III Encontro Pernambucano de Economia

Área temática: Teoria Aplicada

Título: The Olympic Spirit Boost: a Synthetic Control Approach

Autores:

Jorge Henrique Norões Viana
Economista e Mestre pela Universidade Federal da Paraíba. Doutorando em Economia pelo PIMES/UFPE.
Email: henriquenoroes@gmail.com
Phone: (83) 8101.8410

Antonio Vinícius Barros Barbosa
Economista pela Universidade Federal de Pernambuco. Mestre e Doutorando em Economia pelo PIMES/UFPE.
Email: aviniciusbb@hotmail.com

Breno Ramos Sampaio
Doutorado em Economia pela University of Illinois at Urbana-Champaign, EUA. Atualmente é Professor do Departamento de Economia da Universidade Federal de Pernambuco.
Email: brenorsampaio@gmail.com

Gustavo Ramos Sampaio
Doutorado em Economia pela University of Illinois at Urbana-Champaign, EUA. Atualmente é Professor do Departamento de Economia da Universidade Federal de Pernambuco.
Email: gustavorsampaio@gmail.com

The Olympic Spirit Boost:
a Synthetic Control Approach

Abstract

Hosting an Olympics can bring positive effect on sports performance of a country, mainly because the structuring of Olympic venues that encourage sports practice and the consequent emergence of high-level athletes. Using the result of the medals table provided by the International Olympic Committee (IOC), this paper analyzes through the Synthetic Control Method (SCM), the sports performance of a country after it has hosted the Olympics, for the Games between 1964 and 2004. To measure the Olympic performance for each country we constructed an index based on some desirable properties. The results show a large positive impact on the Olympic performance of the host country and this impact, although decreasing with time, remains significantly positive for several subsequent editions. Despite the high costs associated with hosting mega-events such effects indicate the benefits generated from investments in sports and the appeal of public policies for sports development and social welfare.

Keywords: *Olympic Games, Sports Performance, Synthetic Control.*

Resumo

A realização de um evento olímpico pode trazer efeitos positivos sobre a performance esportiva de um país, principalmente pela estruturação de instalações olímpicas que favoreçam a prática de esportes e pelo consequente surgimento de atletas de alto nível. Utilizando o resultado do quadro de medalhas fornecido pelo Comitê Olímpico Internacional (COI), este trabalho analisa, através do Método do Controle Sintético (SCM), o desempenho esportivo de um país após o mesmo ter sediado os Jogos Olímpicos, para as edições compreendidas no período entre 1964 e 2004. Para medir o desempenho olímpico de cada país foi construído um índice a partir de propriedades desejadas para um *ranking* de medalhas. Os resultados mostram um grande impacto positivo sobre a performance esportiva de país que sedia as Olimpíadas e este impacto, ainda que decrescente com o tempo, se matem positivo por várias edições subsequentes. Apesar dos altos custos associados à organização de megaeventos, tais efeitos apontam os benefícios gerados a partir de investimentos nos esportes e o apelo de políticas públicas para a inclusão social.

Palavras-chave: *Jogos Olímpicos, Desempenho Esportivo, Controle Sintético.*

JEL Classification Numbers: C31, D61.

Área Temática: Teoria Aplicada

1 Introduction

The Olympic Games are appointed as the main sporting event in the world, comprising a large number of modalities in different sports and athletes from different nationalities. The current format of the Summer Olympics Games was introduced in 1896 Athens Games and, although originally designed as an event that transcends the nationalist character, the Olympics have played important social and political role for host cities (Zimelis, 2011).

The literature is full of papers that analyze the economics of mega-events, which consists in the economic evaluation of the benefits for the host cities in relation to the high costs associated with implementation and maintenance of Olympic facilities. Using a variety of trade models, Rose and Spiegel (2009) estimated that hosting a mega-event like the Olympics had a permanent and large positive impact on the level of national exports. Viana and Sampaio (2013), on the other hand, estimated the impact of the FIFA World Cup on Gross Domestic Product (GDP) for the host country and showed that the effect is non-positive. Although the focus of most research papers in this literature has been to analyze economic variables, there are a number of intangible benefits, such as well-being or national pride and international visibility that may also react to hosting a large event like the World Cup or the Olympic Games. Kavetsos and Szymanski (2009), for example, show that the effect of hosting major sporting events on the level of happiness reported by individuals, the feel-good factor, was positive and statistically significant.

Given the diversity of costs and benefits generated by mega-events, there is no consensus among economists about the net effect for the host country, specially considering the multidimensionality of social and economic indicators that might be altered at national level. This difficulty in measuring the magnitude of the spillover benefits make it a hard task to justify public expenditures involved in the feasibility of such events (Coates and Humphreys, 2003).

One of the spillover effects that has received little attention in the literature is the effect of hosting the Olympics on athletes performance on subsequent events. As the International Olympic Committee (IOC) reserves the right for the host country to have representation at all played sports, two potential effects may arise through investment in the event. First, the government may invest in the more traditional sports in that particular country, given they present greater medal chances. Second, the implementation of facilities for Olympic sports with lower adhesion at first can generate community benefits able to attract the interest of new athletes over time. The presence of new sports facilities may generate a series of non-economic benefits such as increased accessibility and participation, generating a positive effect on sports success (Grieve and Sherry, 2012). These two factors, in theory, should leverage overall sports performance of a country that decides in favor of hosting the Olympics.

In the 2012 London Summer Olympic Games, for example, medals were contested in 302 events divided in 32 sports (columns 3 and 4 of Table 1 show total events and sports disputed, respectively, in the year that the Olympics occurred, T_0). Great Britain had representation in all sports, as dictated by IOCs rules, a much different scenario from the one observed in the 2008 Beijing Olympics Games, in which Great Britain had representation in only 23 sports (an increase of about 40%), as shown in columns 7 and 8 (Participating in T_{0-1}) of Table 1. This is also observed on other editions of the event.

[Table 1 about here]

Although important, current literature has little to say about the causal link between hosting an Olympic event on country's Olympic performance in subsequent events. Most of the analysis carried out so far compare only the performance of a country on events before and after hosting the event, which may lead to incorrect policy conclusions. That may happen because unaccounted factors, such as GDP, may affect not only the decision to host the event (which, in turn, affects performance), but also performance directly. Therefore, the difference between medal rankings before and after hosting the event would reflect not only the effect of hosting the event but also the effect of better economic conditions (higher GDP) on performance. In this paper, we tackle this question by using a precise methodology to deal with the problem of endogeneity. More specifically, we use the Synthetic Control Method, developed by Abadie and Gardeazebal (2003) and extended by Abadie et al. (2010). This method uses data-driven procedures to construct adequate comparison groups/counterfactuals given that, in practice, it is a difficult task to find a single country unexposed to the policy change (hosting the event) that approximates the most, relevant characteristics of the treated country (country hosting the event) and that would provide a decent control group.

