BBR
BRAZILIAN BUSINESS REVIEW

ARTICLE

Effects of New Stadiums on Attendance of Brazilian Football Clubs

Marke Geisy da Silva Dantas[1]
markegeisy@hotmail.com | 0000-0002-8991-7792

Luciano Menezes Bezerra Sampaio[2]
luciano.sampaio@ufrn.br | 0000-0003-1632-3149

ABSTRACT

This paper aims to analyze the impact of stadium construction and reformation on Brazilian soccer clubs' average attendence. Two methodologies were used: synthetic control and Differences-in-Differences (DD) models. The clubs were divided into treated, nine with new stadiums, and control, with fourteen teams that continued in old stadiums. The synthetic control model indicated that only the attendence average for the team Palmeiras showed any significant impact after the new arena inauguration, which could explain the positive effect on the treated group in the DD model. However, there is also statistical significance when Palmeiras is removed from this group. We concluded that there was a significant impact on attendence averages for Brazilian clubs that built private stadiums, and no impact on clubs that play in public stadiums, despite the government investment made.

KEYWORDS

Attendance Average, 2014 World Cup Stadiums, Synthetic Control, Differences-in-Differences, Brazilian Clubs.

[1]Instituto Federal do Rio Grande do Norte, Canguaretama, RN, Brazil

[2]Universidade Federal do Rio Grande do Norte, Natal, RN, Brazil

Received: 06/16 /2021.
Revised: 01/04/2022.
Accepted: 02/15/2023.
Published: 10/31/2023.
DOI: https://doi.org/10.15728/bbr.2022.1287.en

Efeitos de Novos Estádios no Público dos Clubes Brasileiros de Futebol

RESUMO

O objetivo deste estudo foi analisar o impacto da construção e reforma de estádios na média de público dos clubes brasileiros de futebol. Foram utilizadas duas metodologias: controle sintético e modelos de Diferenças em Diferenças (DD). Os clubes foram divididos em tratados — nove com estádios novos —, e controle, com quatorze equipes que continuaram jogando em estádios antigos. O resultado do modelo de controle sintético indicou que apenas na média de público do Palmeiras se observou algum impacto significante após a inauguração de uma nova arena, o que poderia explicar o efeito positivo sobre o todo o grupo tratado no modelo de DD. Entretanto, também existe significância estatística quando se retira o Palmeiras desse grupo. Concluiu-se de forma geral que houve um impacto significante na média de público dos clubes brasileiros que construíram estádios particulares, não havendo impacto nos clubes que jogam em estádios públicos, apesar do investimento governamental realizado.

PALAVRAS-CHAVE

Média de Público, Estádios da Copa do Mundo, Controle Sintético, Diferenças em Diferenças, Clubes Brasileiros.

1. INTRODUCTION

Twelve arenas were built or reformed for the 2014 Football World Cup held in Brazil, involving public, consortium, or club resources, establishing a new standard in Brazilian football for attending sports events. The expectation was that, after the World Cup, the attendance average of matches would increase (Castilho et al., 2017), which is one of the reasons for holding matches in a greater number of cities than usual in previous World Cups.

Despite an increase in attendance for the Brazilian championship of 27% between 2003 and 2018, its average in 2018 was still the thirteenth (just over 17,000 spectators per game) among 51 leagues around the world (Poli et al., 2019). The Brazilian attendance average is lower than that of less important leagues, such as the Chinese Super League (22,594 people per game), Major League Soccer (United States, 21,358 people per game), and second divisions of Germany and England (Poli et al., 2019).

The academic literature has addressed the impact on attendance due to building and/or moving to new stadiums (Coates & Humphreys, 2005; Gitter & Rhoads, 2014; Leadley & Zygmont, 2006); as well as the effect of a new stadium on attendance and revenue of matches, by the "novelty factor" (honeymoon) (Leadley & Zygmont, 2005; Leadley & Zygmont, 2006; Zygmont & Leadley, 2005), indicating a stabilization or a drop in attendance average after this initial phase, when other motivational and/or restrictive factors become more influential.

Of the twelve stadiums built or reformed for the 2014 World Cup, three are privately owned: Neo Química Arena, Beira-Rio stadium, and Arena da Baixada; while the others are public stadiums managed by state governments (Arena Castelão, Mineirão, Arena Pantanal, Arena da Amazônia and Arena Pernambuco), concessions to companies (Arena Fonte Nova and Arena das Dunas) or clubs (Maracanã). Still, in the period between 2012 and 2014, three other private

arenas were also (re)inaugurated, but not used in World Cup matches: Arena Independência, in 2012; the Arena do Grêmio, inaugurated in 2013; and Allianz Parque, in 2014.

The possibly significant impact of new stadiums has not yet been measured, despite an increase in the attendance average of the Brazilian championship after the World Cup, which was one of the justifications for the construction and reformation of stadiums with public funds. Furthermore, research reporting the effect of new stadiums mostly used linear regression models (Clapp & Hakes, 2005; Coates & Humphreys, 2005; Gitter & Rhoads, 2014; Kahane & Shmanske, 1997; Soebbing et al., 2016), or limited dependent variable models such as Tobit (Leadley & Zygmont, 2005; Leadley & Zygmont, 2006; Zygmont & Leadley, 2005). Generally, the primary dependent variable is total or average season tickets. To achieve the objective, in this paper, synthetic control models (Abadie & Gardeazabal, 2003; Abadie et al., 2010; Abadie et al., 2015) and Differences in Differences, including the two-way fixed effects method, by Goodman-Bacon (2018) were used.

This work aims to contribute to the impact assessment literature, by analyzing care in new arenas based on the Synthetic Control (CS) method, as an alternative to models that have been used. Thus, we were able to evaluate this impact individually for each treated club, which has a new stadium or used a (re)built public property, for use in games. The comparison of SC and DD results allowed identifying the effects individually and jointly for treated clubs.

The results of this research demonstrated that, individually, only Palmeiras' attendance average was impacted by stadium construction, considering the synthetic control method. Furthermore, the differences-in-differences method showed that the treated group in 2013 (clubs that used public stadiums renovated for World Cup) had a significant impact on its attendance average, with a 10% significance level in the estimated model, while, in the group of clubs with private stadiums (dealt with in 2014), built or renovated, the impact was significant in all estimated models.

