



Spatial strategies in Brazilian Franchising; Behavior categories and Performance Outcome

Eugenio Jose Silva Bitti, Muriel Fadairo, Cintya Lanchimba, Vivian-Lara Silva

► To cite this version:

Eugenio Jose Silva Bitti, Muriel Fadairo, Cintya Lanchimba, Vivian-Lara Silva. Spatial strategies in Brazilian Franchising; Behavior categories and Performance Outcome. 2016. halshs-01292057

HAL Id: halshs-01292057

<https://shs.hal.science/halshs-01292057>

Preprint submitted on 22 Mar 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

WP 1614 – March 2016

Spatial strategies in Brazilian Franchising; Behavior categories and Performance Outcome

Eugenio Jose Silva Bitti, Muriel Fadaïro, Cintya Lanchimba,
Vivian-Lara Silva

Abstract:

This empirical article deals with the location strategy and performance outcomes of franchised chains in Brazil. Brazilian franchising has experienced vertiginous process of expansion in recent years, being present in different regions of this country of continental size, including the most remote and less developed. Based on the background literature in economics and management regarding location choices and spatial competition in retailing, we use a new and unique dataset to distinguish several behavior categories in spatial strategies of franchising chains in Brazil, via a two-step cluster analysis. The performance outcomes are then studied with econometric estimations on panel data. Our results provide evidence that the choice for agglomeration, and the location in areas with a high population and a high human development index, lead to higher chain performances.

Keywords:

Franchising, Location Strategy, Spatial Competition, Emerging Market, Panel Data

JEL codes:

M21, C23

Spatial strategies in Brazilian Franchising; Behavior categories and Performance Outcome

Eugenio Jose Silva Bitti

School of Economics, Business Administration and Accounting
University of São Paulo, Brazil
Av. Bandeirantes 3900 - Monte Alegre –
CEP 14040-905 Ribeirão Preto - SP
Telephone: +55 (16) 3315-3899
Email: ebitti@usp.br

Muriel Fadairo

Univ. Lyon, UJM Saint-Etienne
CNRS, GATE L-SE UMR 5824
F-42023 Saint-Etienne, France
Telephone: +336 79 65 88 25
Email: muriel.fadairo@univ-st-etienne.fr

Cintya Lanchimba

Escuela Politécnica Nacional, Ecuador
Ladron de Guevera E11-253, Quito 170517, Ecuador
Telephone: +593 988 61 43 97
Email: cintya.lanchimba@epn.edu.ec

Vivian-Lara Silva

Center for Organization Studies (CORS)
Group for Study and Research in Strategy and Vertical Coordination (GEPEC)
University of São Paulo (USP) / Brazil
Campus Fernando Costa - Pirassununga
Av. Duque de Caxias Norte, 225, ZIP Code 13635-900
Telephone: +55 19 35.65.41.60
Email: vivianlara@usp.br

ABSTRACT

This empirical article deals with the location strategy and performance outcomes of franchised chains in Brazil. Brazilian franchising has experienced vertiginous process of expansion in recent years, being present in different regions of this country of continental size, including the most remote and less developed. Based on the background literature in economics and management regarding location choices and spatial competition in retailing, we use a new and unique dataset to distinguish several behavior categories in spatial strategies of franchising chains in Brazil, via a two-step cluster analysis. The performance outcomes are then studied with econometric estimations on panel data. Our results provide evidence that the choice for agglomeration, and the location in areas with a high population and a high human development index, lead to higher chain performances.

Keywords

Franchising, Location Strategy, Spatial Competition, Emerging Market, Panel Data.

JEL Codes

M21; C23

Acknowledgements:

This research has received funding from the University of São Paulo, Brazil, and from the University of Lyon, France. We would like to acknowledge our respective institutions. We also wish to thank the Brazilian Franchising Association, more particularly Ricardo Camargo, President for the period 2010-2011, Rogério Feijó, Executive Director, and Lilian Greziou, Market Analyst, for their valuable help regarding the construction of our dataset.

We are grateful to Gérard Cliquet and Josef Windsperger for their insight and support. Ms. Bianca Maria Silva provided excellent research assistance. Additional thanks to the participants of the franchise session of the 7th International Conference on Economics and Management of Networks – EMNet 2015, Cape Town, South Africa, December 3-5, for interesting feedback. All errors are ours.

**“Brazil has a large expansion of franchising opportunity, beyond the capital and big cities.
Brazil is a country of continental dimensions, with several pockets of development.”**

Cristina Franco

Chair of the Brazilian Franchising Association (2015)

1. Introduction

This empirical article studies the location strategy and performance outcome of franchised chains in Brazil.

Brazilian franchising has experienced vertiginous process of expansion in recent years, being present in different regions of the country, including the most remote and less developed.

At the moment, Brazil is going through an important economic and political crisis. For 2014, the most optimistic projections attribute growth of less than one digit to the national gross domestic product. As a reflection of this low economic performance, the Brazilian retail sector showed modest growth of 2.2% in 2014. However, franchising in Brazil continues to grow. Thus, according to the Brazilian Franchising Association (ABF), the total income of Brazilian franchising in 2014 grew 7.7% compared to 2013. This growth continued in 2015. Indeed, during the first half of 2015, the income increased 11.2% compared to the same period in 2014, while only for the 2nd quarter of 2015, the income growth was estimated at 13.1% (ABF, 2015).