Our results shows that there is in fact a positive and persistent effect of hosting an Olympics for the host country performance index. Nevertheless, this impact on Olympic performance was clearly greater for the beginning of the post-hosting period and then declined significantly over next editions of the Olympics. For most countries, however, the effect remains positive even after five Olympic Games. The only exceptions were the experiences of Mexico, which hosted the Olympics in 1968, and Greece, which hosted the 2004 edition. For these two countries (Mexico and Greece), the boost on Olympic performance lasted for one edition, specifically the editions they hosted and after that their performance either went back to the pre-hosting level, in the case of Mexico, or even got worse the historical level as was the case for Greece. One possible explanation for the cases of Mexico and specially Greece is that in those countries investment in sports is strongly correlated with the economic performance which, in turn for both, took a turn for the worst right after they hosted the Olympics and, as consequence, brought down the Olympic performance.

A more striking example is the 1988 Summer Olympics held in South Korean. As pointed out by Lim

(2010), although the Olympics were not, at least proclaimed, related to political and social development, the success in hosting the games created an atmosphere of social and economic prosperity, especially for the consolidation of the democracy. The 1988 Seoul Summer Games is deemed by the historians as one of the most successful Olympic Games and considered to have greatly exceeded expectations, especially regarding athletic achievements (Zimelis, 2011). As most reports argue, the environment and access to sports, generated by hosting the event, were responsible for the substantial increase in performance of South Korea on following years.

After this brief introduction, in the next section we present the synthetic control method and describe how to construct the medal weight vector used as the performance index for each country in each event. Section 3 presents the data used in the analysis and Section 4 describes the main results and placebo tests used as robustness check. Finally, section 5 concludes.

2 Methodology

In this section we present the empirical strategy used to identify the causal effect of hosting a Summer Olympic Games on the future sports performance. Let Y_{ct} be the outcome for country c at time t, O_{ct} be a dummy variable that assumes value equal to 1 for the years following the occurrence of an Olympics, and ϵ_{ct} be unobserved determinants of the outcome variable. The parameter of interest, β_1, which represents the effect of the Olympics on the outcome, may be estimated via the following model

$$Y_{ct} = \beta_0 + \beta_1 * O_{ct} + \epsilon_{ct} \tag{1}$$

One can easily verify that by estimating equation 1 using data only for the country that hosted an Olympics, the parameter of interest would equal the average of the outcome variable after the Olympics (when $O_{ct} = 1$) minus the average of the outcome variable before the Olympics (when $O_{ct} = 0$). It is hard to argue, however, that such difference represent the causal effect of the Olympics, given that there are other confounding factors not controlled for that might compromise identification, that is, it might be that $COV(O_{ct}, \epsilon_{ct}) \neq 0$.

To overcome the problems described above, the usual practice in this literature has been to use data on another country (or many other countries) that did not host any Olympics during the years before or after the country currently hosting the Olympics. These countries would then be used as counterfactuals for the country being analyzed and the parameter of interest would be identified via a difference-in-differences (DID) setup. This strategy would remove bias that might result from permanent differences between the country hosting the Olympics and other countries used as counterfactuals, as well as bias from comparison over time in the country that had the Olympics that could be the result of time trends unrelated to the Olympics itself (Imbens and Wooldridge, 2009). In this case, the equation to be estimated is given by

$$Y_{ct} = \alpha_0 + \alpha_1 * O_{ct} + \Theta X_{ct} + \lambda_c + \lambda_t + \mu_{ct} \tag{2}$$

where X_{ct} is a vector of controls, and λ_c and λ_t are, respectively, country and year fixed effects to control non-parametrically for country time-invariant unobservable characteristics and for yearly differences between the outcome of interest. The parameter of interest, α_1, equals the average gain over time in the countries not hosting an Olympics minus the average gain over time in the country hosting the Olympics. One main hypothesis required for the validity of the this approach in identifying the Olympics effect, is that both treated and control countries must have exactly the same time trend in the absence of the Olympics, and it is not clear why this should be the case. If, for example, the countries not hosting an Olympics have different trends compared to the country hosting the Olympics, the researcher will be unable to differentiate between the Olympics effect and the trend difference.

This shortcoming is exactly what we aim to overcome in the present paper by using the synthetic control method to construct a combination of countries that best describes pre-treatment variables for the country hosting the Olympics, i.e., this artificially constructed group is much similar to the treated country in the pre-treatment periods than any of the control country on their own.

2.1 The Synthetic Control Method (SCM)

In this section we describe the synthetic control method (SCM) developed by Abadie and Gardeazabal (2003) and extended in Abadie *et al.* (2010). We also discuss its advantages and limitations when compared to other methodologies used in the literature, paying particular attention to DID strategies. Suppose there are $J + 1$ countries and that only the first country is exposed to the policy change (the country hosting the Olympics), so that there are J remaining regions as potential controls (all other countries not hosting Olympics in the period near the one being analyzed). Let Y^N_{ct} be the outcome that would be observed for region c at time t in the absence of the intervention, for units $c = 1, ..., J + 1$, and time periods $t = 1, ..., T$. Let Y^I_{ct} be the outcome that would be observed for unit c at time t if unit c is exposed to the intervention in periods $T_0 + 1$ to T, where T_0 is the number of pre-intervention periods such that $1 \leq T_0 < T$. It is assumed that the intervention has no effect on the outcome of interest before the implementation period, such that for $t \in 1, ..., T_0$ and all $c \in 1, ..., N$ we have that $Y^I_{ct} = Y^N_{ct}$.

Now let $\alpha_{ct} = Y^I_{ct} - Y^N_{ct}$ the effect of the intervention for unit c at time t, and let D_{ct} be an indicator that takes value one if unit c is exposed to the intervention at time t, and zero otherwise. In this case, the observed outcome for unit c at time t is given by $Y_{ct} = Y_{ct}^N - \alpha_{ct} D_{ct}$. For region one, which is the only region exposed to the intervention after period T_0, it follows that $D_{ct} = 1$ for $t > T_0$ and zero otherwise.

Our objective is to estimate $(\alpha_{1T_0+1}, ..., \alpha_{1T_1})$, which is given by $\alpha_{1t} = Y_{1t}^I - Y_{1t}^N = Y_{1t} - Y_{1t}^N$. The problem in estimating $\alpha' s$ in this case is that Y_{ct}^N is never observed for the treated region once $t > T_0$. Thus, one must estimate its value. To see how a control group might be obtained from the set of untreated regions, suppose as in Abadie *et al.* (2010) that Y_{ct}^N is given by the following model

$$Y_{ct}^N = \delta_t + \theta_t Z_c + \lambda_t \mu_c + \epsilon_{ct} \tag{3}$$

where δ_t is an unknown common factor with constant factor loadings across units, Z_c is a vector of observed covariates (not affected by the intervention), θ_t is a vector of unknown parameters, λ_t is a vector of unobserved common factors, μ_c is an vector of unknown factor loadings, and the error terms ϵ_{ct} are unobserved transitory shocks at the region level with zero mean.

Now consider a $(J \times 1)$ vector of weights $W = (w_2, ..., w_{J+1})'$ such that $w_j \geq 0$ for $j = 2, ..., J+1$ and $w_2 + ... + w_{J+1} = 1$. Each value that W might take represents a synthetic control group for region one. For example, if $w_2 = 1$ and $w_j = 0$ for $j = 3, ..., J+1$, then region 2 works as control for region one (the treated one). If, on the other hand, one sets a subset $J' \subset J$ to have equal weights, such that $w_{j'} = 1/J'$ for $j \subset J$ and 0 otherwise, the comparison would be between the treated region and the average of all other regions that belong to the group J'.