2. LITERATURE REVIEW ABOUT STADIUM CONSTRUCTION AND ITS IMPACT ON ATTENDANCE

The new arenas' impacts on sports clubs' attendance at sports management papers is quite diverse in objectives, methods, and results. The main theme is the analysis of the honeymoon effect, a period in which there is an increase in attendance after the stadium inauguration.

McEvoy et al. (2005) stated that Roger Noll's paper, "Attendance and Price Setting", from 1974, is a seminal study about the honeymoon effect in attendance of sports leagues. The aforementioned authors found a positive effect on the attendance of the clubs in the first years of the existence of new facilities. Because of this, later studies analyzed this honeymoon in several leagues, mainly American sports leagues.

In Kahane and Shmanske's (1997) paper, the authors considered a variable referring to new stadiums in the estimated models for attendance in Major League Baseball, in the 1990-1991 and 1991-1992 seasons. The variable New Stadium ("New Stad") is a dummy variable equal to "1" if the home team's stadium was built in the five years before observation. This variable was identified as positive and statistically significant.

McEvoy et al. (2005) aimed to understand the honeymoon effect on the attendance average of new facilities in Major League Baseball (MLB), relating the facility age and seasonal attendance in the MLB from 1962 to 2001. The results indicated that there was a relationship between facility age and lifetime attendance, negatively significant, and squared stadium age, positively significant.

For the same league, but incorporating the assessment of the impact on club revenue, in addition to attendance, Poitras and Hadley (2006) estimated models from a panel of annual team observations for the period from 1989 to 2001. The dummy for new stadiums and a dummy for the stadium's first year showed a positive relationship with attendance, as well as revenue, in estimated models.

Rascher et al. (2012) evaluated the variation (standard deviation) of attendance and revenues in new arenas of the National Hockey League (NHL), with 15-year data (from 1989 to 2003), and dummies for the first and fifth years as interest variables in regression models. The results showed that attendance standard deviation decreased in new arenas and that both average revenue and its standard deviation are higher in new arenas.

An alternative way of measuring new stadium effects is to include dummies in the model for years following the stadiums' construction, as done by Coates and Humphreys (2005). They investigated the effect of new facilities on attendance for three sports: baseball, basketball, and American football, from 1969 to 2001. The dependent variable is the attendance average for three American leagues (Major League Baseball, the National Football League, and the National Basketball Association). The stadium age showed a negative coefficient in three models and the dummies per year (up to the tenth year) detected a persistent honeymoon until the ninth year for NBA, until the eighth for MLB, and until the fifth year for the NFL.

Using a database covering the period from 1992 to 2006, with more than 200 teams, Gitter and Rhoads (2014) studied minor baseball leagues and the impact of new stadiums on match attendance. Their results showed that there was an increase in attendance over the period in the three leagues analyzed (AAA, AA, and A), and this variation begins to regress more slowly than in Major League Baseball. Despite finding a significant increase up to the tenth year of stadium construction, the authors conclude that the increase in the number of tickets did not compensate for stadium construction costs.

The aforementioned research generally estimated linear regression models for the analysis of their results. In two papers, Leadley and Zygmont (2005, 2006) used Tobit regression to test the honeymoons effect on attendance. In the first of them, for the National Basketball Association (NBA), in the period from 1971 to 2000, they showed that there was a honeymoon effect, with an increase in public attendance of 15% to 20% in the first 4 years of new arena operation, a consequent reduction of impact from fifth year, and without effect for the ninth year. In the second paper (Leadley & Zygmont, 2006), regarding the NHL and considering the period from 1970 to 2003, and a subperiod from 1994 to 2003, the authors showed that the new stadium's effect on attendance lasted for five and eight years in models with longer and shorter periods, respectively.

Because of the research presented, it is clear that there are a variety of results, with a short or long-term impact on the presence of new stadiums, in leagues of different sports, especially in United States leagues. None of the papers presented results with soccer data, even with the World Cups being held every four years and the consequent reformulation of stadiums in different countries. Another issue is that the models estimated in the literature presented did not consider the specific characteristics of each club in different samples.

3. METHODOLOGY

The synthetic control method was developed by Abadie and Gardeazabal (2003) to measure the impact of terrorist conflicts on the GDP (Gross Domestic Product) per capita of the Basque Country region of Spain. The methodology seeks a comparison, estimating a synthetic unit from other units (control group), and that is comparable with the unit that received the treatment. The method was improved based on papers by Abadie et al. (2010) and Abadie et al. (2015). This estimation of a synthetic unit is based on a weighted average of the control units, through weights of outcome variables and relevant predictors in pre-intervention periods (Ando, 2015).

For the analysis, a data set of Brazilian football clubs was assembled, for the period from 2008 to 2019. The initial choice of clubs was made considering those that played in Serie A or Serie B of the Brazilian championship in all years post-World Cup, which were not relegated to Serie C, which resulted in 28 clubs. However, due to a lack of data on some variables, five of them (América-MG, Ceará, Chapecoense, Criciúma, and Oeste) were removed from the data. Therefore, the final sample of twenty-three clubs is composed of Atlético-GO, Atlético-MG, Athletico-PR, Avaí, Bahia, Botafogo, Corinthians, Coritiba, Cruzeiro, Figueirense, Flamengo, Fluminense, Goiás, Grêmio, Internacional, Palmeiras, Paraná, Ponte Preta, Santos, São Paulo, Sport Recife, Vasco, and Vitória.

Of these twenty-three, nine clubs are considered treated units because they play in new or reformed stadiums: five built or renovated their arenas during the analyzed period and another four use public stadiums rebuilt for the 2014 World Cup. The remaining fourteen clubs formed a control group, the donor pool for calculations of the counterfactuals. The year of treatment was considered the one in which the club played its first game in the new arena during the Brazilian series A or B championships, as shown in Table 1.