Despite the current economic crisis, Brazil stands as one of the main emerging economies for the 21st century, and is part of the BRICS, in addition with Russia, India, China and South Africa. One main feature regrouping these emerging countries is their continental dimension, specifically for Brazil, Russia, India and China, which raises the issue of spatial inequalities and agglomeration phenomena in the economic development (Fally et al., 2010; Lessmann, 2014; Cheong & Wu, 2014). Thus, more than elsewhere, the geographical dimension is a critical point in this context and, at the microeconomic level, location choices are key strategical elements for the economic actors.

Regarding Brazilian franchising, the Southeast accounted for 58.8% in the spatial allocation of franchised units during the first half of 2015, followed by the South (14.9%), as presented in Appendix 1. These regions are the richest of the country. However, other regions have attracted the attention for network expansion. The Northeast is a special case, accounting for 14.1% of franchised units in Brazil in 2015, while it was estimated at around 2% in 2013/2014 (ABF, 2015). Central West and North, in turn, represent respectively 8% and 4.2% in 2015.

However, while franchising issues in the BRICS economies is a major field of interest at the moment (Dant & Grünhagen, 2014), and despite several interesting articles dealing more specifically with Brazilian franchising (Azevedo & Silva, 2005; Dant et al. 2008; Falbe & Welsh,

1998; Lafontaine & Oxley, 2004; Silva et al. 2009; Streed & Cliquet, 2013; Vance et al. 2011; Fadaïro & Lanchimba, 2013; Michael, 2014; Resende Melo, 2015), the locational aspects of franchising in Brazil remains mainly unexplored - the only exception being an article by Bitti et al. (2010) -, and represents a gap in the literature.

In the retailing literature, location is clearly, and for a long time, considered as a major determinant of firm performance (Jones & Simmons 1987), and as one of the most critical factor of performance in the case of retail networks (Cliquet, 1998). Thus, in their central articles in the retail location literature, Applebaum (1966) and Cliquet (1998) define the key concepts of store trade areas, market penetration, and potential sales (Applebaum, 1966), in addition with the important concept of entropy, measuring the territory coverage of a retail network (Cliquet, 1998; Streed & Cliquet, 2013).

However, as underlined by Ehrmann & Meiseberg (2011), locational strategy for network expansion has received little attention in the franchise literature. When addressed, the geographic dimension is most often taken into account as a way to evaluate the monitoring costs between headquarters and the outlets of franchise chains (Brickley & Dark, 1987; Norton, 1988; Minkler, 1990; Maness, 1996; Bitti & al, 2010; Combs and Ketchen, 2003; Perryman & Combs 2012).

Considering these gaps in the literature, and based on previous researches in economics and strategic management, this paper addresses the following issue: in the Brazilian case of franchised networks, is it possible to highlight different profiles of location choices, and to distinguish the most efficient location strategies?

To characterize the different profiles of networks, we use the concept of behavior categories, referring to Porter (1980)'s strategic groups, defined as sets of firms in an industry that display similar competitive profiles, and to Meyer et al. (1993)'s configurations, relating to groups of firms with a common organizational profile. Behavior categories in the locational strategy of franchising networks in Brazil are distinguished by means of a two-step cluster analysis. We use a unique and recent panel dataset enclosing interesting information in terms of geographical location. The study variables for the empirical analysis are derived from the literature relating retailing and location choices. Finally, we perform econometric estimations, and relate behavior categories in location choices to network performance outcomes.

The paper is organized as follows. Section 2 presents the background literature, and identifies relevant choices determining the locational strategy of multi-outlet chains. Section 3 describes the data, providing information on the data collection and the study variables, in addition with summary statistics. The cluster analysis and the estimations are respectively presented in Section 4 and Section 5. Section 6 concludes.

2. Related research

In view of the objective to investigate behavior categories in Brazilian franchising, this article is related to the clustering tradition of networks in the franchise literature. Several important contributions can be mentioned, Carney & Gedajlovic (1991), Combs & Ketchen (1999), Combs et al. (2004), Gonzalez-Diaz & Solis-Rodriguez (2015), which relate specific strategies characterizing groups of franchise chains with performance outcomes. As mentioned by Combs and Ketchen (1999), the clusters can be derived inductively or theoretically. We give priority to the second case, where the variables used for the clustering process are related to theoretical results. We present hereafter the theoretical background of our analysis.

Given the importance of location choices, previously emphasized, a vast literature dealing with consumer spatial behavior and retail location exists. The aim of this analytical section is not to review this wide research field, but to highlight the theoretical results, in economics and management, relevant to distinguish behavior categories in the location strategy of franchised chains in Brazil. This strategy relates to the choice of where to locate retail outlets. From the background literature, four main determinants of the geographical location strategy of franchised networks can be distinguished. We discuss them in the following.

2.1. The choice for agglomeration *versus* dispersion

Spatial competition, where customer behaviors are influenced by the cost of transportation and the relative locations of purchaser and purveyor, can give rise to a variety of retail location patterns. Agglomeration of retail outlets is one of the most relevant. Indeed, contrary to the aspatial competition of standard microeconomics models, spatial competition has monopolistic features leading to agglomeration. Space introduces monopolistic elements, as in spatial markets all firms enjoy some degree of monopoly power over their immediate market area.