Using W as weights to construct a weighted average of equation 3, one obtains the following expression

$$\sum_{j=2}^{J+1} w_j Y_{jt} = \delta_t + \theta_t \sum_{j=2}^{J+1} w_j Z_j + \lambda_t \sum_{j=2}^{J+1} w_j \mu_c + \sum_{j=2}^{J+1} w_j \epsilon_{jt}. \tag{4}$$

If one assumes that exists weights $(w_2^*, ..., w_{J+1}^*)$ such that the following holds, $\sum_{j=2}^{J+1} w_j Y_{j1} = Y_{11}, ...,$ $\sum_{j=2}^{J+1} w_j^* Y_{jT_0} = Y_{1T_0}$ and $\sum_{j=2}^{J+1} w_j^* Z_j = Z_1$ then Abadie *et al.* (2010) prove that the following equation is true

$$Y_{1t}^N - \sum_{j=2}^{J+1} w_j^* Y_{jt} = \sum_{j=2}^{J+1} w_j \sum_{s=1}^{T_0} \lambda_t \left(\sum_{n=1}^{T_0} \lambda_n^{'} \lambda_n \right)^{-1} \lambda_s^{'} (\epsilon_{js} - \epsilon_{1s}) - \sum_{j=2}^{J+1} w_j^* (\epsilon_{jt} - \epsilon_{1t}) \tag{5}$$

and that its right hand side will be close to zero if the number of pre-intervention periods is large relative to the scale of the transitory shocks. This implies that $Y_{1t}^N = \sum_{j=2}^{J+1} w_j^* Y_{jt}$ which suggests the following estimator for the α vector:

$$\hat{\alpha_{1t}} = Y_{1t} - \sum_{j=2}^{J+1} w_j^* Y_{jt} \quad (6)$$

To obtain the vector of optimal weights W, let $X_1 = (Z_1^{'}, Y_{11}, ..., Y_{1_{T0}})^{'}$ be a vector of pre-intervention characteristics for the treated region an X_0 be a matrix that contains the same variables for the untreated regions, such that jth column of X_0 is $(Z_j^{'}, Y_{ij}, ..., Y_{jT_0})^{'}$. Then W^* is chosen to minimize the distance, $\|X_1 - X_0 W\|_V = \sqrt{(X_1 - X_0 W)' V (X_1 - X_0 W)}$ between X_1 and $X_0 W$ subject to $w_j \geq 0$ and $w_2 + ... + w_{J+1} = 1$, where V is symmetric and positive semidefinite matrix chosen in a way that the resulting synthetic control region approximates the trajectory of the outcome variable of the affected region in the pre-intervention periods.

The model described above has several advantages when compared to other approaches used in the literature. As pointed out by Nannicini and Ricciuti (2010), the model is transparent, given the weights $(w_2^*, \ldots, w_{J+1}^*)$ identify the regions that are used to construct counterfactuals for the treated region, and the model is flexible, as the set of potential control regions can be appropriately restricted to make the comparisons sensible. Also, the model relaxes the assumption that confounding factors are time invariant (fixed effects) or share a common trend (differences-in-differences), given the effect of unobservable confounding factors is allowed to vary with time.

On the other hand, this approach has the limitation that it does not allow one to assess the significance of the results using standard inferential techniques, given the number of untreated regions and the number of periods considered are small. Abadie *et al.* (2010) suggest that inference should be carried out by implementing placebo experiments. In this case, inference is based on comparisons between the magnitude of the gaps generated by the placebo studies and the magnitude of the gap generated for the treated state. Thus, if the gap estimated for the treated state is large compared to the gap estimated for the placebo experiments, then the analysis would suggest that the treatment had an effect on the outcome of interest and is not driven by chance.

2.2 System of Points for Medal Ranking

The International Olympic Committee does not recognize an official performance ranking by country. Nonetheless, this has not prevented each country to present its own ranking, generally based on the number of gold or total medals won. For this study the definition of a measurement used to assess the performance of each country is a key point, because as seen in the previous section, the results presented by the synthetic control method depend on the choice of a variable of interest (Y_{ct}), which given our goal needed to be a performance index. In this sense, the purely lexicographical criteria used to rank countries based in their performance

(Barba-Romero and Pomerol, 1997) seems simplistic. For example, following this criteria, if country A wins only a gold medal and country B wins several silver medals, country A would still be better ranked when compared to country B. Many studies recognize the problem reported above when applying lexicographical ranking to the Olympic performance and most suggest alternative ways of qualifying countries. Most of these studies use Data Envelopment Analysis (DEA), which is a non-parametric method of constructing efficiency frontiers from information about variables that serve as "input" for the "production" of a good Olympic performance (Lins *et al.*, 2003; Soares de Mello *et al.* 2008 and 2009).

This paper sought an alternative method proposed by Sitarz (2013) which establishes minimum constraints on a set of medal scores. The first constraint requires that the gold medal is awarded highest score regarding the silver medal; and that the silver medal is given higher scores when compared to the bronze medal. The second constraint is that the difference between the score of the gold medal and silver medal is greater than the difference between the score of the silver medal and bronze medal (Hai, 2007). Thus the set of the weighting for medals is given by:

$$K = \left\{(x_1, x_2, x_3) \in \Re^3 : x_1 \geq x_2 \geq x_3 \geq 0 \ and \ x_1 - x_2 \geq x_3 - x_3\right\} \quad (7)$$

where x_1 denotes the score for the gold medal, x_2 the score for the silver medal and x_3 the score for the bronze medal . Sitarz (2013) draws attention to the fact that the set K is an unbounded convex cone and the corresponding mean for these types of set is its incenter. Using numerical methods to calculate the incenter of this set, Sitarz (2013) comes to the following scoring system for the Olympics:

$$(\bar{x_1}, \bar{x_2}, \bar{x_3}) = ((\sqrt{2}+1)(\sqrt{3}+2) - (\sqrt{3}+1), (\sqrt{2}+1), 1) \approx (6.3, 2.4, 1) \quad (8)$$

which was used in this paper to calculate the Olympic performance index for each country to be used as the variable of interest (Y_{ct}) in the synthetic control method.

3 Data

To compose the sport performance index used as outcome of interest, we use individual-level data obtained from the IOC Database about the number of medals won by athletes from a specific country between the Games from 1964 to 2004. An overall medal table is obtained adding up the gold, silver and bronze medals for each modality by country. Given the interest of identifying the impact of hosting the Olympic Games on the future outcome of the treated country, the Independent Olympic Participants and the Combined Nation Teams were not here considered.

The synthetic control methodology consists of "creating" a country that is similar to the characteristics of the treated country. The choice of predictors is of fundamental importance for the construction of the synthetic control; it should reflect economic and social aspects that may influence the sport performance of a country during the period of analysis. The variables used as predictors are described below. The number of participants in the competing countries in all Olympic Games was obtained from the IOC Database. Information on the total population were obtained from the Penn World Tables 8.0 and data related to life expectancy and the GDP per capita were taken from the World Bank Database. Also, we use data about secondary school enrollment as Barro and Lee (2000) as covariate in the vector of pre-intervention characteristics. Abadie *et al.* (2010) argues that the inclusion of lagged performance index should be used to improve the fit of the pre-treatment period.

