

Insert Table 3 about here

The first column shows the country total (top row, if the medal haul is predicted for the country as a whole), the total for each of the 10 classes and the total summed across all classes (bottom row). The 'SUCCESS' column shows the number of successes in absolute terms; 'N' shows the sample size (i.e. the number of awarded medals); and 'RATIO' is the success ratio. The last three columns show the z-value for the hypothesis testing. The two z-values marked with an asterisk correspond to significant differences (0.05 level). Bwin beats ECON in terms of success ratio over all sport classes, although there were no differences at the class level. Bwin also beats Betfair for 'Elimination' sports.

The predictions were not significantly different enough to invalidate the proposed framework. Moreover, the overall success ratios were high and similar across predictions (85%–86%). The three predictions were therefore consistent enough to form the basis to identify over- or under-performance in terms of meeting expectations at Rio 2016 by country and by sport class.

6. Meeting expectations by country

To measure the extent to which a validated set of expectations for each country was met, we compared expected outcomes with observed results at Rio 2016. The expected results were compared with the awarded medals using the signed relative squared distance (sRSD). The relative squared deviation (RSD) was used because its sample statistic follows a chi-squared distribution. The sign (sRSD) was added to determine over- or under-performance in terms of meeting expectations:

$$RSD_i = \begin{cases} \dfrac{(n_i - E_i)^2}{E_i} & if E_i \geq u \\ 0 & otherwise \end{cases}$$

$$sRSD_i = \varphi_i * RSD_i$$

$$\varphi_i = \begin{cases} 1 & if (n_i - E_i) > 0 \\ 0 & if (n_i - E_i) = 0 \\ -1 & if (n_i - E_i) < 0 \end{cases}$$

where n_i are the observed sample values and E_i are the expected values for each comparison i = 1, …, I. It is recommended that each $E_i \geq u$, so a minimum sample size u, usually 5, should be used to provide robustness to the analysis (Martin-Pliego & Ruiz Maya, 2005). In our case, I was the number of countries. All countries whose expected number of medals was less than 5 were grouped into a 'super-country' called 'TAIL < 5'. This allowed us to perform an additional comparison of countries with low expected totals.

The RSD_i, and therefore $|sRSD_i|$, converges to a χ^2 distribution with 1 degree of freedom. For extraordinary performance of a given country, the absolute value of sRSD had to be larger than a value obtained from the chi-squared distribution for a given significance level α:

$$|sRSD_i| \geq \chi^2_{\alpha}(1)$$

Countries with statistically significant signed squared deviation were therefore selected because they failed to meet expectations. These countries either excelled or failed in preparation and/or performance. Meeting expectations was assessed at the overall country level and at the sport class level. Figure 2 shows the countries with at least 5 predicted medals ($E \geq 5$) with significant sRSD at the 5% level ($|sRSD| > \chi^2_{0.05}(1) = 3.841$).

Figure 2 shows the extent to which expectations are met, considering econometric predictions as an indicator of country preparation and Betfair predictions as an indicator of country performance at the Games. The first 10 graphs (from left to right, from top to bottom) show the extent to which expectations were met by sport class. The 11th graph shows how well each country met expectations. For completeness, the 12th graph shows the results for bwin at the country level. The graphs show how well countries met expectations based on Betfair odds rather than bwin odds because the sample was larger for Betfair. We nevertheless performed the analysis for each sport class based on each set of predictions. Table 4 provides the names of the countries with significant deviations, including bwin and Betfair predictions.

Insert Figure 2 about here

Insert Table 4 about here

Few countries had significant discrepancies. The reasons for these discrepancies varied and are included to foster a debate on sport policy. It would be very interesting to include all these possible development policies as factors in the econometric models, but it is unattainable to develop indicators for each and every participating country based on publicly available data. What follows is an ad-hoc analysis of the factors that should lead to sport policy development based on countries' ability to meet expectations at Rio 2016.

The host effect can be significant. Using London 2012 as a springboard, Great Britain (GBR) outperformed econometric predictions by 15 medals (67 awarded; 52 predicted), especially in Cycling and Triathlon (15; 9). If supported by good planning, the host effect can provide the foundations to develop good infrastructures (training facilities and coaches, for example) by making additional investment in preparation for the Games in an attempt to win more medals. Also, the host country participates in all events, thereby improving its chances of winning medals. However, this was not the case for Brazil. Notably, Brazil's performance reflects how Brazil failed to meet expectations according to expected economic behaviour (19; 38). Conversely, Brazil (BRA) had the worst performance of any host for several Games, with Great Britain excelling in London 2012, China in Beijing 2008, Greece in Athens 2004 and Australia in Sydney 2000.

Two emerging countries, Azerbaijan (AZE) and Uzbekistan (UZB) both ex-members of the Soviet Union, outperformed econometric predictions in Rio. Azerbaijan (AZE) (18; 9) is a wealthy country that spends heavily on sport, paying its athletes 465,000 euros per Olympic gold medal. Unsurprisingly, Azerbaijan's results are better than ever before. Uzbekistan (UZB) (13; 5) is the only country that exceeded long- and short-term expectations. Uzbekistan outperformed expectations, again, by paying its athletes 800,000 euros per Olympic gold medal. Denmark (DEN) (15; 9) also won more medals than expected in every sport class except 'Weightlifting' and 'Judge'. Its NOC has built a strong sports culture by centralising major elite sports decisions and investing in more sports. This approach has led to better results (De Bosscher, Shibli & Westerbeek, 2015). Also, in terms of expected results based on short-term competitiveness, Canada performed well. Canada won more medals than the total predicted by Bwin. Its sport policies are similar to those of Denmark (De Bosscher, Shibli & Westerbeek, 2015). On the other hand, Australian athletes performed poorly and failed to meet expectations. Its star athletes reported that they wasted their opportunities and lost concentration.