In this analytical context, Hotelling (1929)'s pioneering contribution formally demonstrates the benefits of clustering. Hotelling's model shows that in a duopoly of maximizing firms selling identical products from fixed locations, in a bounded linear market, firms cluster at the center of the market area, guaranteeing themselves one-half share of the market. The modeled noncollusive duopolistic competition results in price stability, and is characterized by the principle of minimum differentiation regarding locational decisions.

This demonstration can be extended to several types of locations, not only spatial, for example location in terms of product differences. Still, Hotelling's analysis stands as the main normative

theory of retail location, and the principle of minimum differentiation as the most relevant explanation of the agglomeration of similar retail firms.

However, in addition to the benefits of clustering, the risk of cannibalization must be taken into account (Pancras et al. 2012), since we focus on the location choices of retail outlets within the same chain. Cannibalization relates to spatial competition between geographically proximate retail outlets, when the presence of a new unit captures sales of other outlets. The spatial strategy of franchising chains, depending on the choice of where to locate their retail units, is affected by this cannibalization effect. This intra-chain competition effect may thus influence the spatial choices in the way of dispersion *versus* agglomeration.

Pancras et al. (2012) estimate this cannibalization effect and provide evidence for a significant decay in cannibalization with distance. In a context of network expansion, they measure the net impact of a new store opening on the overall chain performance. This impact results from the joint and contradictory effect of incremental sales *versus* cannibalization. Estimating via econometrics and panel data the parameters of a demand model that captures spatial competition, they highlight a clear empirical link between cannibalization and distance.

With the concept of entropy, Cliquet (1998) defines a tool to measure the territory of geographic coverage of a retail chain. If a population of stores is concentrated in one spatial area, entropy is minimum. It is maximum with dispersion.

While the choice for agglomeration *versus* dispersion concerns the density of retail outlets, spatial strategies also deal with site selection decision, in other words with the choice to locate stores in a specific region or city.

2.2. The choice for a specific region or city

With the nearest center hypothesis postulated by the classical central place theory (Christaller, 1933; Losch, 1954), customers necessarily shop at the closest store. Gravity models underline also the importance of geographical distance, but demonstrate that the attractiveness of a retailing location, and therefore the choice to locate in a specific place, is not exclusively related to it. These models include thus a variety of other factors in the attraction concept.

Gravity models are derived from the laws of Newtonian physics. In this literature, Reilly (1931), Huff (1964) and Huff & Batsell (1977) played a pioneering role. Reilly's law of retail gravitation states that consumers trade off the cost of travel to a retail outlet with the attractiveness of alternative shopping opportunities.

More generally, gravitational models allow to study attractiveness phenomena and spatial interactions, considering other attributes in addition to distance related to the potential of a local

or regional area, as demographic weight, gross domestic product, or socioeconomic characteristics.

These models suggest thus that there are differences in location quality. As emphasized by Ehrmann & Meiseberg (2011), this is the case with the whole location theory (Ghosh & McLafferty 1982; Lee & McCracken 1982; Craig et al., 1984; Ghosh & McLafferty, 1987; Jones & Simmons, 1990; Simons 1992; Kelly et al., 1993; Peterson 2003; Christensen & Drejer, 2005; Park & Khan, 2006; Khan 1999; Park & Khan 2006).

Recognizing that some places are more lucrative, or have a greater potential to be profitable, than others, Ehrmann & Meiseberg (2011) study the location decisions in the expansion of franchising networks. These authors show that, in addition with endogenous network characteristics, exogenous location factors related to market perspective criteria, as population density or local income, affect location decisions.

2.3. Additional determinants of the location strategy

2.3.1. First mover advantage *versus* new entrant advantage

The timing is also part of spatial strategies. Indeed, the first firm to locate in a market can capture strategic sites. This first mover advantage acts thus as a barrier to entry and delays the entry of new competitors by affecting their locational choices.

Distinguishing competitive advantages at several stages of a franchise life (Lillis et al., 1976), and studying territorial expansion of franchised chains in geographically fragmented markets (Julian & Castrogiovanni 1995), Lillis et al. (1976) and Julian & Castrogiovanni (1995) develop the idea that when franchisors attempt to capitalize on the first mover advantage by preempting the most desirable locations and saturating the market, locations are dispersed.

In addition, the empirical findings of Carney & Gedajlovic (1991) highlight the heterogeneity of strategies in franchising chains. These authors argue that the need for spatial preemption, and the accompanying rapid growth, is particularly important when franchisors have innovative retail concepts.

Finally, considering the short- and long-run incentives to franchise, and monitoring problems related to geographic dispersion, Martin & Justis (1993) explain the intensive use of franchising in the early growth stages, arguing that growth and preemptive strategies are most important for young franchisors.

However, the alternative strategy can be interesting too. Later entrants may have an advantage since they are better able to react to changing conditions

2.3.2. Location choices and network ownership structure

As previously mentioned, in the franchise literature, the location issue is most often taken into account within the agency theory framework and the related monitoring cost problem.

Since the pioneering contributions of Rey & Tirole (1986a, 1986b), Mathewson & Winter (1983a, 1983b, 1984, 1985), agency theory, focusing on the delegation of tasks, is the main theoretical context to study franchising. The literature highlights free-riding behaviors in retail networks, and underlines the deriving need for the headquarters to monitor the outlets (Brickley & Dark 1987; Lafontaine & Slade 2001; Lal 1990; Michael 1999).