The next step in the construction of synthetic control is the choice of the time horizon for adjusting the pre-treatment period and for analyzing the performance in subsequent Olympic games. For the pre-treatment period we use data from three previous games. To estimate the treatment effect, we compared the host country with its synthetic for a maximum period of five Olympic Games, given the availability of data. Thus, the treatment period T_0 is set to immediately prior to the holding of the Olympic Games period for the treated country in order to capture the home advantage effects (Pollard, 1986).

The increase of the number of Olympic medals won by a nation may be related to the best sports performance in a particular sport or increasing the number of events in each sport. In order to capture only the first effect, the performance improvement in a particular sport, the data analyzed eliminates the second effect, homogenizing the number of sports played throughout the time horizon of the analysis.

Regarding the choice of countries potentially used as control for the treated country, we use those who participated in all games during the analysis period, but who did not receive treatment during the period. That is, for each analyzed event were excluded from the pool of controls all the countries that hosted the Olympic Games during the period. Still, were excluded from the pool of controls countries that did not have enough information to construct the synthetic and those who did not have at least one observation in the period of adjustment for each of the predictors. Thus, the pool of controls is formed by countries that participated concomitantly with the treated country during the period of analysis.

An important limitation is the availability of data for the period, especially for the 1968 Mexico and 1964 Tokyo Games. For these two Olympic Games, we did not use for predicting the life expectancy of the population and data on GDP per capita. The 1972 Munich and the 1980 Moscow Games were not analyzed here by reason of restructuring their political systems (German reunification in 1990 and the collapse of the countries that were part of the Soviet block). Some works try to precisely predict the economic performance of these countries after the intervention using the synthetic control approach (Abadie, Diamond and Hainmueller, 2010; Kennedy, 2012).

The 1980 Moscow and 1984 Los Angeles Games were carried out in a troubled world political period. The political position of the countries in the midst of the Cold War generated a series of boycotts. The result of these choices could be observed in Moscow, 1980, when only 80 countries attended the event. The 1984 Los Angeles and 1996 Atlanta Games were not analyzed because the U.S. have always the highest rank, so we would not gain any information studying the U.S. case and even if we did studied, we would not be able to construct a synthetic control for it. Still, the Olympics of 2008 Beijing and 2012 London Games were not analyzed for the reason we do not have additional information for comparison.

This paper analyzes the 1964 Tokyo, the 1968 Mexico, the 1976 Montreal, the 1988 Seoul, the 1992 Barcelona, the 2000 Sydney and the 2004 Athens Summer Olympic Games.

4 Results

Before looking at the estimated effects of the Summer Olympic Games, let us first look at the countries that compose the synthetic country for each of the treated countries and how their pre-treatment characteristics compare to the pre-treatment characteristics of the real hosting country. Table 2 presents the estimated weights for each country in the set of potential control countries. Synthetic Greece and synthetic Australia, for example, are convex combinations of 17 and 16 countries respectively, while synthetic Japan is composed of only 10 countries. Hence, as pointed out above, the model is transparent, given the weights clearly identify the countries that are used to construct the counterfactuals.

[Table 2 about here.]

In Table 3 we provide a comparison by explanatory variable between each treated country and the constructed synthetic control. A general conclusion is that the synthetic countries seem to provide a better control group than only comparing the treated country with the average characteristics of all other countries in the donor pool or with a single country, given the optimization process used. This is exactly the justification for the use of the synthetic control method approach in the first place.

[Table 3 about here.]

The graphs on the left-hand side of Figures 1-7 represent the time series of the Olympic index performance, the outcome variable, for the treated country (solid line) and the synthetic control country (dashed line), both in the entire pre-treatment period, the previous three Olympics, and for the post-period, the five Olympics after the hosting. The dotted vertical red line represents the year of the Olympics just before the one hosted in the country under analysis and the black one represents the year that the Olympics took place. For example, in Figure 1 the dotted vertical red line highlights the year 2000, which was the year of the last Olympic before Athens 2004 and the dotted vertical black line highlights the year 2004, when the Olympics Games really occurred. The comparison between the solid and dashed line before treatment shows the quality of adjustment in the time series of the outcome variable for the country hosting the Olympics and the time series of the outcome variable for the synthetic country. The after-treatment period comparison estimates the dynamic treatment effects of interest.

This can also be seen on the graph on the right-hand side of Figures 1-7, where we plot the gap between the outcome variable of the country hosting the event (solid line of the graph on the left-hand side) and the outcome variable of the synthetic control (dashed line of the graph on the left-hand side). Again, we plot two dashed vertical lines representing the year of the Olympics just before the one hosted in the country under analysis and the year that it occurred.

As pointed out by Abadie *et al.* (2010) and by Abadie and Gardeazabal (2003), one must "evaluate the significance" of the estimates using the SCM, given "results could be driven entirely by chance". Thus, they propose that the SCM should be applied to all other countries that did not hosted the Olympics before in the period analyzed (donor pool) and inference is based on comparisons between the magnitude of the gaps generated by the placebo studies and the magnitude of the gap generated for the real treated country. Therefore, we implement this idea and add placebo gaps, which are represented by grey lines, to the graphs on the right-hand side. Note that these placebo gaps consider only our baseline specification, i.e., that treatment is defined as the year of the last Olympics. We should emphasize also that we discarded placebo countries with pre-intervention mean squared prediction error - MSPE (the average of the squared discrepancies between performance index in the treated country and in its synthetic counterpart during the pre-intervention period) five times larger than the hosting country. This is because placebo countries with poor fit prior to the Olympics do not provide information to measure the relative "rarity" of estimating a large post-event gap for a country which is well fitted prior to the intervention (Abadie *et al.*, 2010).

In general, pre-treatment adjustment between real and synthetic countries performance index was quite good, although a few countries presented poor pre-treatment adjustment compromising the inferential value of the case study. We now describe in more detail the results and provide some contextual background to justify potential heterogeneities.

The results for Greece 2004 are presented graphically in Figure 1. The pre-treatment adjustment between real and synthetic Greece was quite good for the period after 1992, with performance index time series almost overlapping that of synthetic performance index time series. The estimated dynamic treatment effects were quite different along time. In the short term there is a clearly "home-advantage effect", with a massive boost on the Olympic performance of Greece in 2004. However, in the next edition (2008 Beijing Games) that boost transformed in an even greater decline, which only got bigger in the next edition (2012 London Games). Note that these conclusions endure the placebo tests in the right panel of Figure 1, where the gap of Greece Olympic performance is either a upper bound (for 2004) or a lower bound (for 2008 and 2012) of all other gaps calculated for the pool control.

A possible explanation for this poor Olympic performance boost for Greece after 2004 is that the costs involved in hosting the games in 2004 were one of the many causes of the Greek debt crisis by the end of 2009. According to the Greek Finance Ministry of 8.9 billion euros, expend in hosting the Olympics, approximately 80% were financed by the government local and national. And even if the financial cost of the Olympics was not one of the causes of the Greek debt crisis, in a country living a crisis such that the finance of new and old athletes is definitely injured, which in turn have serious consequences in the performance of that country in such a competitive event as the Summer Olympics.