Table 1
Built or reformed stadiums used in this paper and their respective home clubs

Stadium	Home Club	Date of (re) Opening	First Year	2014 World Cup Stadium?
Arena do Grêmio	Grêmio	12/08/2012	2013	No
Mineirão	Cruzeiro	02/03/2013	2013	Yes
Arena Fonte Nova	Bahia	04/07/2013	2013	Yes
Maracanã	Flamengo	04/27/2013	2013	Yes
Maracanã	Fluminense	04/27/2013	2013	Yes
Estádio Beira-Rio	Internacional	02/15/2014	2014	Yes
Arena da Baixada	Athletico-PR	05/14/2014	2014	Yes
Neo Química Arena	Corinthians	05/18/2014	2014	Yes
Allianz Parque	Palmeiras	11/19/2014	2014	No

Note. Source: developed by authors

The attendance average of series A and B of the Brazilian Championships is the outcome variable, and it was collected from three sources. First, between 2013 and 2019, matches financial bulletins, and attendance average was built by adding the number of used tickets for all matches played at home in a season and dividing by the number of matches played at home (19). As these bulletins were not available for the years 2008 and 2012, the attendance average for these games was collected directly on the "Transfermarkt" website, and when not available on this website, on the "Bola n@ Área" website.

Four variables were used as predictors: Points Percentage, Ranking CBF, Players' Value, and Total Assets. The points percentage was calculated from the final tables of Series A and B championships, dividing the number of points by the total number of possible points (38 rounds multiplied by 3 points, resulting in 114 possible points). The CBF Ranking is made available by the Brazilian Football Confederation (CBF) annually, a score based on the club's position in championships promoted by the entity. Club performances as a way to explain attendance stadiums were also considered by Coates and Humphreys (2007) and Villa et al. (2011). The player's value was collected on the Transfermarkt website in €(euros), a variable proposed in Serrano et al. (2015); and Total Assets were included to show the clubs size and were collected from clubs` financial statements.

The synthetic control method allows inferences to be made from placebo tests (Abadie et al., 2010): it is verified through the emulation of the treatment in units that were not initially treated if there are different effects for the treated ones. Furthermore, it is possible to calculate p-values for estimated impact (Cavallo et al., 2013).

As an alternative method, Differences-in-Differences were estimated to detect aggregate effects on attendance in clubs that started to play in new or refurbished stadiums compared to the group that continued to play in old stadiums. As in the treated group with treatment in 2013 and 2014, different regression models were also used for these two subgroups.

For each of these treated subgroups, three regression models were estimated. The first with the dependent variable Attendance Average and with three dummies as predictor variables: one to identify the treated club ("1" for treated club j, "0" for others); another for the treatment period ("1" for the years after treatment $t > T_{0,}$ and "0" for others); and a third for the interaction between two previous ones (assuming "1" for club j in period $t > T_0$). In the second regression model, the dependent variable was the logarithm of Attendance Average; and in the third regression, the dependent variable was also the Attendance Average in logarithm, but with data only from 2011, and with the addition of four controls in logarithm: Points Percentage, CBF Ranking, Players Value, and Total Assets.

Considering the difference between the two treatment groups (with treatment in 2013 and 2014) and between the treated and untreated groups, two two-way fixed effects DD models by Goodman-Bacon (2018) were used. Thus, the resulting model can be presented, in general, by the following equation:

$$Attendance\ Average_{jt} = \alpha_j + \alpha_j + \beta^{DD} D_{jt} + u_{jt}$$

Where α_j and α_j are fixed effects components for units and time, respectively, while D_{jt} is a dummy where the value "1" refers to treatment period t for treated club j. In this case, β^{DD} represents the mean estimator of Diff-in-Diff for the following comparisons among different groups: Early Group x Late Group; Early Group x Untreated Group; and Late Group x Untreated Group. As two periods of treatment are presented, there are two groups called: the Early Group,

composed of teams that received treatment (new stadium) in 2013; and the Late Group, for clubs that received treatment in 2014. Both groups are compared with specific control groups, according to a year of treatment and inclusion and exclusion of clubs that are treated in a year, but not in another.

We used to obtain "Synth" package from R (Abadie et al., 2011) to obtain results of the Synthetic Control Method, while we used SCTools" package (Silva, 2020), also from R, for the calculation of the placebo and p-value tests. For the difference-in-differences method and two-way fixed effects model by Goodman-Bacon (2018), we used Stata 13, through the "bacondecomp" package.

4. RESULTS

Table 2 presents the variables' average and in which period these variables were analyzed. The attendance Average was higher in the years 2018 and 2019. Points Percentage average ranged between 47% and 50%. Players' Value average evolved from € 32 million in 2009 to just over € 53 million in 2019. The CBF Ranking average in 2011 is lower than in other years because the calculation methodology was different in that year. Finally, average Total Assets increased from R$ 208 million to approximately R$ 403 million from 2010 to 2019.

Table 2
Outcome and predictor variables average

Year	Attendance Average	Points Percentage (%)	Players Value (€)	CBF Ranking	Total Assets (R$)
2008	15,560	50.30	-	-	-
2009	17,104	48.32	32,199,130	-	-
2010	13,714	48.32	40,662,609	-	207,867,651
2011	13,964	47.56	37,901,304	2,553	297,748,728
2012	12,915	48.51	39,015,435	11,579	350,463,770
2013	14,753	48.55	40,863,478	11,293	327,466,325
2014	15,042	48.70	31,407,826	10,900	340,448,054
2015	16,290	48.05	52,223,043	10,693	362,650,936
2016	14,731	49.31	48,913,043	10,510	379,692,271
2017	16,099	46.19	45,066,522	10,544	376,413,963
2018	16,800	47.29	49,087,826	10,359	374,501,861
2019	18,842	48.47	53,433,913	10,223	402,511,867

Note. Source: developed by authors

Table 3 presents the estimated weights that each team in the control group represents in combination for calculating the counterfactual of the treated clubs.

Table 3
Donor pool weights for building synthetic control of treated clubs

Clubs	APR	BAH	COR	CRU	FLA	FLU	GRE	INT	PAL
Atlético-GO	0.001	0.429	-	0.026	-	0.001	0.013	0.012	0.021
Atlético-MG	0.182	-	-	0.087	0.591	0.001	0.014	0.346	0.160
Avaí	0.001	0.001	-	0.028	-	0.001	0.014	0.012	0.022
Botafogo	0.116	0.001	-	0.043	-	0.001	0.014	0.011	0.039
Coritiba	0.001	0.278	-	0.031	-	0.001	0.013	0.009	0.030
Figueirense	0.001	-	-	0.027	-	0.001	0.013	0.013	0.020
Goiás	0.001	0.001	-	0.027	-	0.001	0.013	0.008	0.019
Paraná	0.352	0.004	-	0.030	-	0.001	0.013	0.010	0.027
Ponte Preta	0.001	0.001	-	0.028	-	0.001	0.013	0.013	0.028
Santos	-	0.001	-	0.041	0.001	0.048	0.039	0.186	0.133
São Paulo	0.029	-	1	0.468	0.331	0.338	0.807	0.351	0.404
Sport Recife	0.001	0.282	-	0.031	-	0.001	0.014	0.014	0.031
Vasco	0.307	0.001	-	0.105	0.076	0.607	0.005	0.007	0.045
Vitória	0.005	0.001	-	0.028	-	0.001	0.013	0.008	0.021