India's (IND) failure to meet expectations according to the country's socioeconomic potential was considerable. India's performance was overpredicted by econometric modelling (2; 16) because of the country's massive population. Indian athletes participated in Rio in more disciplines than in the previous four Olympic Games. If the trend continues, India may win more medals in future Games, as predicted by its large population. Other countries that performed poorly according to their socioeconomic potential were Ukraine (11; 22) and Romania (5; 12). Ukraine has suffered the effects of protracted conflict. One example of these effects is the loss of its main training centre, located in Crimea to Russia. Its NOC sent fewer athletes to the Games after suffering funding restrictions. Romania has lost is edge in Gymnastics because its best coaches now train foreign athletes abroad. Germany also performed poorly in Swimming (0; 5), failing to win a single medal for the first time since 1932. Russia was a special case in this Olympics. Because of doping sanctions, fewer Russian athletes competed than in prior editions of the Summer Olympic Games. This low turnout resulted in a high number of predicted medals but a low likelihood of meeting expectations. 'Weightlifting' (0; 5) offers a notable example. Neither betting house predicted that Russia would win many medals.

Finally, the super-country 'TAIL< 5', which comprised all countries that were predicted to win fewer than five medals, also exceeded long- and short-term expectations. This finding indicates that newcomers or emerging countries won more medals overall and in the 'Swimming' class. Serbia (8 awarded; 3 predicted) had the highest difference with respect to its socioeconomic potential. Several countries also won medals despite predictions that they would not win any (Bahrein, Burundi, Fiji, Individual Olympic Athletes, Israel, Ivory Coast, Jordan, Kosovo, Niger, Philippines, Taipei, Tajikistan, United Arab Emirates and Vietnam). Some of these countries are poor, others are new countries (e.g. Kosovo), and others (e.g. Fiji) won medals in new sports such as Rugby Sevens. The effect of 'TAIL' was also significant for Betfair, reflecting the difficulty in predicting the success of small newcomers.

7. Conclusions

Our theoretical contribution is to provide an analysis framework on how long- and short-term expectations are met at sporting events. This framework reflects the way information is converted into knowledge in the 21st century. Traditional comparative statistics from government agencies provide the input for complex models, as does Big Data provided by expert knowledge and the wisdom of crowds. We compared expected performance with actual results at the 2016 Rio Olympic Games to determine countries that significantly over- or under-performed. This insight could help to create effective sport policies. We used publicly available information for each participating country (more than 200 IOC members). We used indicators of countries' socioeconomic power and online odds for participating athletes. Uzbekistan was the only country that was awarded more medals than expected, not only according to its socioeconomic potential and previous success at the Olympic Games, but also in terms of the current form of its competitors, as reflected by betting houses' odds. The underlying assumption is that the country investment has paid off because the athletes' preparation and performance was excellent.

The joint use of indicators and expert odds has proven satisfactory to understand the Rio 2016 results. The two forecasting methods were similarly effective at forecasting the overall country medal haul. Econometric modelling had an 85.3% success ratio, whereas bwin and Betfair had success ratios of 85.1% and 86.3%, respectively. Because bettors take financial risks, their input should act as a quasi-objective predictor of sporting event outcomes. Similarly, 'sport-agnostic' mathematical models, which are used extensively to forecast econometric series, should also offer good predictions. On that regard, socioeconomic variables offer good proxies for estimating performance. Consequently, there is considerable room for improving medal performance by enhancing long-term socioeconomic indicators. At the sport level, use of the latest information (e.g. sports with a short history or short-term policies such as persuading athletes to change their sport nationality or offering government grants to train abroad in richer countries) gives betting houses a predictive advantage.

Econometric forecasting focuses on long-run success. However, three known factors might account for short-term success. First, top individual athletes might train and compete with rich nations' teams. Second, governments might give young talents the opportunity to train abroad in the best facilities. Doing so could improve their ability and competitiveness, thereby fully developing the talent pipeline. Third, countries might arrange changes of sport nationality for star athletes. Accordingly, at the sport level, odds have a predictive advantage.

This analysis may apply to other major sporting events. For example, for the world championships of any sport, socioeconomic models and odds can predict the number of medals awarded to participating countries. The framework must be adapted however for tournaments where only three medals or prizes are awarded. Prediction at the event or match level is far more complex, and modelling techniques are less accurate. However, odds are always provided, so the wisdom of crowds may be used to forecast and determine how well expectations are met.

Finally, there is room for further study. New econometric models could be developed to improve accuracy when predicting outcomes in sports that award few medals. More detailed information about sport expenditure for each country would have been helpful. Certain databases with additional expert knowledge may be added to the framework to increase the accuracy of the short-term expectations. Long-term 'sport-agnostic' models may also be improved by implementing new techniques or by incorporating new variables.