[Figures 1 about here]

[Figures 2 about here]

[Figures 3 about here]

In Figure 2 we present results for the 2000 Sydney Olympic Games. Differently from what occurred in Greece, Australia experienced a substantial increase in the Olympic performance that clearly last for, at least, two editions of the games. Is true, however, that the performance boost was decreasing over time but, even in 2012, the performance was significantly better than before hosting, as pointed out in the two panels of Figure 2, specially the panel B.

A similar case to Australia was the experience of Spain after host the 1992 Barcelona Games. After that period, the performance massively improved, partly because of a very poor performance in the editions before 1988. However, after 1992 the same performance index entered in a clear downslope path, that ceased in 2000 and started again in 2008. As was the case for Australia, the Spain performance kept itself better than the years before hosting for the whole period, as can be seen in the placebo test in panel B of Figure 3. We do, however, take this result for Spain more cautiously, because the synthetic control found for Spain presented a poorer pre-treatment fit than the one presented for Australia, but was still rich enough so that we could not reject the results presented just above. What also call attention in the results for Spain and Australia is the timing of the second negative turn in the Olympic performance, which took place after the 2008 and suggests that the 2008 global financial crises injured this countries not only economic, but as a byproduct of the economic downturn, it also injured their Olympic performance.

The South Korea case, among our results, is the one that shows a greater success of hosting an Olympics over the performance index of the hosting country. As can be seen on the Figure 4, this is the case because since 1988 the performance index of Korea has been kept significantly greater than the same index for synthetic Korea and other countries placebo study. The Korean case can then be flagged as on hope of undisputed success for the countries who want to host an edition of the Summer Olympics in the future. However, as our other results shows, the Korean experience is not easy to reproduce. This is so because one possible explanation for the success of Korea in maintain part of its hosting boost was the Korean formidable economic development since the 1970. In turn this economic development success was strongly associated with an even more formidable advance in education, one of the pillars of a sort performance that is both successful and stable (Zimelis, 2011).

Canada hosted the 1976 Montreal Olympic Games, and its effect on the Olympic performance on the country is shown on Figure 5. The history told by Figure 5 is one of great variability, one reason behind this is that Canada did not attend in the 1980 Moscow Olympics, so there is no observation for Canada in that year. This is one cause for the change from Canada‘s performance index from 1976 to its performance in 1984 was so drastic. Another reason for this variability of Canada performance was the boycott from the communist countries to the 1984 Los Angeles Olympics, which helped the countries which participate of that Olympics to attain a much better performance, by means of less competitors. So even the Figure 5 show a great boost in Canada performance between 1976 and 1984 this boost was not estimated accurately. However, from 1984 and upward Canada indeed maintained a performance significantly better then its synthetic counterpart.

The 1968 Mexico Games was the first edition held in a less-developed country since the 1896 Athens Games. Mexico has experienced an average of 6% increase in per capita between 1940 and 1980, an episode referred as the “Mexican Miracle”. The development and modernization of the country - mainly in the urban areas -, however, was not followed by advances in the education system and health distribution. The results for the Mexican sport performance is shown in figure 6. The index trend for Mexico after hosting the Games is not clear, given the increase of the synthetic index performance. For example, in the 1980 Moscow and 1984 Los Angeles Games, the performance of Mexico alternate in sign compared with its synthetic counterpart. The placebo test in Panel B confirms the undefined effect of hosting the Olympics.

[Figures 4-7 about here]

Finally, figure 7 shows the results for the 1964 Tokyo Games. As we can see, the trend of the index performance is increasing since the 1952 Helsinki Games with peak in the Tokyo Olympics, reinforcing the hypothesis of home advantages. However, the effect disappears after 1984 (Japan did not participate in the 1980 Moscow Games). Panel B of Figure 7 shows that the effect is positive only for the following three editions of the Games.

5 Concluding Remarks

There is no consensus among economists in relation to the benefits generated by mega-events, like the Olympic Games, due the high costs involved in the viability of the infrastructure of the sporting events and the burden generated on the population, which calls for improvements in health, education and income distribution. A number of studies have examined the benefits generated on economic activity, on the creation of new jobs and on international trade (Baade and Mathenson, 2002; Rose and Spiegel, 2009; Gathani, Santini and Stoelinga 2013).

This paper shows an intangible effect generated by the development of new non-traditional sports and the consolidation of others. The net effect here is aimed at improving sports performance of athletes in different types of sports.

Our results show that the effect is positive in periods subsequent the treatment period (when the country hosted the Olympic Games).[1] A positive and persistent effect in hosting the Olympic Games took place in South Korea. Before hosting the Games in 1988, the country configured in a modest position in the overall medals table with poor athletic performance. Currently, has consolidated as a strong participant in various forms, in both individual and team sports. The exceptions are Greece and Mexico (in 2004 and 1968, respectively). Our results show that for these two countries, investment in sports is strongly related to economic performance.

Thus, the assessment of the effects of mega-events must take into account the potential long-term effects. When considered the effects on sports performance, one must evaluate the importance of public policies of investment in sports and its effects on the welfare of the population and social inclusion.

References

[1] Abadie, A., Diamond, A., Hainmueller, J., 2010. Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program. Journal of the American Statistical Association, 104, 493-505.

[2] Abadie, A., Gardeazabal, J., 2003. The Economic Costs of Conflict: A Case Studyof the Basque Country. American Economic Review, 93, 113-132.

[1] Although the Olympic Games are held in a single host city, we believe in the existence of splillovers for athletes around the country. Thus, we assign the national sport performance in each event.

[3] Angrist, J., and J.S. Pischke, 2008. Mostly Harmless Econometrics: An Empiricist's Companion. Princeton University Press

[4] Baade, R., Matheson, V., 2002. Bidding for the Olympics: Fool's Gold? in C. P. Barros, M. Ibrahimo and S. Szymanski (eds.) Transatlantic Sport: The Comparative Economics of North American and European Sports, Cheltenham: Edward Elgar.

[5] Barba-Romero, S. and Pomerol, J. C. (1997). Decisiones multicriterio: Fundamentos teóricos e utilización práctica: Universidad de Alcalá.

[6] Barro, R. J., Lee, J-W., 2010. A new data set of educational attainment in the world, 1950-2010. NBER Working Paper No. 15902, Cambridge, Massachusetts.

[7] Coates, D., Humphreys, B., 2008. Do Economists Reach a Conclusion on Subsidies for Sports Franchises, Stadiums, and Mega-Events?, Vol. 5(3), 294-315.

[8] Coates, D., Humphreys, B., 2003. Professional Sports Facilities, Franchises and Urban Economic Development. UMBC Economics Department Working Paper, 03-103.

[9] Fremeth, A. R., Holburn, G. L., Richter, B. K, 2013. Making Causal Inferences in Small Samples using Synthetic Control Methodology: Did Chrysler Benefit from Government Assistance? Working papers series.

[10] Gathani, S., Santini, M., Stoelinga, D., 2013. Innovative techniques to evaluate the Impact of private sector development reforms: An application to Rwanda and 11 other countries. MPSA Annual Conference.

[11] Glynn, M., 2008. Configuring the Field of Play: How Hosting the Olympic Games Impacts Civic Community. Journal of Management Studies, Vol. 45, Issue 6, pp. 1117-1146.