Note. Source: by authors
APR = Athletico-PR; BAH = Bahia; COR = Corinthians; CRU = Cruzeiro; FLA = Flamengo; FLU = Fluminense; GRE = Grêmio; INT = Internacional; PAL = Palmeiras

Table 4 presents averages of each predictor variable, comparing them for treated clubs and their synthetics, as well as the established weight for each of the variables to form respective synthetic. The calculation of synthetic versions of Athletico-PR and Cruzeiro assigned greater weight to the CBF Ranking variable; for Bahia, Fluminense, and Grêmio, to Total Assets; the points percentage had the greatest importance for Corinthians and Internacional; while for Flamengo and Palmeiras, the highest weights of their synthetic versions were attributed to the value of their players.

In graphs presented from Figures 1-9, in alphabetical order of clubs, the black line represents the observed value of the attendance average and the black dashed line in the horizontal direction is the synthetic value of the same variable, while the vertical dashed line reflects the year of treatment. Of the treated clubs, Corinthians and Grêmio are the only teams that did not need to change stadiums in the pre-treatment period. Corinthians did not have its arena, with no impact in this period regarding the temporary move to a new arena. In general, we observe an upward trend in the club's attendance average. However, this observed level is close to the synthetic value in most years analyzed post-treatment. Grêmio built a new stadium but kept matches in the old stadium (Estádio Olímpico) until Arena do Grêmio's inauguration, at end of 2012. The observed values and Synthetic Grêmio line. This pattern persists until 2016 when the observed value becomes lower than the synthetic value in later years.

Table 4

Averages and weights of predictive variables for treated clubs and their synthetics

Clubs	Variables	Treated	Synthetic	Weight
Athletico-PR	Points Percentage (%)	48.10	47.78	0.257
	CBF Ranking	8,486	8,398	0.272
	Players Value (€ millions)	22.98	27.48	0.309
	Total Assets (€ millions)	345.97	314.80	0.161
Bahia	Points Percentage (%)	45.79	45.80	0.039
	CBF Ranking	5,434	5,430	0.029
	Players Value (€ millions)	13.35	14.53	0.004
	Total Assets (€ millions)	74.23	87.38	0.928
Corinthians	Points Percentage (%)	55.99	54.09	0.494
	CBF Ranking	10,956	10,202	0.001
	Players Value (€ millions)	60.90	50.25	0.135
	Total Assets (€ milllions)	965.05	930.56	0.370
Cruzeiro	Points Percentage (%)	51.40	51.40	0.236
	CBF Ranking	7,605	7,716	0.391
	Players Value (€ millions)	57.26	57.27	0.373
	Total Assets (€ millions)	355.26	355.19	0.000
Flamengo	Points Percentage (%)	50.17	50.14	0.163
	CBF Ranking	8,280	7,491	0.048
	Players Value (€ millions)	59.50	60.78	0.786
	Total Assets (€ millions)	592.07	591.44	0.003
Fluminense	Points Percentage (%)	52.98	52.55	0.047
	CBF Ranking	9,025	8,994	0.075
	Players Value (€ millions)	53.57	55.11	0.001
	Total Assets (€ milllions)	375.94	326.13	0.877
Grêmio	Points Percentage (%)	54.21	54.21	0.137
	CBF Ranking	8,334	8,334	0.000
	Players Value (€ millions)	51.75	76.57	0.000
	Total Assets (€ millions)	228.31	421.37	0.863
Internacional	Points Percentage (%)	49.27	49.27	0.792
	CBF Ranking	10,129	10,222	0.110
	Players Value (€ millions)	71.01	69.83	0.000
	Total Assets (€ millions)	725.91	469.96	0.098
Palmeiras	Points Percentage (%)	49.71	49.71	0.264
	CBF Ranking	9,767	9,727	0.000
	Players Value (€ millions)	54.89	64.21	0.693
	Total Assets (€ millions)	384.17	384.03	0.043

Note. Source: by authors

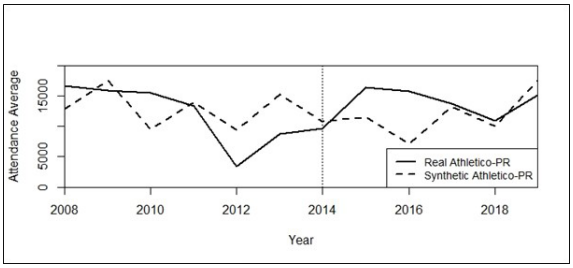

Figure 1. Attendance averages of Real Athletico-PR and Synthetic Athletico-PR
Note. Source: developed by authors

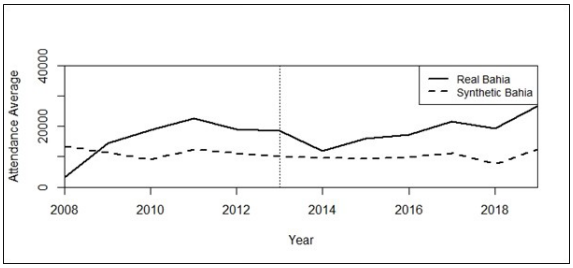

Figure 2. Attendance averages of Real Bahia and Synthetic Bahia
Note. Source: developed by authors

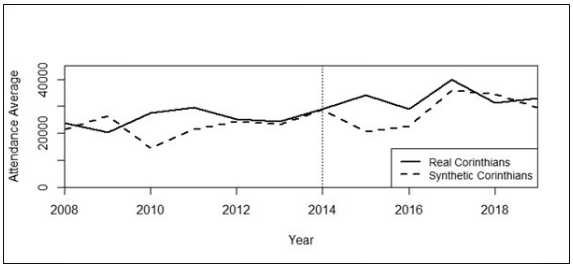

Figure 3. Attendance averages of Real Corinthians and Synthetic Corinthians
Note. Source: developed by authors