In addition, analysis of odds by type of medal could be used to shed light on countries' socioeconomic potential, detect individual talent or compare top athletes. An approach that we have introduced in this study is to perform the analysis by sport class. Sport policies for team sports might differ considerably from policies for individual sports. Our proposed taxonomy sheds light on the areas where an increase in investment may lead to high returns. For example, the host of the London 2012 Summer Olympic Games, Great Britain, invested resources to create a strong sporting infrastructure. This investment continues to pay dividends in terms of the high number of medals that Great Britain won at Rio 2016. When considered together, small countries were also successful, probably after detecting and promoting new talent. Some new countries, including Uzbekistan and Azerbaijan, have used sports to improve their visibility and raise the country's profile in the new era of the global economy.

Acknowledgement

The authors wish to thank an anonymous referee, whose comments greatly contributed to improve the quality of this work.

References

Andersson P., Memmert, D., & Popowicz E. (2009). Forecasting outcomes of the World Cup 2006 in football: Performance and confidence of bettors and laypeople. *Psychology of Sport and Exercise*, 10, 116-123.

Andreff, W. (2013). Economic development as major determinant of Olympic medal wins: predicting performances of Russian and Chinese teams at Sochi Games. *International Journal of Economic Policy in Emerging Economies*, 6, 314-340.

Andreff, M., Andreff, W. & Poupaux, S. (2008). Les determinants économiques de la performance olympique: Prévison des médailles qui seront gagnées aux Jeux de Pékin. *Revue d'Economie Politique*, 118, 135-169.

Bale, J., & Sang, J. (1996). *Kenyan running. Movement culture, geography and global change*. London: Frank Cass.

Baker, R. D., & McHale, I. G. (2013). Forecasting exact scores in National Football League games. *International Journal of Forecasting*, 29, 122-130.

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

Blais-Morisset, P., Boucher, V. & Fortin, B. (2017). The Impact of Public Expenditure in Sports on the Olympic Medals. *Revue Économique*, 68, 623-642.

Boulier, B. L., & Stekler, H. O. (2003). Predicting the outcomes of National Football League games. *International Journal of Forecasting*, 19, 257-270.

Brownlie, L.W., and Kyle, C.R. (2012). Evidence that skin suits affects long track speed skating performance.. *Procedia Engineering-ISEA 2012,* 34, 26-31.

Cantinotti, M., Ladouceur, R., & Jacques, C. (2004). Sports betting: can gamblers beat randomness? *Psychology of Addictive Behaviors*, 18, 143-147.

Celik, O. B. & Gius, M. (2014). Estimating the Determinants of Summer Olympic Game Performance. *International Journal of Applied Economics*, 11, 39-47.

Chen, W., Li, X., & Zeng, D. (2011). *Estimating Collective Belief in Fixed Odds Betting*. Pacific-Asia Workshop on Intelligence and Security Informatics, 54-63.

Clarke, S. R. (2000). Home advantage in the Olympic Games. In Cohen, G. & Langtry, T. (Eds.), *Proceedings of the fifth Australian Conference on Mathematics and Computers on Sport*, University of Technology, Sydney, Australia, 76-85.

Constantinou, A. C., Fenton, N. E., & Martin, N. (2013). Profiting from an inefficient association football gambling market: Prediction, risk and uncertainty using Bayesian networks. *Knowledge-Based Systems*, 50, 60-86.

De Bosscher, V., Shibli S., H. Westerbeek, H., and M. van Bottenburg (2015). *Successful Elite Sport Policies. An international comparison of the Sport Policy factors Leading to international Sporting Success (SPLISS 2.0) in 15 nations*. Aarchen: Meyer & Meyer.

Easton, S., & Uylangco, K. (2010). Forecasting outcomes in tennis matches using within-match betting markets. *International Journal of Forecasting*, 26, 564-575.

Forrest, D., Goddard, J., & Simmons R. (2005). Odds-setters as Forecasters: The Case of English Football. *International Journal of Forecasting*, 21, 551-564.

Forrest, D., McHale, I. G., Sanz, I. & Tena, J.D. (2017). An analysis of country medal shares in individual sports at the Olympics. *European Sport Management Quarterly*, 17, 117–131.

Forrest, D., Sanz, I. & Tena, J. D. (2010). Forecasting national team medal totals at the Summer Olympic Games. *International Journal of Forecasting*, 26, 576-588.

Forrest, D., & Simmons, R. (2008). Sentiment in the betting markets on Spanish Football. *Applied Economics*, 40, 119-126.

Garicano, L., Palacios, I. & Prendergast, C. (2001). *Favorism under Social Pressure*. Working paper 8376, NBER, Cambridge, MA.

Gramm, M., & Owens, D. H. (2005). Determinants of Betting Market Efficiency. *Applied Economic Letters*, 12, 181-185.

Grimes, A. R., Kelly, W. J., & Rubin, P. H. (1974). A Socioeconomic Model of National Olympic Performance. *Social Science Quarterly*, 55, 777-782.

Haaker, S. J. (2009). The impact of technology on sporting performance in Olympic sports. *Journal of Sport Sciences*, 27, 1421-1431.

Haan, M., Dijkstra, S. G., & Dijkstra P. T. (2005). Expert Judgment Versus Public Opinion – Evidence from the Eurovision Song Contest. *Journal of Culture Economics*, 29, 59-78.

Herzog, S. M., & Hartwig, R. (2011). The wisdom of ignorant crowds: Predicting sport outcomes by mere recognition. *Judgement and Decision Making*, 6, 58-72.

Hoffman, R., Ging, L., & Ramasamy, B. (2002). Public Policy and Olympic Success. *Applied Economic Letters*, 9, 545-548.

Hvattum, L. M., & Arntzen H. (2010). Using ELO ratings for match result prediction in association football. *International Journal of Forecasting*, 26, 460-470.