Network expansion and spatial dispersion increase the monitoring agency problem. In other words, with the geographical dispersion of the network, the costs of monitoring managers and enforcing quality standards increase. Monitoring outlets in geographic markets that are far from headquarters is more costly, due to the number of monitoring personnel required and to the related travel expenses (Carney and Gedajlovic, 1991).

Because they are independent firms, franchisees are more likely to provide efforts, requiring thus less monitoring. This situation would impact the ownership structure of retail networks, mixing strategically company-owned units and franchised units as part of their location strategy. This argument finds empirical support in the econometric work on franchise data.

Indeed, as emphasized by Perryman and Combs (2012), the most common test of the monitoring cost prediction in the framework of the agency theory investigates the relationship between the geographic dispersion of a chain and its ownership composition, that is the proportion of franchised outlets or, on the reverse, the proportion of company-owned outlets.

Brickley & Dark (1987), Norton (1988), Lafontaine (1992), Maness (1996), Combs and Ketchen (2003), Combs and Ketchen (2003), Bitti et al. (2010), provide evidence that outlets located close to headquarters are more likely to be company owned, while retail units in more distant locations are more likely to be franchised. Moreover, monitoring costs arising from the geographical dispersion of the chain would be sufficient to explain the franchise option. More generally, at the chain level, the higher the geographical dispersion, and thus the related monitoring cost, the lower the proportion of company-owned outlets.

Therefore, a clear link is established between spatial strategies of franchising networks and the chain ownership structure.

3. The Data

We use recent panel data for the period 2009-2014, regarding 335 franchise networks. The sample franchise chains occupy 1,397 Brazilian municipalities, and are present in the 26 Brazilian states.

Our unique dataset compiles information from three main distinct sources: the ABF's Official Franchise Guides, the websites of franchise chains, and the Brazilian Institute of Geography and Statistics (IBGE).

We present hereafter the data collection process, and then the study variables derived from the analytical framework. We discuss the relevance of the sample, comparing with the full population. Finally we provide summary statistics.

3.1. Data collection

Our first source is the ABF's Official Franchise Guide. This annual publication provides valuable information on the Brazilian franchise sector. More precisely, the guide contains data concerning chain characteristics, as the franchise fee, the total number of outlets, the required investment level of the franchisees, the number of employees, and so on. We obtained data for the period 2009-2014.

In order to achieve a better understanding of how the ABF collected the information used to construct our dataset, we performed interviews with executives from this organization. Specifically, we performed interviews with the President of the ABF, his executive director, and his marketing analyst, who is responsible for both collecting and pooling information provided by the franchised networks.

The ABF has an internal information system accessible to associated franchised chains, which are free to update information regarding their network in real time. The chains have an interest in keeping up-to-date information, since the Guide is an important communication channel with potential franchisees. Divergent or even fake information could damage the reputation of the chain, or make the search for potential franchisees more difficult. Therefore, there is a strong incentive for franchising networks to offer good quality information.

In addition, the ABF controls for the consistency of the data, analyzing systematically the annual variations before publishing official information. Any abnormal variation is checked and corrected.

However, the ABF does not already have a scientific advising council. For this reason, the ABF marketing department performs both the monitoring and the validation of provided data. If necessary, specialized consultants are hired for more specific studies.

The websites of the sample franchise networks are our second main source of information. The manual capturing of website data was used to perform control checks, and to complete the missing values from the ABF guides. In addition, we used this data source to collect spatial information regarding the Zip codes of cities where the franchised chains are established, in addition with the proportion of network outlets operated in shopping malls. When the network websites were incomplete, we used the PEGN publication, which is a SME specialized Brazilian media.

As our analysis requires localized data, we use a third source of information, the Brazilian Institute of Geography and Statistics (IBGE). We collected information about georeferencing (longitude and latitude) of the municipalities where franchising networks are established, the Gross Domestic Product (GDP), the population, and the Human Development Index (HDI).

The IBGE information comes from censuses conducted in 1991, 2000 and 2010. Municipalities HDI data are also available in the United Nations Development Program (UNDP) website for the years 1990, 2000 and 2010.

We pooled the different sources of information, performed crosschecks, and matched the data in order to stabilize the final sample enclosing the required information for our empirical analysis.

Finally, as all the Brazilian chains are not associated to the ABF, we controlled for the bias in our sample compared to the franchise network population in Brazil. To do so, we performed an initial analysis with some Brazilian franchising consolidated information available on the ABF website. Table 1 presents the 11 segments of the franchising industry in Brazil, based on the ABF taxonomy. For each segment, we tabulated data regarding the total income of the segment, the number of chains and the number of outlets (company-owned and franchised) in 2014. It is thus possible to compare the franchising sector in Brazil as a whole, to both the whole “ABF chains” and our final sample. Table 1 presents the absolute values and the ratios for each sector in comparison to the total for each of the three variables used in the bias analysis¹.

Two segments appear over-represented in our sample (as in ABF as a whole): Food and Shoes & Accessories. This over-representation occurs for the three variables. The proportion of chains in these segments is between ABF associated chains compared to the Brazilian franchising sector

¹ To calculate the total sales (income) of the networks associated to the ABF (and therefore of the sample chains), we computed the average income per store, which is multiplied by the total number of stores. We treated the missing values using the average value for each segment.

as a whole; and it is greater in the sample compared to ABF. This is also the case for the total income and for the total number of outlets. Conversely, segments of House & Building and Business, Services & Other Retails appear underrepresented. The sample contains 11.3% of all chains operating in Brazil, 26.6% of the outlets. The sample networks represent 31.2% of all sector incomes.