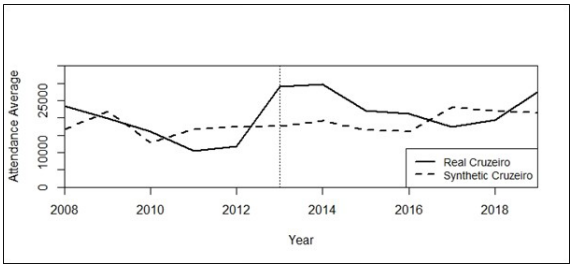

Figure 4. Attendance averages of Real Cruzeiro and Synthetic Cruzeiro
Note. Source: developed by authors

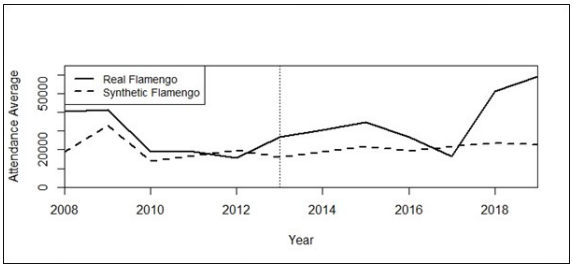

Figure 5. Attendance averages of Real Flamengo and Synthetic Flamengo
Note. Source: developed by authors

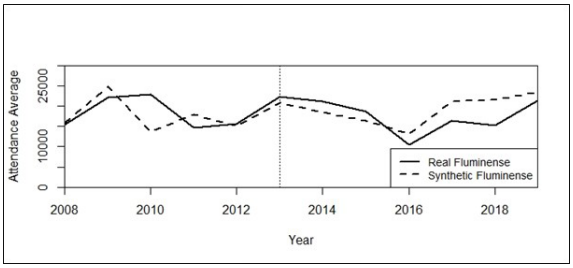

Figure 6. Attendance averages of Real Fluminense and Synthetic Fluminense
Note. Source: developed by authors

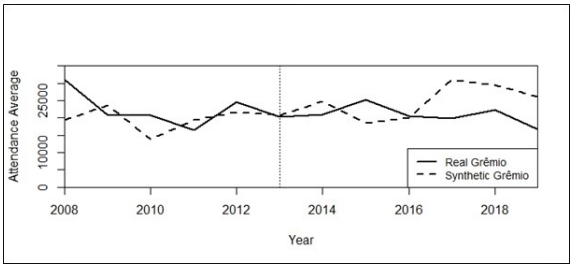

Figure 7. Attendance averages of Real Grêmio and Synthetic Grêmio
Note. Source: developed by authors

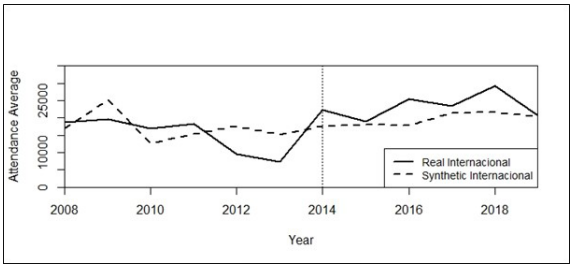

Figure 8. Attendance averages of Real Internacional and Synthetic Internacional
Note. Source: developed by authors

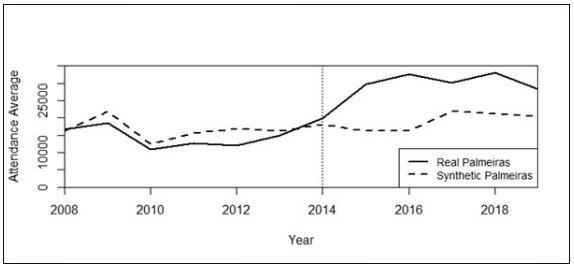

Figure 9. Attendance averages of Real Palmeiras and Synthetic Palmeiras
Note. Source: developed by authors

The estimated p-values through the placebo tests and presented in Table 5 for each club confirm that only treatment referring to Palmeiras rejected the null hypothesis, but only 10%, considering then that club was the only one to obtain some significant difference in your Attendance Average in the post-treatment period from building a stadium.

Table 5
MSPE indicator and p-values

Clubs	**MSPE**	***p-value***
Athletico-PR	1.0908	0.53
Bahia	1.2717	0.67
Corinthians	1.3351	0.40
Cruzeiro	0.9326	0.73
Flamengo	3.5547	0.27
Fluminense	0.8155	0.73
Grêmio	1.5454	0.60
Internacional	0.8059	0.73
Palmeiras	20.0039	0.07*

Note. Source: developed by authors
p-value < 0.01 = ***; p-value < 0.05 = **; p-value < 0.10 = *

Difference-in-difference models were first estimated for each of the treated groups and then followed by a two-way fixed effects model (Goodman-Bacon, 2018). By coincidence, stadiums inaugurated in 2014 are all private, while among arenas inaugurated in 2013, only Arena do Grêmio is private, the others are public stadiums used in the 2014 World Cup. Thus, removing Grêmio from the sample, we have two different groups: in 2013, public arenas were built for the World Cup; and in 2014, private stadiums, three of them used in the 2014 World Cup. So, we verified if there was a difference in the effects on attendance among new and renovated, public and private stadiums (Table 6).

Table 6
Treatment Effect by Differences-in-Differences

	Group 1 (2013)		Group 2 (2014)	
Before	**Attendance Average**	**p-value**	**Attendance Average**	**p-value**
Control (C)	≅13,000		≅14,000	
Treated (T)	≅19,000		≅17,000	
A = Diff (T-C)	6,184	0.003***	2,902	0.133
After	**Attendance Average**	**p-value**	**Attendance Average**	**p-value**
Control (C)	≅14,000		≅14,000	
Treated (T)	≅24,000		≅25,000	
B = Diff (T-C)	9,817	0.000***	10,429	0.000***
Diff-in-Diff (B-A)	**3,632**	**0.184**	**7,527**	**0.006*** **
Observations	264		264	
R^2	0.14		0.12	

Note. Source: developed by authors
p-value < 0.01 = ***; p-value < 0.05 = **; p-value < 0.10 = *

In Group 1, the difference between the control group and the treated group is significant in both periods. However, the Differences-in-Differences coefficient was not statistically significant. The difference between the differences within Group 1 is 3,632 on average. For group 2, results show that the mean difference in the pre-treatment period between the treated and control group is 2,902 (attendance average). This difference is not significant. For the post-treatment period, the mean has a value of 10,429, and the attendance average of differences-in-differences is 7,527, significant at 5% in both cases.