Jiang, M., & Xu, L. C. (2005). Models in transition: explaining medal performance and inequality of Chinese provinces. *Journal of Comparative Economics*, 33, 158-172.

Johnson, D. K. N., & Ali, A. (2004). A Tale of Two Seasons: Participation and Medal Counts at the Summer and Winter Olympic Games. *Social Science Quarterly*, 85, 974-993.

Khazaal, Y., Chatton, A., Billieux, J., Bizzini, L., Monney, G., Fresard, E.,…Khan, R. (2012). Effects of expertise on football betting. *Substance Abuse Treatment, Prevention and Policy*, 7, 18.

Kiernan, J., & Daley, A. (1965). *The Story of the Olympic Games, 776 B.C. to 1964*. New York: J.B. Lippincott.

Kindt, J.W., & Asmar, T. (2001). College and amateur sports gambling: Gambling away our youth? *Villanova Sports & Entertainment Law Journal*, 8, 221-252.

Klemperer, P. (2004). Auctions: *Theory and Practice*, Princeton: Princeton University Press.

Klemperer, P. (1999). Auction Theory: A Guide to the Literature. *Journal of Economic Surveys*, 13, 227-286.

Kuper, G. H., & Sterken, E. (2010). Who is going to win in Vancouver? Third International Sport Business Symposium, Vancouver, British Columbia, Canada.

Kuper, G. H., & Sterken, E. (2001). Olympic participation and performance since 1896. Available at SSRN: https://ssrn.com/abstract=274295.

Lahvicka, J. (2015). Using Monte Carlo Simulation to Calculate Match Importance: The Case of English Premier League. *Journal of Sport Economics*, 16, 390-409.

Leeds, E. M. & Leeds, M. A. (2012). Gold, Silver, and Bronze: Determining National Success in men's and women's Summer Olympic events. *Journal of Economics and Statistics*, 232, 279-292.

Leitner, C., Zeileis, A., & Hornik, K. (2010). Forecasting sports tournaments by ratings of probabilities: A comparison for the EURO 2008. *International Journal of Forecasting*, 26, 471-481.

Levitt, S. (2004). Why are gambling markets organised so differently from financial markets? *Economic Journal*, 114, 223-246.

Lisi, F., & Zanella, G. (2017). Tennis betting: can statistics beat bookmakers? *Electronic Journal of Applied Statistical Analysis*, 10, 790-808.

List, C. (2012). Collective Wisdom: Lessons from the Theory of Judgment Aggregation. In H. Landemore & J. Elster (Eds.) *Collective Wisdom: Principles and Mechanisms*. Cambridge: Cambridge University Press.

Lowen, A., Deaner, R. O. & Schmitt, E. (2016). Guys and Gals Going for Gold: The Role of Women's Empowerment in Olympic Success. *Journal of Sports Economics*, 17, 260-285

Ma, T., Tang, L., McGroarty, F., Sung, M.-C., & Johnson, J. E. V. (2016). Time is money: Costing the impact of duration misperception in market prices. *European Journal of Operational Research*, 255, 397-410.

Maher, M. (2013). Predicting the outcome of the Ryder cup. *IMA Journal of Management Mathematics*, 24, 301-309.

Martin, R.J., Nelson, S.E. & Gallucci, A.R. (2016). Game On: Past Year Gambling, Gambling-Related Problems, and Fantasy Sports Gambling Among College Athletes and Non-athletes. *Journal of Gambling Studies*, 32, 567-579.

Martin-Pliego, F.J., & Ruiz Maya, L. (2005). *Fundamentos de Inferencia Estadística*. (3rd ed.) Madrid: Thomson-Paraninfo.

McHale, I., & Morton, A. (2011). A Bradley-Terry type model for forecasting tennis match results. *International Journal of Forecasting*, 27, 619-630.

Menéndez, H. D., Vázquez, M., and Camacho, D. (2015). Mixed Clustering Methods to Forecast Baseball Trends. In D. Camacho, L. Braubach, S. Venticinque and L. Badica (Eds.) *Intelligent Distributed Computing VIII, Studies in Computational Intelligence*, 570, Springer, Cham.

Munafo, M. R., Pfeiffer, T., Altmejd, A., Heikensten, E., Almenberg, J., Bird, A., & Dreber, A. (2015). Using prediction markets to forecast research evaluations. *Royal Society Open Science* 2, nº 150287.

Noland, M., & Stahler, K. (2017). An old boys club no more: pluralism in participation and performance at the Olympic Games. *Journal of Sports Economics*, 18 (5), 506-536.

Otamendi, J., & Doncel, L. M. (2018), Can Economists Beat Sport Experts? Analysis of Medal Predictions for Sochi 2014. *Social Science Quarterly*, 99, 1699-1732.

Otamendi, J., & Doncel, L. M. (2014a), Medal Shares in Winter Olympic Games by Sport: Socioeconomic Analysis After Vancouver 2010. *Social Science Quarterly*, 95, 598–614.

Otamendi, J., & Doncel, L. M. (2014b), By Sport Predictions Through Socio Economic Factors and Tradition in Summer Olympic Games: the case of London 2012. In P. M. Pardalos & V. Zamaraev (Eds.), *Social Networks and the Economics of Sports*. New York: Springer, pp. 125-147.

Pfau, W. D. (2006). Predicting the medal wins by country at the 2006 winter Olympic games: An econometric approach. *The Korea Economic Review*, 22, 1-15.