[12] Hai H. L., 2007. Using vote-ranking and cross-evaluation methods to assess the performance of nations at the Olympics, WSEAS Transactions on Systems 6 (6), 1196â1205.

[13] Henrion, R., Seeger, A., 2010. On properties of different notions of centers for convex cones, Set-Valued and Variational Analysis 182, pp. 205-231.

[14] Kavetsos, G; Szymanski, S., 2013. National Well-being and International Sports Events, North American Association of Sports Economists. Working Paper Series, Paper No. 08-04.

[15] Kennedy, R., 2012. Fading Colours? A Synthetic Comparative Case Study of the Impact of 'Colour Revolutions'. APSA 2012 Annual Meeting Paper.

[16] Lins, M.P.E.; Gomes, E; J.C.C.B. Soares de Mello. Olympic ranking based on a zero sum gains DEA model. European Journal of Operational Research 148 (2003) 312-322.

[17] Liou, Yu-Ming, Musgrave, P., 2013. Refining the Oil Curse: Country-Level Evidence From Exogenous Variations in Resource Income. Comparative Political Studies, 1-27.

[18] Matheson, V. Mega-Events: The effect of the worldâs biggest sporting events on local, regional, and national economies. Holy Cross Working Paper Series (2006).

[19] Nannicini, T., Ricciuti, R., 2010. Autocratic Transitions to Growth. CESifo Working Paper No. 2967.

[20] Pollard R., 1986. Home advantage in soccer: a retrospective analysis. J Sports Sci 1986; 4: 237-48.

[21] Rose, A. K., Spiegel, M. M., 2011. The Olympic Effect. The Economic Journal, 121, 652-677.

[22] Sitarz, S., 2013. The medal points' incenter for rankings in sport. Applied Mathematics Letters 26, pp. 408-412.

[23] Soares de Mello, J.C.C.B.; Gomes, E.G.; Angulo-Meza, L.; L. Biondi Neto, L. Cross evaluation using weight restrictions in unitary input DEA models: theoretical aspects and application to Olympic Games ranking, WSEAS Transactions on Systems 7 (1) (2008) 31-39.

[24] Viana, J; Sampaio, B. Does the World Cup Get the Economics Ball Rolling? Evidence from a Synthetic Control Approach, PIMES, UFPE.

[25] Zimelis, A., 2011. Let the Games Begin. Politics of Olympic Games in Mexico and South Korea. India Quarterly, 67(3), 263-278.

Tables

Table 1: Number of sporting events matched and attended by the host country in current Olympic year and previous editions.

		Total in T_0		Participating in T_0		Participating in T_{0-1}		Participating in T_{0-2}	
Year	Host Country	Events	Sports	Events	Sports	Events	Sports	Events	Sports
2012	Great Britain	302	32	249	32	185	23	162	24
2008	China	303	34	257	34	204	31	163	28
2004	Greece	301	34	215	33	114	23	94	18
2000	Australia	300	34	370	34	211	26	153	25
1996	USA	271	31	263	31	248	28	230	27
1992	Spain	257	29	195	29	130	24	104	23
1988	South Korea	237	27	218	27	97	19	-	-
1984	USA	221	26	217	25	-	-	189	19
1980	Soviet Union	203	23	202	23	189	22	180	22
1976	Canada	198	23	173	23	136	18	124	14
1972	West Germany	193	23	183	23	154	17	-	-
1968	Mexico	172	20	146	20	58	15	54	14
1964	Japan	163	21	155	21	96	17	71	13

Note: Data available from the International Olympic Committee (IOC). The table shows the total events and sports disputed in all Olympics since 1964 Japan Games. The IOCs rules reserves the right to the Host Country have representation in all sports. We show the sport attendance in the current hosting period and in two games before the treatment period.

Table 2: Country Weights for Synthetic Controls

Pool of Controls	**Synthetic Greece weights**	**Synthetic Australia weights**	**Synthetic Spain weights**	**Synthetic South Korea weights**	**Synthetic Canada weights**	**Synthetic Mexico weights**	**Synthetic Japan weights**
ARG	0.001			0.001	0.595		
AUT	0.001				0.215	0.088	
BRA				0.557		0.703	
CAN	0.007	0.377					
CRC	0.534			0.001			
CYP			0.749				
FRA	0.003	0.623				0.046	
GBR				0.074	0.190		0.946
HKG	0.168						
IND			0.023	0.046			0.054
ISL	0.002		0.064				
ITA	0.254		0.161				
JAM				0.077			
MEX				0.002			
NZL	0.002						
PER						0.041	
POR	0.001			0.001		0.119	
ROU				0.224			
SWE	0.002						

Note: This table presents the estimated weights for each host country that compose the synthetic control. We show the weights equal or greater than 0.001 for a clear presentation of the counterfactual construction.

Table 3: Olympic Performance Index Predictor Means

	Greece	**Synthetic Greece**	**Sample Mean**
Participants	111	105.345	87.272
Population	10.698	18.769	43.211
Life expectancy	77.652	77.596	70.253
GDP per capita	11657.51	11725.33	9977.875
Sec. enrollment	2515951	6205427	6552835
Index 1992.2000	27.9	27.884	19.222
	Australia	**Synthetic Australia**	**Sample Mean**
Participants	333.333	320.753	72.437
Population	17.482	47.228	23.965
Life expectancy	77.296	77.203	69.763
GDP per capita	18287.05	21389.7	9047.28
Sec. enrollment	7201665	13878076	3813606
Index 1988-1996	71.433	71.424	13.203
	Spain	**Synthetic Spain**	**Sample Mean**
participants	193.667	49.677	58.895
Population	38.193	27.532	36.563
Life expectancy	76.131	75.106	66.698
GDP per capita	6611.872	6789.521	5750.971
Sec. enrollment	4243043	3380433	3309720
Index 1980.1988	12.3	12.346	11.533
	South korea	**Synthetic South Korea**	**Sample Mean**
Participants	123	121.811	60.274
Population	37.61	110.123	36.541
Life expectancy	65.446	65.373	65.819
GDP per capita	1565.642	1739.473	3709.454
Sec. enrollment	6418956	6352260	3105909
Index 1976.1984	35.95	37.94	11.81
	Canada	**Synthetic Canada**	**Sample Mean**
Participants	156.333	120.019	56.193
Population	21.045	25.94	27.159
Life expectancy	72.355	68.166	61.416
GDP per capita	3663.02	1565.947	1194.429
Sec. enrollment	4791551	3113120	1266137
Index 1964.1972	11.467	11.569	8.77
	Mexico	**Synthetic Mexico**	**Sample Mean**
Participants	63.667	62.366	62.833
Population	38.1	55.323	39.873
Life expectancy	59.056	59.058	64.727
GDP per capita	456.852	458.213	1257.575
Sec. enrollment	705022.1	1410375	1367321
Index 1956.1964	3.1	3.839	10.991
	Japan	**Synthetic Japan**	**Sample Mean**
Participants	116.333	233.186	68.76
Population	89.889	70.727	32.605
Sec enrollment	11444972	7094588	1069663
Index 1952.1960	41.967	39.311	11.789

Note: The table shows the mean values of the covariates and outcome variables used in the analysis. The outcome variable is the sport performance index and the covariates include the participants athletes, the total population, life expectancy, real GDP per capita and secondary school enrollment. All values are obtained using the *R* package “Synth”, developed by Abadie, Diamond and Hainmueller(2011).