In summary, data indicate a bias in the sample relative to the population: the sample networks are typically larger, and the sector representation does not reflect exactly the population. For this reason, some dose of parsimony should be included when analyzing the results.

However, we insist on the good quality of the data, compiling three different and complementary sources of information. Our unique sample consists of a six-year panel, in a wide range of activities, and contains geographical information. The aforementioned bias regarding some specific segments calls for the use of sector dummies in the econometric estimations in order to control it.

Table 1 – Bias Analysis – Comparison between the Brazilian franchising sector, ABF associated chains, and the sample (source: ABF 2014)

Segments ABF Classification	Total Income (R\$ million)						Total Number of Outlets						Total Number of Chains					
	Sector		ABF		Sample		Sector		ABF		Sample		Sector		ABF		Sample	
	R\$	%	R\$	%	R\$	%	Un.	%	Un.	%	Un.	%	Chains	%	Chains	%	Chains	%
Shoes & Accessories	8.847	6,9	7.247	8,8	5.982	14,9	8.285	6,6	6.390	8,1	4.787	14,3	216	7,3	92	9,8	39	11,6
Food	25.635	19,9	22.472	27,3	18.562	46,2	21.720	17,3	19.394	24,6	11.085	33,2	685	23,3	258	27,4	119	35,5
House & Building	7.320	5,7	4.489	5,5	701	1,7	5.380	4,3	4.262	5,4	424	1,3	159	5,4	69	7,3	4	1,2
Advertising, Informatics & Electronics	3.867	3	525	0,6	230	0,6	3.336	2,7	1.049	1,3	207	0,6	142	4,8	34	3,6	4	1,2
Education	8.538	6,6	6.885	8,4	4.264	10,6	14.732	11,7	10.714	13,6	7.307	21,9	298	10,1	98	10,4	44	13,1
Sports, Health, Beauty & Leisure	23.288	18,1	11.436	13,9	4.581	11,4	20.670	16,5	11.928	15,1	4.956	14,8	547	18,6	136	14,4	54	16,1
Hoteling & Tourism	9.329	7,2	3.483	4,2	1.023	2,5	2.674	2,1	791	1	160	0,5	49	1,7	14	1,5	3	0,9
Cleaning	1.159	0,9	1.403	1,7	382	0,9	3.426	2,7	2.586	3,3	800	2,4	97	3,3	32	3,4	5	1,5
Business, Services & Other Retails	26.726	20,7	15.332	18,6	828	2,1	28.616	22,8	5.041	6,4	1.015	3	305	10,4	87	9,2	22	6,6
Cars	4.490	3,5	3.872	4,7	1.244	3,1	8.032	6,4	12.733	16,2	648	1,9	115	3,9	29	3,1	4	1,2
Clothes	9.677	7,5	5.115	6,2	2.369	5,9	8.770	7	3.939	5	1.998	6	329	11,2	94	10	37	11
TOTAL	128.876		82.259		40.166		125.641		78.827		33.387		2.942		943		335	

3.2. Measurement

We present hereafter the study variables, used in the empirical estimations, and in the cluster analysis determining behavior categories in spatial strategies of franchising chains in Brazil.

The variables for the cluster analysis are based on the background literature regarding location choices in franchising or, more broadly in retailing. In line with our analytical framework, these variables are the following: the network geographic dispersion, the percentage of network outlets located in shopping malls, the weighted average human development index of the network location area, and the weighted average population of the network location area. In addition with these localized variables, we include complementary variables related to the location strategy: the network age, the percentage of company-owned units, and the chain brand value.

In the econometric estimations, the dependent variable is the average monthly income per outlet. The core explanatory variable is the network location strategy constructed via the cluster analysis. We add variables to control for the influence of the sector, of the headquarter nationality, and of the ABF certification of excellence.

More details are provided in the following.

3.2.1. Variables defining the chain location strategy (Cluster Analysis)

3.2.1.1. Location variables

Network Geographic Dispersion

The construction of this variable is based on the geographic location of the network outlets. First, we surveyed the addresses of all the sample network outlets in 2011 and 2014. We created thus two specific data sets containing information regarding the Brazilian municipalities where the franchising chains are established. Then, we georeferenced the cities where the stores are located, getting their latitude and their longitude.

On this basis, we constructed geographic clusters for each network regrouping stores located in the same area. Each cluster is thus an area of agglomeration of network outlets. We defined the center for each cluster so that the cumulative distance between the stores and the cluster center is the smallest possible¹.

The distance criterion is the Euclidean distance based on the latitude and the longitude of the cities where each store is located. For example, assuming that a store A is located in the point $(x_A,$

¹ To split the data and to determine the appropriate number of clusters for each chain, we used the “R” package “NbClust” and the “Kmeans” R function.

y_A) and a store B is located in the point (x_B, y_B) , the Euclidean distance between them is defined as follows:

$$D(A, B) = \sqrt{((x_A - x_B)^2 + (y_A - y_B)^2)}$$

Our variable, the network geographic dispersion, is measured as the distance between the stores in the clusters, to which we add the distance between the centers of the chain clusters.