Robles Rodriguez, J., Abad Robles, M. T., & Giménez F. J. (2009) Concepto, características, orientaciones y clasificaciones del deporte actual, *efdeportes*, 14 (138).

Sauer, R. D. (2005). The State of Research on Markets for Sports Betting and Suggested Future Directions. *Journal of Economics and Finance*, 29, 206-213.

Sauer, R. D. (1998). The Economics of Wagering Markets. *Journal of Economic Literature*, 36, 2021-2064.

Scheibehenne, B., & Broeder, A. (2007). Predicting Wimbledon 2005 tennis results by mere player name recognition. *International Journal of Forecasting*, 23, 415-426.

Schumaker, R. P. (2013). Machine learning the harness track: Crowdsourcing and varying race history. *Decision Support Systems*, 54, 1370-1379.

Seemann, T., & Hungenberg, H. (2008). Capturing Public Knowledge to Forecast Future Events. *Computer and Information Science*, 131, 87-95.

Shachar, A. (2011). Picking winners: Olympic citizenship and the global race for talent. *Yale Law Journal*, 120, 2088-2139.

Simmons, J. P., Nelson, L. D., Galak, J., & Frederick, S. (2011). Intuitive Biases in Choice versus Estimation: Implications for the Wisdom of Crowds. *Journal of Consumer Research*, 38, 1-15.

Smith, M. A., Paton, D., & Williams, L. V. (2009). Do Bookmaker possess superior skills to bettors in predicting outcomes? *Journal of Economic Behavior & Organization*, 71, 539-549.

Štrumbelj, E., & Šikonja M. R. (2010). Online bookmakers' odds as forecasts: The case of European soccer leagues. *International Journal of Forecasting*, 26, 482-488.

Sung, M. C., McDonald, D. C. J., & Johnson, J. E. V. (2016). Probabilistic forecasting with discrete choice models: Evaluating predictions with pseudo-coefficients of determination. *European Journal of Operational Research*, 248, 1021-1030.

Surowiecki, J. (2004). *The Wisdom of Crowds: Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business, Economies, Societies and Nations*. New York: Doubleday.

Tcha, M & Pershin, V. (2003). Reconsidering Performance at the Summer Olympics and Revealed Comparative Advantage. *Journal of Sports Economics*, 4, 216-239.

Trivedi, P. K., & Zimmer, D. M. (2014). Success at the Summer Olympics: How much do Economic Factors Explain? *Econometrics*, 2, 169-202.

Valero A., & Gómez-Mármol, A. (2016). Los deportes individuales. Sus características y taxonomía. EmásF, *Revista Digital de Educación Física,* 7, 42.

Vaz de Melo, P. O. S., Almeida, V. A. F., Loureiro, A. A. F., & Faloutsos, C. (2012). Forecasting in the NBA and Other Team Sports: Network effects in action. *ACM Transactions on Knowledge Discovery from Data*. Vol. 6, 3, Article 13 (October 2012), 27 pages.

Williams, A. M., & Ford, P. R. (2009). Promoting a skills-based agenda in Olympic sports: The role of skill-acquisition specialists. *Journal of Sport Sciences*, 27, 1381-1392.

Williams, L. V. (2005). *Information Efficiency in Financial and Betting Markets*. Cambridge: Cambridge University Press.

Williams, L. V. (1999). Information Efficiency in Betting Markets: a Survey. *Bulletin of Economic Research*, 55, 1-30.

Wise, S., Miric M., & Valliere, V. (2010). Testing the effectiveness of Semi-Predictive Markets: Are fight fans smarter than expert bookies? *Procedia-Social and Behavioral Sciences*, 2, 6497-6502.