Table 7 presents these previous results and expanded estimation of these models, with the inclusion of other controls. The regressions adopt robustness for heteroscedasticity.

Models 1 and 4 exactly represent the results in Table 5 for the two treated groups. Models 2 and 5 were estimated with attendance average in ln as the dependent variable. Model 3, referring to Group 1, with the addition of control variables, did not show significance at 5% in the interaction variable, but at 10%. Thus, the result changes with the transformation of the dependent variable and with the addition of other variables.

Regarding the results of models 3 and 6, with the addition of controls, Points Percentage, Players Value, and Total Assets were statistically significant, with a positive coefficient, in both estimations. Points Percentage significance indicates that the sporting performance is decisive for an increase in attendance, while other significant variables pointed to an effect of the size of the clubs on variation in attendance average.

The results for Group 2 showed that stadiums inaugurated in 2014 had a significant impact on the attendance average. These arenas are private, belonging to Athletico-PR, Corinthians, Internacional, and Palmeiras. However, synthetic control results showed a statistical significance at 10% only for Palmeiras attendance average.

Considering the regression model results, as well as synthetic control, the presence of Palmeiras in the sample may be responsible for the 2014 treated group obtaining statistical significance in the interaction variable. To test this hypothesis, we removed Palmeiras from the sample, and the results are shown in Tables 8 and 9.

Table 7

Differences in Differences Regression Models by Treated Groups in Different Years

	Group 1 (2013)			Group 2 (2014)		
	Model 1	**Model 2**	**Model 3**	**Model 4**	**Model 5**	**Model 6**
Treated	6,184***	0.4310***	0.0860	2,902*	0.2665**	-0.4131**
	(2,078)	(0.1301)	(0.1921)	(1,498)	(0.1130)	(0.1576)
Time	920	0.0139	-0.0226	402	-0.0230	-0.1736*
	(1,072)	(0.0870)	(0.1126)	(1,197)	(0.0888)	(0.0895)
Treated*Time	3,633	0.2308	0.3383*	7,528***	0.4398***	0.6123***
	(2,979)	(0.0625)	(0.2025)	(2,417)	(0.1520)	(0.1720)
Ln Points Perc.			0.3894***			0.3246**
			(0.1272)			(0.1315)
Ln CBF Ranking			-0.0452			-0.0173
			(0.0593)			(0.0522)
Ln Players Value			0.2820***			0.3238***
			(0.0455)			(0.0472)
Ln Total Assets			0.1767***			0.1713***
			(0.0385)			(0.0425)
Constant	13,160***	9.3309***	0.0579	13,793***	9.3649***	-0.3989
	(728)	(0.0625)	(0.6960)	(746)	(0.0589)	(0.7522)
Dependent Variable	Attendance Average	Ln Attendance Average	Ln Attendance Average	Attendance Average	Ln Attendance Average	Ln Attendance Average
Period	2008/2019	2008/2019	2011/2019	2008/2019	2008/2019	2011/2019
Observations	264	264	198	264	264	198
R^2	0.1395	0.1200	0.5845	0.1169	0.1031	0.5678

Note. Source: developed by authors
p-value < 0.01 = ***; p-value < 0.05 = **; p-value < 0.10 = *

Table 8

Treatment effect by Diff-in-Diff without Palmeiras

	Group 1 (2013)		Group 2 (2014)	
Before	**Attendance Average**	**p-value**	**Attendance Average**	**p-value**
Control (C)	≅13,000		≅14,000	
Treated (T)	≅19,000		≅17,000	
A = Diff (T-C)	6,242	0.003***	3,701	0.094*
After	**Attendance Average**	**p-value**	**Attendance Average**	**p-value**
Control (C)	≅13,000		≅14,000	
Treated (T)	≅24,000		≅23,000	
B = Diff (T-C)	11,000	0.000***	9,005	0.000***
Diff-in-Diff (B-A)	**4,328**	**0.109**	**5,303**	**0.090***
Observations	252		252	
R^2	0.16		0.08	

Note. Source: developed by authors
p-value < 0.01 = ***; p-value < 0.05 = **; p-value < 0.10 = *

The estimate of Differences in Differences of Group 2 of the model with Palmeiras was 7,527, while without Palmeiras it drops to 5,303, being significant at 10%. That is, according to the DD model, even without Palmeiras, new private stadiums were of significant importance in attendance average increasing. This result is corroborated by the regression models presented in Table 9. The interaction Treated*Time is statistically significant when controls are added (in a model with smaller data, with three years less in the pre-treatment period).

Table 10 shows the mean estimates of the two-way fixed effects model for data with and without Palmeiras.

Table 9

Differences-in-Differences Regression Models by Treated Groups in Different Years Without Palmeiras

	Group 1 (2013)			Group 2 (2014)		
	Model 7	**Model 8**	**Model 9**	**Model 10**	**Model 11**	**Model 12**
Treated	6,242***	0.4432***	0.0841	3,701**	0.2935**	-0.4486**
	(2,093)	(0.1317)	(0.1918)	(1,811)	(0.1386)	(0.2003)
Time	224	-0.0223	-0.0423	402	-0.0230	-0.1690*
	(1,085)	(0.0896)	(0.1163)	(1,197)	(0.0888)	(0.0902)
Treated*Time	4,328	0.2669	0.3704*	5,303*	0.3428*	0.5736***
	(2,985)	(0.1615)	(0.2015)	(2,860)	(0.1803)	(0.2126)
Constant	13,101***	9.3187***	0.4159	13,793***	9.3649***	-0.4428
	(767)	(0.0657)	(0.7206)	(746)	(0.0589)	(0.7767)
Dependent Variable	Attendance Average	Ln Attendance Average	Ln Attendance Average	Attendance Average	Ln Attendance Average	Ln Attendance Average
Control Variables	No	No	Yes	No	No	Yes
Period	2008/2019	2008/2019	2011/2019	2008/2019	2008/2019	2011/2019
Observations	252	252	189	252	252	189
R^2	0.1604	0.1382	0.5802	0.0768	0.0718	0.5500

Note. Source: by authors
p-value < 0.01 = ***; p-value < 0.05 = **; p-value < 0.10 = *