Percentage of network outlets located in shopping malls

Jointly with the outlet addresses, we also captured whether the store is located in some kind of shopping mall. Thus, for each chain in our sample, we are able to define the proportion of outlets located in shopping centers, in a percentage base.

Weighted average HDI of the network location areas

The Human Development Index (HDI) is a socio-economic indicator developed by the United Nations Development Program (UNDP) in the early 90's to measure both the countries socioeconomic development of countries and their constituent units (states, provinces, municipalities, etc.).

This multidimensional indicator aims at going beyond the only economic issue, and captures, in addition with economic features, some aspects related to public health and education. Moreover, given its synthetic character, the index allows establishing a ranking and, as a result, certain comparability between countries and regions. More precisely, the HDI compiles three indexes relating to three distinct socioeconomic dimensions:

- Longevity – which measures life expectancy at birth and also reflects the health of the population;
- Education – an index operationalized by a combination of the adult literacy rate and the combined enrollment rate in primary, secondary and higher education;
- Income – a measure of population's purchasing power based on GDP per capita adjusted to the local cost of living to make it comparable across countries and regions, through the methodology known as purchasing power parity.

The IBGE provides information regarding the HDI for each Brazilian state and municipality. Since its inception in the early 90's, the HDI has been recorded three times (1990, 2000 and 2010). In our study, we use the index calculated by the UNDP in 2010.

Our study variable is the weighted average HDI of cities where the network is established. We use as weight the number of stores per town.

Weighted average population of the network location areas

The Brazilian Institute of Geography and Statistics (IBGE), implements for the Brazilian government a population census approximately every 10 years, with the last three censuses conducted in 1991, 2000 and 2010. Out of this period, the IBGE estimates the population per municipality by linearizing the trend indicated by the last two censuses. We used the IBGE estimates for 2011 regarding the population of the cities where the sample chains were established, in 2011 to calculate the weighted average, using the same procedure as with the HDI.

3.2.1.2 Additional variables related to the location strategy

Network age

The age of the network is the difference between the year of the first franchised unit and each year in our panel.

Percentage of company owned-units

This variable represents for each franchising chain the proportion of company-owned units in the total of the chain outlets.

Brand value

The ABF yearbooks provide information about a total of thirteen types of supports that the associated franchisors offer to their franchisees. Parts of them are obligatory, as defined by the Brazilian Law of Franchising (No. 8955/94). The other supports are optional. They are: legal support, choice of plant and equipment, support for financing leverage, promotional support, financing of project, marketing project, organizational design, design/production layout of the new store, and advertising and marketing.

In our dataset, all these types of supports are represented by a dummy variable. We add these dummies to construct a proxy of the brand value, related to the franchisor's effort.

3.2.2. Variables used to study the performance outcome (Econometric estimations)

3.2.2.1. Dependent variable

Network average monthly income per outlet

The performance indicator is the dependent variable of our econometric models. We used the network average monthly income per outlet. This information is available in the ABF Franchising Guide. We collected the information from 2009 to 2014.

3.2.2.2. Core explanatory variable

Location strategy

The network location strategy is the main explanatory variable of our econometric models. We constructed this variable with a two-step cluster analysis (section 4), using the variables previously presented (subsection 3.2.1.).

3.2.2.3. Control variables

Sector dummies

The ABF Guide offers a sector classification. We made the choice to use this official classification of the franchising chains in Brazil to construct sector dummies. The classification is presented in Table 1.

Foreign

We constructed an additional dummy variable to indicate whether chains operating in Brazil have a Brazilian or a foreign headquarter.

ABF label “certification of excellence”

The ABF Seal of Excellence is awarded to distinguish chains achieving a minimum level of satisfaction among its franchisees. This certification is based on a survey of the franchisees, conducted by the ABF. The results are published annually in the Franchising Guide.

Franchising chains that reached the specified satisfaction ratings are identified with the label. It is important to mention that chains have to pay to be part of the contest. Franchisees satisfaction indexes include the overall chain performance (the franchisor quality in coordinating

the whole chain), the brand performance with regard to profitability, and operating performance of the chain (class support to franchisees).

We used a dummy variable to indicate the chain situation during the period 2009-2014.

3.3. Summary statistics

Table 2 reports the means, standard deviation, and correlation matrix. Because the presence of a few missing values may bias the estimations, we completed the data with the average values.

Descriptive statistics indicate that the variables are quite homogenous in the sample, as shown by the means higher than the standard deviations. However, this is not the case with the variables *dispersion* and *percentage of company-owned units*, characterized by some degree of heterogeneity.