Table 1. Rio 2016 - Observed and predicted medal table by sport class

RIO2016

Country	Total	Athletics	Swimming	Team	Judge	Canoe&Kayak&Rowing	Bouts	Cycling	Elimination	Weightlifting	Rounds
USA	**121**	32	33	5	16	2	9	6	10	1	7
CHN	**70**	6	6	1	17	2	10	1	11	7	9
GBR	**67**	7	6	2	12	9	7	15	2	0	7
RUS	**56**	0	4	2	13	1	18	3	9	0	6
GER	**42**	3	0	5	6	10	3	2	5	0	8
FRA	**42**	6	3	2	0	5	12	1	4	0	9
JPN	**41**	2	7	0	5	1	19	0	6	1	0
AUS	**29**	2	10	1	1	5	0	2	1	0	7
ITA	**28**	0	4	3	2	2	3	2	5	0	7
CAN	**22**	6	6	2	3	1	1	2	0	0	1
KOR	**21**	0	0	0	0	0	9	0	8	1	3
BRA	**19**	1	1	2	3	3	5	0	2	0	2
NED	**19**	1	2	1	1	3	2	6	1	0	2
NZL	**18**	4	0	1	0	6	0	1	0	0	6
AZE	**18**	0	0	0	0	2	16	0	0	0	0
ESP	**17**	2	2	2	1	4	2	1	2	1	0
KAZ	**17**	1	1	0	0	0	10	0	0	5	0
HUN	**15**	1	7	0	0	3	0	0	4	0	0
DEN	**15**	1	2	1	0	3	1	3	2	0	2
KEN	**13**	13	0	0	0	0	0	0	0	0	0
UZB	**13**	0	0	0	0	0	12	0	0	1	0
CUB	**11**	1	0	0	0	0	10	0	0	0	0
POL	**11**	3	0	0	0	4	1	2	0	0	1
SWE	**11**	0	3	1	0	0	2	2	0	0	3
UKR	**11**	1	0	0	3	2	1	0	2	0	2
JAM	**11**	11	0	0	0	0	0	0	0	0	0
RSA	**10**	4	3	1	0	1	0	1	0	0	0
CRO	**10**	3	0	1	0	2	1	0	0	0	3
CZE	**10**	1	0	0	0	4	1	1	3	0	0
BLR	**9**	1	1	0	1	1	3	0	0	2	0
COL	**8**	1	0	0	0	0	3	2	0	2	0
IRI	**8**	0	0	0	0	0	6	0	0	2	0
TUR	**8**	1	0	0	0	0	6	0	0	1	0
ETH	**8**	8	0	0	0	0	0	0	0	0	0
SRB	**8**	1	0	4	0	1	2	0	0	0	0
PRK	**7**	0	0	0	1	0	0	0	1	4	1
SUI	**7**	0	0	0	1	1	0	3	1	0	1
GEO	**7**	0	0	0	0	0	5	0	0	2	0
THA	**6**	0	0	0	0	0	2	0	0	4	0
BEL	**6**	1	1	1	0	0	1	2	0	0	0
GRE	**6**	1	1	0	1	0	0	0	0	0	3
MAS	**5**	0	0	0	1	0	0	1	3	0	0
ROU	**5**	0	0	0	0	1	1	0	2	1	0
MEX	**5**	1	0	0	1	0	2	0	0	0	1
IND	**2**	0	0	0	0	0	1	0	1	0	0
MGL	**2**	0	0	0	0	0	2	0	0	0	0
IRL	**2**	0	0	0	0	1	0	0	0	0	1
TAIL < 5	**83**	14	1	4	1	12	26	1	6	10	8
Total	**974**	**141**	**104**	**42**	**90**	**91**	**212**	**60**	**90**	**45**	**99**

ECONOMETRIC

Country	Total	Athletics	Swimming	Team	Judge	Canoe&Kayak&Rowing	Bouts	Cycling	Elimination	Weightlifting	Rounds
USA	**111**	32	32	6	9	3	11	4	8	1	5
CHN	**83**	5	8	1	20	1	11	3	18	8	8
GBR	**52**	7	3	1	3	12	9	9	2	0	6
RUS	**60**	0	4	2	12	3	23	2	6	5	3
GER	**51**	8	5	2	5	11	5	6	5	0	4
FRA	**33**	2	5	2	1	3	9	5	3	0	3
JPN	**34**	1	7	1	3	0	15	1	4	1	1
AUS	**39**	3	14	4	1	7	0	6	0	0	4
ITA	**27**	2	2	2	1	3	4	2	7	0	4
CAN	**19**	1	2	1	3	5	3	2	0	1	1
KOR	**23**	0	0	1	1	0	6	0	11	0	4
BRA	**38**	1	4	4	6	2	11	2	4	0	4
NED	**18**	0	4	2	2	1	2	3	0	0	4
NZL	**12**	1	1	0	0	5	0	2	0	0	3
AZE	**9**	0	0	0	0	0	8	0	0	1	0
ESP	**19**	1	1	3	2	3	4	2	1	0	2
KAZ	**13**	1	0	0	0	0	7	1	0	3	1
HUN	**18**	1	4	0	1	6	4	0	1	0	1
DEN	**9**	0	0	1	0	3	0	1	2	0	2
KEN	**12**	12	0	0	0	0	0	0	0	0	0
UZB	**5**	0	0	0	1	0	4	0	0	0	0
CUB	**16**	4	0	0	0	0	10	0	0	1	1
POL	**10**	2	1	0	1	2	0	0	0	2	2
SWE	**8**	0	1	1	0	0	2	1	0	0	3
UKR	**22**	4	0	0	2	4	6	0	2	2	2
JAM	**12**	12	0	0	0	0	0	0	0	0	0
RSA	**7**	2	3	0	0	2	0	0	0	0	0
CRO	**6**	1	0	2	0	1	1	0	0	0	1
CZE	**10**	2	0	0	0	4	0	1	1	0	2
BLR	**12**	2	1	0	1	3	0	0	2	2	1
COL	**7**	1	0	0	0	0	4	1	0	1	0
IRI	**13**	1	0	0	1	0	7	0	0	3	1
TUR	**8**	2	0	0	1	0	3	0	0	1	1
ETH	**8**	8	0	0	0	0	0	0	0	0	0
SRB											
PRK											
SUI											
GEO	**7**	0	0	0	0	0	7	0	0	0	0
THA											
BEL											
GRE											
MAS											
ROU	**12**	1	1	0	4	1	2	0	1	2	0
MEX	**8**	1	0	1	2	0	1	0	2	0	1
IND	**16**	0	0	1	4	0	5	0	2	1	3
MGL	**5**	0	0	0	0	0	5	0	0	0	0
IRL	**5**	0	0	0	0	0	4	0	0	0	1
TAIL < 5	**123**	20	1	5	7	6	33	6	10	11	24
Total	**974**	**141**	**104**	**42**	**90**	**91**	**212**	**60**	**90**	**45**	**99**