Table 10

Differences-in-Differences Estimates with and without Palmeiras by two-way fixed effects model

Diff-in-Diff Comparison	**With Palmeiras**		**Without Palmeiras**	
	Weight	**Mean Estimator**	**Weight**	**Mean Estimator**
Earlier T vs. Later C	0.019	8,124	0.017	9,517
Later T vs. Earlier C	0.023	10,973	0.020	9,910
T vs. Never treated	0.958	6,782	0.963	5,567
Diff-in-diff estimate		6,904		5,719

Note. Source: developed by authors
T = Treated group; C = Control group

Figures 10 and 11 show Differences-in-Differences estimates using a two-way fixed effects model, considering data with and without Palmeiras, and only a dummy variable composed of the years in which there was treatment per club. The first graph shows a difference when Group 2 (2014, Later Group) is treated and Group 1 (2013, Earlier Group) is the control group. This difference presents an average estimator of 10,973. When treatment and control positions are reversed, the mean estimator is 8,124. This comparison in the second graph, without data from Palmeiras, corroborates the importance of the club for the significance of the Group 2 treatment. We noticed that marks between the two comparisons (Group 2 Treated and Group 1 Control and Group 1 Control and Group 2 Treated) are closer, with mean estimators of 9,910 and 9,517 respectively.

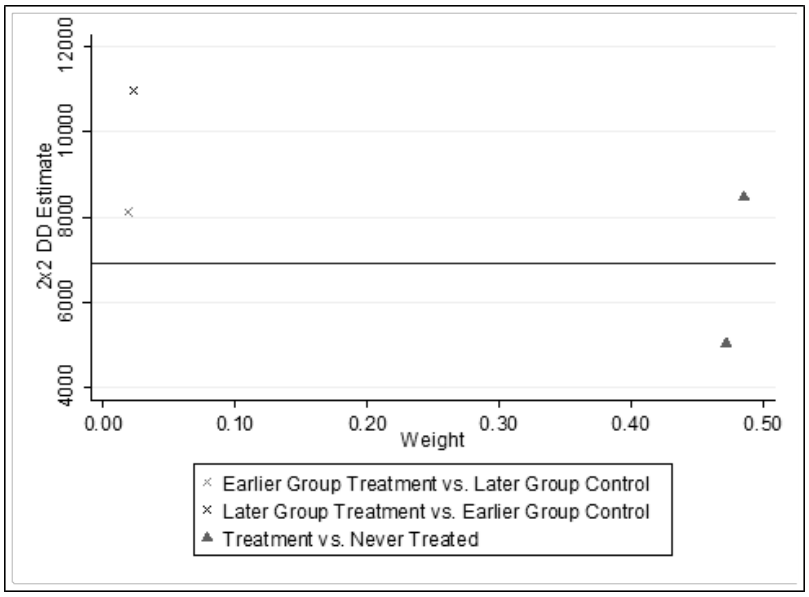

Figure 10. Graphs of Differences-in-Differences Estimates with Palmeiras by two-way fixed effects model
Note. Source: developed by authors

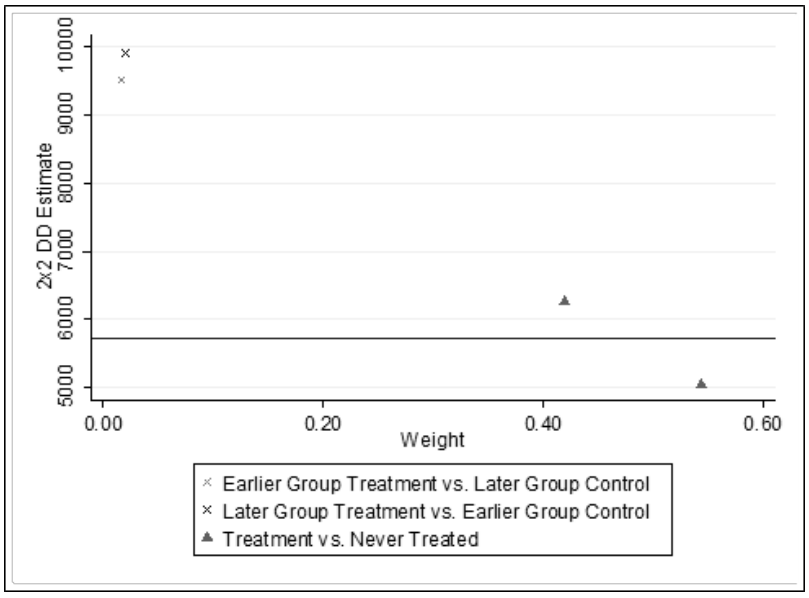

Figure 11. Graphs of Differences-in-Differences Estimates without Palmeiras by two-way fixed effects model
Note. Source: developed by authors

5. CONCLUSIONS

The purpose of this paper was to analyze the impact of stadium construction or renovation, especially for the 2014 World Cup, on the attendance average of Brazilian football clubs. However, some clubs, in addition to those whose stadium was used in World Cup, also proposed to change their structures to better serve their public, and consequently attract more people to the matches and obtain more revenue.

Considering the result of synthetic control models, eight clubs analyzed did not have their attendance average impacted by new arenas, Palmeiras being the exception. We identified an impact in attendance average for the group of clubs that started to play their matches in their stadiums, considering the most complete estimation, with controls, in the Difference-in-Differences models.

The results also indicate that, even with large investments in infrastructure, there was no significant increase in the attendance average of clubs playing matches in public stadiums. The sports performance proved to be positively significant to explain the attendance average logarithm. Thus, we considered that some of the treated clubs in the sample may have increased or decreased attendance averages according to variations in their sporting performance.

We tried to present a different way of analyzing the impact on attendance of new stadiums from those found in literature, observing impacts by a club (with synthetic control method) and in aggregate form for a set of clubs that started to perform their matches in new or refurbished arenas (with difference-in-differences models).

The paper's main limitation is the data. Firstly, the treated clubs are considered major clubs in Brazil, in terms of size and importance. Although there are other large teams in the donor pool, much of the control group is not at the treaty level, but this problem is mitigated by the methodology chosen and weights assigned in synthetic control. Also, the Attendance Average output variable was assembled from multiple sources. In addition, predictive variables also suffer from a lack of data, such as CBF Ranking and Total Assets, since the methodology of the first one was implemented in 2011, and Total Assets depend on the club's balance sheets, which are not always available during the entire sample period. Another important issue is the difference among treated clubs. Only Corinthians and Grêmio changed stadiums with no need for renovations in their previous ones. The other clubs in the period analyzed before the treatment started to play their matches in other places, even in other cities, which may have changed the synthetic control estimation, such as an abnormal reduction of this attendance average in the pre-treatment period. However, the results themselves showed that there was no impact of the presence of new stadiums in most clubs.