Table 1 – Summary statistics and correlations

	Average	SD	1	2	3	4	5	6	7	8	9	10
1. Location Strategy	2.43	0.62	1	,000**	.012	,000**	.000	,000**	.033	,000**	.042	.025
2. Network Dispersion	0.40	2.49	,000**	1	-,080**	,000**	-,100**	-,083**	-,061**	-,096**	-,053*	-,042
3. % of network outlets in Shopping Malls	0.42	0.39	.012	-,080**	1	-,039	,089**	-,052*	,000**	,000**	,000**	,000**
4. Network Age	21.66	15.09	,000**	,000**	-,039	1	-,016	-,046*	.006	-,050*	-,003	,000**
5. % of Company-Owned Units	0.21	0.25	.000	-,100**	,089**	-,016	1	,000**	,000**	,044*	,095**	,000**
6. Brand value	10.78	2.36	,000**	-,083**	-,052*	-,046*	,000**	1	.024	,065**	,052*	.005
7. Mean Population of the network location area	3.29	2.32	.033	-,061**	,000**	.006	,000**	.024	1	,000**	,000**	,000**
8. Mean HDI of the network location area	0.78	0.01	,000**	-,096**	,000**	-,050*	,044*	,065**	,000**	1	,000**	,082**
9. Mean GDP of the network location area	125.01	98.84	.042	,053*	,000**	-,003	,095**	,052*	,000**	,000**	1	,100**
10. Performance	84.76	76.79	.025	-,042	,000**	,000**	,000**	.005	,000**	,082**	,100**	1

* Three variables have a higher scale than the others, and are thus divided by 1000 000: Network Dispersion, Mean GDP of the network location area, Mean Population of the network location area.

4. Behavior categories in spatial strategies

4.1. Methodology

In order to construct the variable *location strategy*, that will be the core explanatory variable in the estimations, we use a two-step cluster analysis. This statistical multivariate technique is indeed developed in two phases. The aim is to achieve the maximum intra-group homogeneity and the biggest inter-group heterogeneity that would otherwise not be apparent.

To carry out this analysis, in the first stage, the individuals are distributed into pre-clusters, which become single individuals in the second stage. This second stage consists in applying a hierarchical algorithm in the pre-clusters. The advantage of this method is that categorical and continuous variables can be included. In addition, the number of clusters is automatically selected and the method allows controlling for the quality of the clustering process.

As previously mentioned, our analysis is based the following variables: the network geographic dispersion, the percentage of network outlets located in shopping malls, the weighted average human development index of the network location area, and the weighted average population of the network location area. In addition with these localized variables, we include complementary variables related to the location strategy: the network age, the percentage of company-owned units, and the chain brand value.

As we use a panel dataset, we consider here each franchised network for a specific year as a different individual the other years. This methodology enables to capture changes in strategy over time.

Finally, we use the variable *network geographic dispersion* as criterion to order the clusters, and to construct on this basis the ordered qualitative variable *location strategy*.

We perform two additional cluster analyses i.e. two additional ways to construct the variable *location strategy*, for robustness checks in the econometric estimations. All the clustering processes present a good quality level, as shown in Appendix 2.

4.2. Data analysis and features of the three behavior categories

The first and main clustering process allows us to distinguish three groups of franchising networks in Brazil. On the basis of descriptive statistics for each group², the following comments can be offered to specify the three behavior categories.

² The whole descriptive statistics are available upon request.

The first group that we label "*concentrated beginners franchisors*" contains the youngest networks (15.9 years), with a low geographic dispersion. These chains are established in cities with a moderate population and HDI. The franchisors in this group have established franchised units almost immediately. These networks are also characterized by a moderate presence in shopping malls. Regarding the contract design, the royalty rate and the upfront fee of these chains are smaller in average compared to the two other groups.

We label the second group "*concentrated mature franchisors*". This second category consists of older networks (27.1 years), with a moderate geographic dispersion. This group is distinguished by a positive and significant relationship between the HDI and the total number of units; the units are located in cities with a high population. In addition, the networks are mainly present in shopping malls. In this group, the presence in shopping malls is positively and significantly related to the royalty rate. The reason of this correlation could be the strong brand name value, compared to the other two groups.

The third group is labeled "*dispersed mature franchisors*". The networks in this group have a high level of geographic dispersion, and they are the oldest (29.78 years) in average in the sample. The age is positively and significantly related to the number of cities where the network is established. In addition, the number of cities where the network is established is significantly and negatively related to the percentage of company-owned units. This correlation is clearly higher compared to the other two groups. This seems to indicate that the expansion process experienced by these chains is strongly driven by franchising.

From this analysis, two main discriminating variables are highlighted: the network geographic dispersion and the network age. As previously mentioned, we make the choice to order the ordered qualitative variable *location strategy*, constructed with the two-step cluster analysis, using the network geographic dispersion. This variable is the main regressor in our econometric models, as presented hereafter.

5. Performance outcome of spatial strategies

5.1. Methodology

Regarding the choice of the econometric model, as we use panel data, we compare first the random effects model and the fixed effects model. Both of them address the problem of the unobserved heterogeneity by specifying an error term constant over time for each unit (fixed effects model) or randomly distributed over time for each unit (random effects model).

With short periods panels, as with the sample, the random effects model may produce better estimators than the fixed effects model (Heckman, 1981). In addition, the random effects model is consistent in the presence of time-invariant variables (Greene, 2000). This is not the case of the fixed effects model. Indeed, time-invariant variables can be perfectly collinear with the fixed effects, while most of the contract variables are by nature almost time-invariant.

The Hausman test confirms our intuition, and shows that the random effects model is more appropriate to the data (χ^2 : 1.44, p-value: 0.487)

Additional checks are performed, which confirm the choice for the random effects model. We use the Lagrange multiplier test to see if the variance across the franchised networks is zero. This test supports the random effect model, since it provides evidence of significant differences across the networks (χ^2 : 3228.64, p-value: 0.000).