Figures

Figure 1: Trends in Greece index performance (Panel A) and Gap in Greece index performance (Panel B).

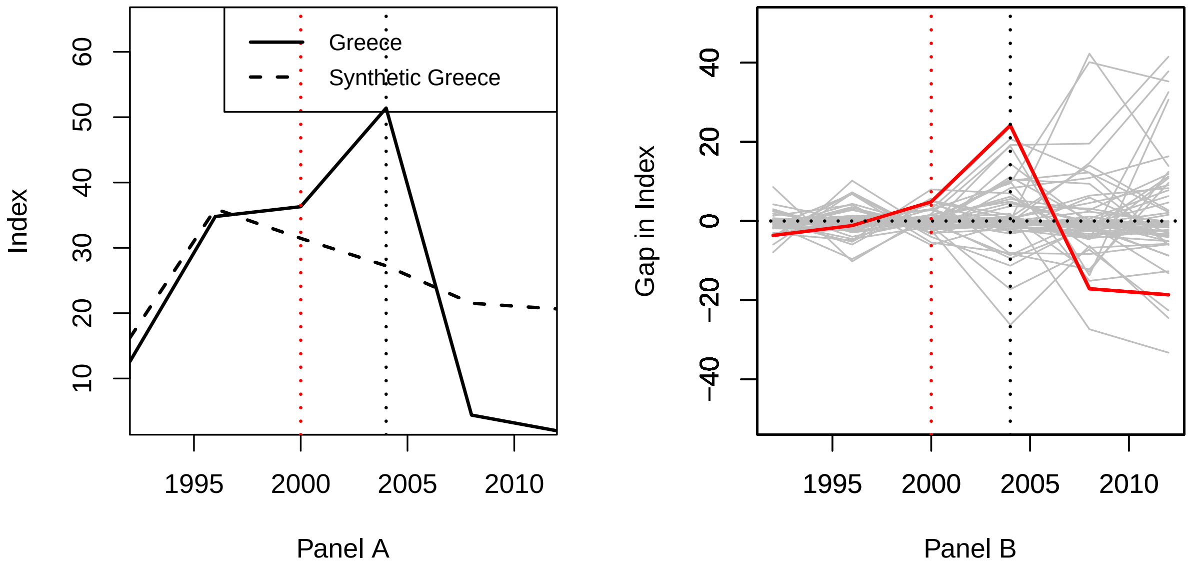

Figure 2: Trends in Australia Index Performance (Panel A) and Gap in Australia Index Performance (Panel B).

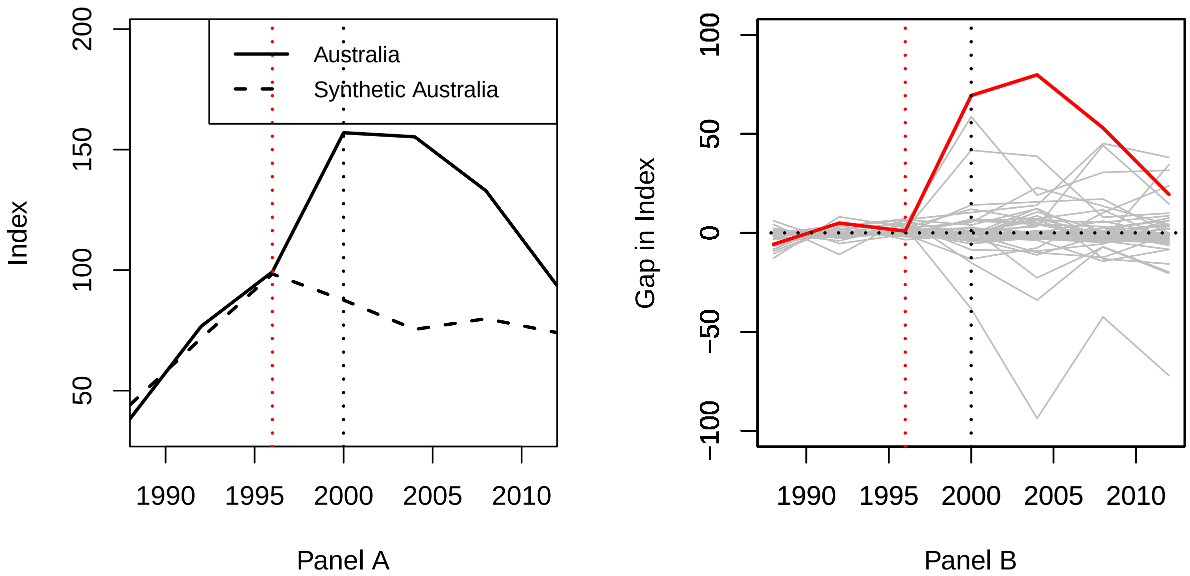

Figure 3: Trends in Spain Index Performance (Panel A) and Gap in Spain Index Performance (Panel B).

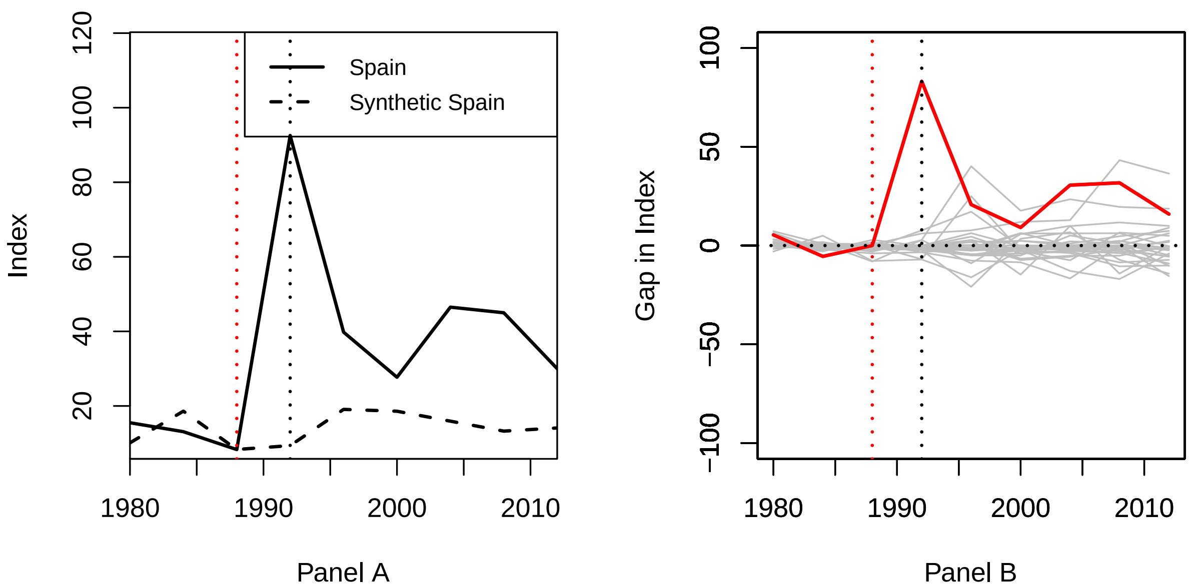

Figure 4: Trends in South Korea Index Performance (Panel A) and Gap in South Korea Index Performance (Panel B).