As a suggestion for future research, other authors can use synthetic control methodology for other leagues and other sports, as an alternative for papers on the impact of new arenas on attendance average, as also, with synthetic control, can estimate the effects on other variables of interest such as Ticket Revenue, which perhaps better justifies the new stadiums.

REFERENCES

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*, *105*(490), 493-505. https://doi.org/10.1198/jasa.2009.ap08746

Abadie, A., Diamond, A., & Hainmueller, J. (2011). Synth: An R package for synthetic control methods in comparative case studies. *Journal of Statistical Software*, *42*(13), 1-17. https://doi.org/10.18637/jss.v042.i13

Abadie, A., Diamond, A., & Hainmueller, J. (2015). Comparative politics and the synthetic control method. *American Journal of Political Science*, *59*(2), 495-510. https://doi.org/10.1111/ajps.12116

Abadie, A., & Gardeazabal, J. (2003). The economic costs of conflict: A case study of the Basque Country. *American Economic Review*, *93*(1), 113-132.

Ando, M. (2015). Dreams of urbanization: Quantitative case studies on the local impacts of nuclear power facilities using the synthetic control method. *Journal of Urban Economics*, *85*, 68-85. https://doi.org/10.1016/j.jue.2014.10.005

Castilho, C. T., Evrard, B., & Charrier, D. (2017). 2014 FIFA World Cup in Brazil: Gentrification of Brazilian football. *Sociology and Anthropology*, *5*(9), 703-712. http://doi.org/10.13189/sa.2017.050902

Cavallo, E., Galiani, S., Noy, I., & Pantano, J. (2013). Catastrophic natural disasters and economic growth. *Review of Economics and Statistics*, *95*(5), 1549-1561. https://doi.org/10.1162/REST_a_00413

Clapp, C. M., & Hakes, J. K. (2005). How long a honeymoon? The effect of new stadiums on attendance in Major League Baseball. *Journal of Sports Economics*, *6*(3), 237-263. https://doi.org/10.1177/1527002504265957

Coates, D., & Humphreys, B. R. (2005). Novelty effects of new facilities on attendance at professional sporting events. *Contemporary Economic Policy*, *23*(3), 436-455. https://doi.org/10.1093/cep/byi033

Coates, D., & Humphreys, B. R. (2007). Ticket prices, concessions and attendance at professional sporting events. *International Journal of Sport Finance*, *2*(3), 161-172.

Gitter, S. R., & Rhoads, T. A. (2014). Stadium construction and minor league baseball attendance. *Contemporary Economic Policy*, *32*(1), 144-154. https://doi.org/10.1111/coep.12016

Goodman-Bacon, A. (2018). *Difference-in-differences with variation in treatment timing* (Working Paper No. w25018). National Bureau of Economic Research.

Kahane, L., & Shmanske, S. (1997). Team roster turnover and attendance in major league baseball. *Applied Economics*, *29*(4), 425-431. https://doi.org/10.1080/000368497326921

Leadley, J. C., & Zygmont, Z. X. (2005). When is the honeymoon over? National Basketball Association attendance 1971-2000. *Journal of Sports Economics*, *6*(2), 203-221. https://doi.org/10.1177/1527002504263399

Leadley, J. C., & Zygmont, Z. X. (2006). When is the honeymoon over? National Hockey League attendance, 1970-2003. *Canadian Public Policy*, *32*(2), 213-232.

McEvoy, C. D., Nagel, M. S., DeSchriver, T. D., & Brown, M. T. (2005). Facility age and attendance in Major League Baseball. *Sport Management Review*, *8*(1), 19-41. https://doi.org/10.1016/S1441-3523(05)70031-0

Poitras, M., & Hadley, L. (2006). Do new major league ballparks pay for themselves? *The Journal of Business*, *79*(5), 2275-2299. https://doi.org/10.1086/505235

Poli, R., Ravenel, L., & Besson, R. (2019). *Attendances in football stadia (2003-2018)*. The 44th CIES Football Observatory Monthly Report.

Rascher, D. A., Brown, M. T., Nagel, M. S., & McEvoy, C. D. (2012). Financial risk management: The role of a new stadium in minimizing the variation in franchise revenues. *Journal of Sports Economics*, *13*(4), 431-450. http://doi.org/10.1177/1527002512450281

Serrano, R., García-Bernal, J., Fernández-Olmos, M., & Espitia-Escuer, M. A. (2015). Expected quality in European football attendance: Market value and uncertainty reconsidered. *Applied Economics Letters*, *22*(13), 1051-1054. https://doi.org/10.1080/13504851.2014.997919

Silva, B. C. (2020). *Package 'SCTools': extensions for Synthetic Controls Analysis*. Retrieved March 25, 2021, from https://cran.r-roject.org/web/packages/SCtools/SCtools.pdf.

Soebbing, B. P., Mason, D. S., & Humphreys, B. R. (2016). Novelty effects and sports facilities in smaller cities: Evidence from Canadian hockey arenas. *Urban Studies*, *53*(8), 1674-1690.

Villa, G., Molina, I., & Fried, R. (2011). Modeling attendance at Spanish professional football league. *Journal of Applied Statistics*, *38*(6), 1189-1206. https://doi.org/10.1080/02664763.2010.491859

Zygmont, Z. X., & Leadley, J. C. (2005). When is the honeymoon over? Major League Baseball attendance 1970-2000. *Journal of Sport Management*, *19*(3), 278-299. http://doi.org/10.1123/jsm.19.3.278

AUTHOR'S CONTRIBUTION

MD was responsible for the theoretical-methodological construction, production, data collection, and analysis, as well as the final writing. LS was responsible for the theoretical-methodological construction and final review.

ACKNOWLEDGMENTS

The present work was carried out with the support of the Coordination for the Improvement of Higher Education Personnel – Brazil (CAPES) – Financing Code 001, and the National Council for Scientific and Technological Development (CNPq).

CONFLICTS OF INTEREST

The authors declare no conflicts of interest.