bwin

Country	Total	Athletics	Swimming	Team	Judge	Canoe&Kayak&Rowing	Bouts	Cycling	Elimination	Weightlifting	Rounds
USA	76.33	31.08	28.00	4.00	0.00	0.00	0.00	5.50	5.75	0.00	2.00
CHN	20.25	1.00	6.00	1.00	0.00	0.00	1.00	4.00	5.25	0.00	2.00
GBR	35.00	7.25	6.00	0.00	2.00	0.00	3.00	10.50	0.25	0.00	6.00
RUS	8.75	0.00	0.00	1.00	0.00	0.00	1.00	1.00	5.75	0.00	0.00
GER	26.83	5.83	2.00	3.00	3.00	0.00	1.00	4.00	2.00	0.00	6.00
FRA	20.50	3.00	3.00	3.00	0.00	0.00	1.00	5.50	2.00	0.00	3.00
JPN	10.00	0.00	6.00	0.00	0.00	0.00	1.00	0.00	3.00	0.00	0.00
AUS	34.83	0.00	21.50	3.00	0.00	0.00	0.00	7.00	0.00	0.00	3.33
ITA	9.50	0.00	2.00	1.00	0.00	0.00	0.00	2.00	4.50	0.00	0.00
CAN	5.50	4.00	0.50	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00
KOR	11.00	0.00	0.00	0.00	0.00	0.00	3.00	0.00	8.00	0.00	0.00
BRA	6.33	0.00	0.00	4.00	0.00	0.00	0.00	0.00	2.00	0.00	0.33
NED	16.67	3.00	2.00	2.00	1.00	0.00	0.00	5.67	2.00	0.00	1.00
NZL	9.08	1.50	1.00	0.00	0.00	0.00	0.00	4.25	0.00	0.00	2.33
AZE											
ESP	9.50	2.00	1.00	2.00	0.00	0.00	1.50	2.00	0.00	0.00	1.00
KAZ											
HUN	11.50	0.00	8.00	0.50	0.00	0.00	0.00	0.00	2.00	0.00	1.00
DEN											
KEN	13.25	13.25	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
UZB											
CUB											
POL	5.00	4.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
SWE											
UKR											
JAM	11.50	11.50	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
RSA	6.50	3.00	3.00	0.00	0.00	0.00	0.00	0.50	0.00	0.00	0.00
CRO	5.50	1.50	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	3.00
CZE											
BLR											
COL											
IRI											
TUR											
ETH	6.83	6.83	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
SRB											
PRK											
SUI											
GEO											
THA											
BEL											
GRE											
MAS											
ROU											
MEX											
IND											
MGL											
IRL											
TAIL < 5	70.83	27.25	7.00	4.50	0.00	0.00	11.50	7.08	8.50	0.00	5.00
Total	**431**	**126**	**98**	**30**	**6**	**0**	**24**	**60**	**51**	**0**	**36**

Betfair

Country	Total	Athletics	Swimming	Team	Judge	Canoe&Kayak&Rowing	Bouts	Cycling	Elimination	Weightlifting	Rounds
USA	72.70	24.00	20.87	3.00	5.50	3.00	4.33	2.50	3.00	0.00	6.50
CHN	51.20	0.00	2.20	0.00	16.00	1.00	4.00	4.00	10.00	7.00	7.00
GBR	47.83	6.00	4.67	1.00	2.00	7.67	9.00	9.00	1.00	0.00	7.50
RUS	28.03	0.00	0.20	1.00	3.50	0.00	16.50	2.33	0.50	2.00	2.00
GER	33.17	5.00	1.00	2.00	2.00	6.67	3.00	6.00	2.50	0.00	5.00
FRA	36.17	2.00	4.00	2.00	0.00	5.00	11.50	5.33	1.00	0.00	5.33
JPN	21.53	0.00	6.87	0.00	2.00	0.00	11.17	0.00	1.50	0.00	0.00
AUS	41.00	1.00	13.00	5.00	2.00	8.00	0.00	6.00	0.00	0.00	6.00
ITA	13.50	0.00	3.00	0.00	1.00	1.00	2.00	1.50	2.00	0.00	3.00
CAN	10.33	2.50	0.00	1.00	3.00	1.00	2.00	0.83	0.00	0.00	0.00
KOR	17.00	0.00	0.00	0.00	0.00	0.00	10.00	0.00	5.50	0.00	1.50
BRA	10.00	1.00	0.00	1.00	0.00	2.00	3.00	0.00	0.00	0.00	3.00
NED	21.83	3.00	0.00	2.00	2.00	2.00	4.33	5.50	2.00	0.00	1.00
NZL	18.00	1.00	0.00	2.00	0.00	10.00	0.00	3.00	0.00	0.00	2.00
AZE	12.50	0.00	0.00	0.00	0.00	0.00	12.50	0.00	0.00	0.00	0.00
ESP	9.50	0.50	0.00	2.00	0.00	0.00	0.00	2.00	2.00	0.00	3.00
KAZ	6.67	0.00	0.00	0.00	0.00	0.00	5.67	0.00	0.00	1.00	0.00
HUN	16.17	1.50	5.00	1.00	0.00	5.00	1.00	0.00	1.00	0.00	1.67
DEN	9.50	0.00	1.00	1.00	0.00	3.00	1.00	1.00	0.50	0.00	2.00
KEN	15.50	15.50	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
UZB	5.00	0.00	0.00	0.00	0.00	0.00	5.00	0.00	0.00	0.00	0.00
CUB	14.00	2.00	0.00	0.00	1.00	0.00	11.00	0.00	0.00	0.00	0.00
POL	7.70	4.50	0.20	0.00	0.00	3.00	0.00	0.00	0.00	0.00	0.00
SWE	5.00	0.00	3.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	1.00
UKR	10.00	1.00	0.00	0.00	1.00	1.00	4.00	0.00	1.00	0.00	2.00
JAM	12.00	12.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
RSA	8.00	3.00	3.00	1.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00
CRO	8.33	1.00	0.00	1.00	0.00	2.00	0.33	0.00	0.00	0.00	4.00
CZE	9.00	3.00	0.00	0.00	0.00	6.00	0.00	0.00	0.00	0.00	0.00
BLR											
COL											
IRI	8.00	0.00	0.00	0.00	0.00	0.00	7.00	0.00	0.00	1.00	0.00
TUR	6.00	1.00	1.00	0.00	0.00	0.00	4.00	0.00	0.00	0.00	0.00
ETH	8.50	8.50	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
SRB	9.33	1.00	0.00	1.00	0.00	3.67	1.67	0.00	1.00	0.00	1.00
PRK											
SUI											
GEO											
THA											
BEL											
GRE											
MAS											
ROU											
MEX											
IND											
MGL											
IRL											
TAIL < 5	92.00	14.00	2.00	3.00	4.00	12.00	30.00	8.00	1.50	7.00	10.50
Total	**695**	**114**	**71**	**30**	**45**	**85**	**164**	**57**	**36**	**18**	**75**