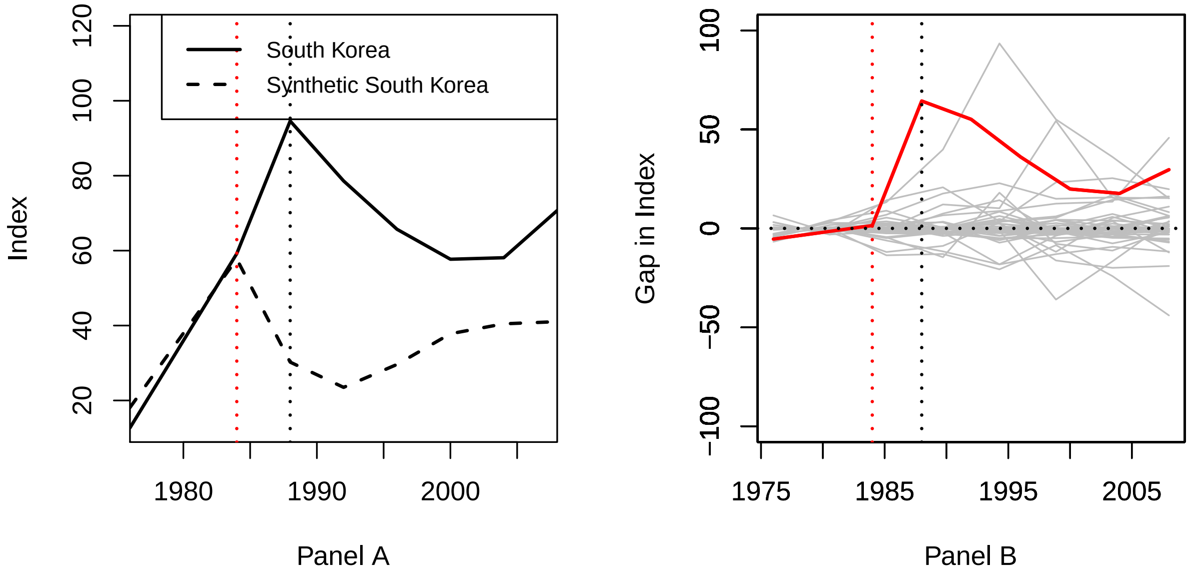

Figure 5: Trends in Canada Index Performance (Panel A) and Gap in Canada Index Performance (Panel B).

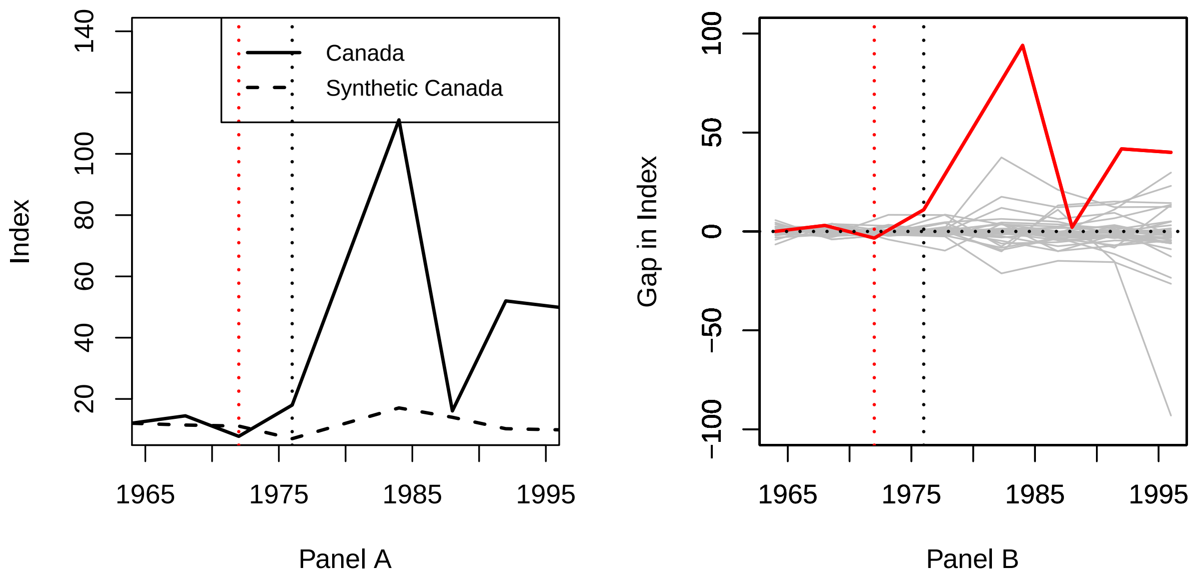

Figure 6: Trends in Mexico Index Performance (Panel A) and Gap in Mexico Index Performance (Panel B).

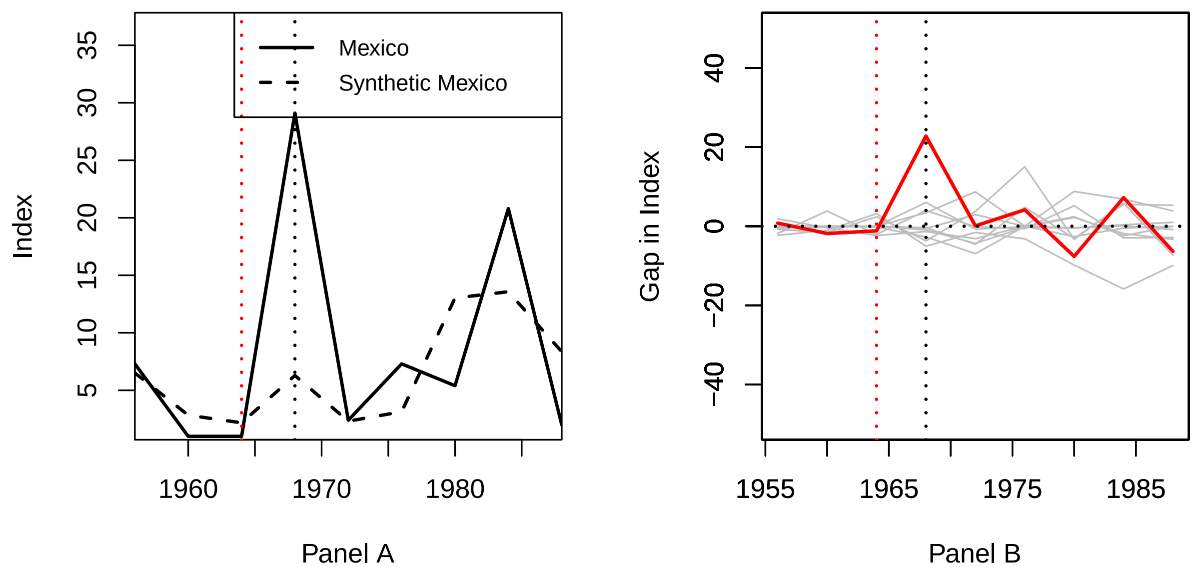

Figure 7: Trends in Japan Index Performance (Panel A) and Gap in Japan Index Performance (Panel B).