Table 2. Example odds for one event

NAME	CODE	ODDS	PREDICTION
DAVID RUDISHA	KEN	2.62	1
NIJEL AMOS	BOT	5	1
FERGUSON ROTICH	KEN	6	0.5
ADAM KSZCZOT	POL	6	0.5
ALFRED KIPKETER	KEN	11	0
BORIS BERIAN	USA	15	0
AMEL TUKA	BIH	15	0

KEN	1.5	2
BOT	1	1
POL	0.5	1
TOT	3	4

Table 3. Hypothesis testing for difference in proportions

	ECON			Bwin			Betfair			ECON-BWIN	ECON-BETFAIR	BWIN-BETFAIR
	SUCCESS	N	RATIO	SUCCESS	N	RATIO	SUCCESS	N	RATIO	Z-VALUE	Z-VALUE	Z-VALUE
Country (total)	831	974	85,32%	366,58	431,00	85,05%	599,63	695,00	86,28%	0,1283	-0,5552	-0,5675
Athletics	107	141	75,89%	107,00	126,00	84,92%	93,00	114,00	81,58%	-1,8780	-1,1130	0,6916
Swimming	82	104	78,85%	76,50	98,00	78,06%	54,27	71,00	76,43%	0,1356	0,3752	0,2489
Team	28	42	66,67%	21,00	30,00	70,00%	21,00	30,00	70,00%	-0,3007	-0,3007	0,0000
Judge	63	90	70,00%	5,00	6,00	83,33%	32,00	45,00	71,11%	-0,8353	-0,1338	0,7342
Canoe&Kayak&Rowing	66	91	72,53%				63,33	85,00	74,51%		-0,2980	
Bouts	156	212	73,58%	14,50	24,00	60,42%	129,50	164,00	78,96%	1,2624	-1,2244	-1,7702
Cycling & Triathlon	37	60	61,67%	39,92	60,00	66,53%	35,67	57,00	62,57%	-0,5557	-0,1010	0,4472
Elimination	67	90	74,44%	43,00	51,00	84,31%	23,00	36,00	63,89%	-1,4385	1,1434	2.1527*
Weightlifting	25	45	55,56%				8,00	18,00	44,44%		0,8018	
Rounds	64	99	64,65%	22,33	36,00	62,04%	52,33	75,00	69,78%	0,2774	-0,7171	-0,8004
Country (Sum of classes)	695	974	71,36%	329,25	431,00	76,39%	512,10	695,00	73,68%	-2.0095*	-1,0530	1,0257

Table 4. Countries that failed to meet expectations

			Total	Athletics	Swimming	Team	Judge	Canoe&Kayak &Rowing	Bouts	Cycling & Triathlon	Elimination	Weightlifting	Rounds
PREPARATION	ECON	OVER	GBR AZE DEN UZB TAIL		TAIL					GBR			
		UNDER	BRA UKR ROU IND		GER							RUS	
PERFORMANCE	BWIN	OVER	RUS CAN										
		UNDER	AUS		AUS								
	BETFAIR	OVER	UZB TAIL										
		UNDER	AUS		AUS								

[i] Results by sport are available upon request.

[ii] Three countries (India, Mongolia and Ireland) won only two, but they are highlighted since they were predicted to win at least five by at least one prediction method.

[iii] We use this model for consistency, just in case a comparison over time wants to be performed. We acknowledge that some improvements might be made on the model, including or removing certain variables, or even using other forecasting techniques, like logit or probit regression. But once again, the aim of this article was to compare the predictions with those provided by odds, and not to provide the "best" predictions. Besides, the success ratio over time of the model being used has been consistently above a high 80%, and statistically comparable to the predictions based on odds (see Section 5).

[iv] We performed a prediction exercise prior to the Games. The expected totals per sport and country were published the week prior to the Games in different media formats. The total number of predicted medals was 918, corresponding to just three per event.

[v] Coefficients and r^2 values are available upon request. All coefficients are significant in at least one sport, except for POL, with the r2 > 0.564 in each case.

[vi] Table per sport is available upon request.

[vii] Medal table by sport is available upon request.

[viii] Results by sport are available upon request.