Using again the Hausman test, we check for potential problems of endogeneity, since the regressors may indeed raise endogeneity problems, because they are geographical characteristics from where the outlet is located deriving from the franchisor's choice. We compare an instrumental model including as instrument the lagged variable, with the previous results. The results presented in Appendix 4 show that there is no problem of endogeneity.

In addition, we perform a likelihood-ratio test regarding heteroskedasticity at the panel level. The results (χ^2 : 3307.8, p-value: 0.000) confirm that the data in the sample do not have a common disturbance variance, providing thus another support for the random effect model. We also test for autocorrelation. The results (χ^2 : 13.23, p-value: 0.000) show that we have to correct this problem.

In order to correct these different problems we use the generalized least squares method (GLS). This method enables us to estimate the unknown parameters of a linear regression model. The GLS is applied when the variances of the observations are unequal (heteroscedasticity problem), or/and when there is a certain degree of correlation between the observations. It is well known that in both cases the ordinary least squares (OLS) can be statistically inefficient, or even provide misleading inferences.

Finally, we test if all year coefficients are not jointly equal to zero. The results (χ^2 : 215.11, p-value: 0.000) show that the time-fixed effects are needed.

All the results of the specification tests are presented in Appendix 4.

5.2. Estimation results

Our estimation results are reported in Table 3, and concern the influence of the location strategy on franchised chain performances (models 1 and 2).

For robustness checks, we perform estimations with alternative construct of the location strategy³ (models 3-6).

Table 2 – Performance outcome of spatial strategies (1)
Random-effects models for the average monthly income per outlet

	(1) (3 clusters)	(2) (3 clusters)	(3) (7 clusters)	(4) (7 clusters)	(5) (9 clusters)	(6) (9 clusters)
Location Strategy	-1.839* (1.011)	-3.304*** (0.840)	-1.840*** (0.304)	-0.572* (0.298)	-1.173*** (0.205)	-0.724*** (0.205)
Foreign		-26.71*** (3.718)		-26.83*** (3.980)		-25.92*** (3.927)
ABF label		1.547** (0.704)		1.559** (0.701)		1.664** (0.709)
<i>Sector dummies</i>	<i>no</i>	<i>yes</i>	<i>no</i>	<i>yes</i>	<i>no</i>	<i>yes</i>
<i>Year dummies</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>
Constant	77.70*** (2.645)	92.38*** (2.439)	80.22*** (1.232)	84.99*** (1.145)	79.66*** (1.165)	86.20*** (1.186)
N	2010	2010	2010	2010	2010	2010
Wald chi2	124.62***	883.62***	191.77***	954.71***	188.93***	950.88***

Standard errors in brackets

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

We comment first the good global significance of the models, highlighted by the Wald Chi2 tests. In addition, comparing the estimations of the models based on three, seven and nine clusters, the models including or not sector dummies, and the models including or not control variables, we conclude that the results are robust.

The main result of this set of estimations is the clear significant influence of the location strategy on the network performance. As clusters are ranked regarding the *dispersion* variable, the negative sign means that higher network geographic dispersion generates a lower performance. In other words, agglomeration would favor chain performance.

³ Cluster analyses with seven and nine groups, in model (3-4) and (5-6), respectively.

The two control variables have also a significant impact on the network performance. Logically, a positive relationship is highlighted between ABF labeled chains and performance. The headquarter nationality (Brazilian *versus* no-Brazilian) has also an effect on the chain performance. The negative sign indicates here that chains with Brazilian headquarters are more likely to achieve higher performances.

As the *location strategy* variable sums up several strategic decisions of the franchised chain, we estimate additional models including as regressors the variables initially used in the clustering. The aim is to go further in the comments, based on complementary results. The full estimation results of these models (models 7-10) are presented in Appendix 3. We provide a summary of these results in Table 4.

Here again we obtain a good global significance of the estimated models, and the results are robust. The significant and negative sign of the variable *Network geographic dispersion* provides support to the preceding empirical results, indicating that agglomeration favors chain performance.

These estimation results provide additional evidence.

First, the location in shopping center impacts significantly and positively chain performance. Multiplicative variables have no significant impact, suggesting that the interaction of *dispersion* with the network *age* or with the *proportion of company-owned units* does not affect the performance. Consistent with the gravity models, location in area with a high HDI, or a high population, increases chain performance. We obtain a reverse and surprising sign with the GDP (model 7). Concerning the control variables, the sign regarding the variable *foreign* is consistent with preceding results; the network *age* and the *brand value* favor the chain performance, which stands to reason.

Table 3 – Performance outcome of spatial strategies (2) Summary of additional estimation results

Random-effects models for the average monthly income / outlet		
Mean HDI of the network location area	+	***
Mean population of the network location area	+	***
% of network outlets in shopping malls	+	***
Network geographic dispersion	–	***

Conclusion

The literature in economics and management underlines the role of location choice as one of the most important decision regarding retailing. Indeed, location is paramount as it defines the access and attraction of large numbers of customers, impacting thus the chain market-share and profitability.

Studying Brazilian franchising, we focus on spatial strategies of franchising chains in this continental sized emerging country.

In order to highlight the potential determinants of location choices, we first survey previous literature related to location decision and spatial competition between the franchisees of a same chain.

Our analytical framework takes into account the complexity of store location problems, and the different factors potentially relevant. Four determinants of the location strategy are distinguished: the choice for agglomeration of the chain outlets *versus* dispersion in the space, the choice for a specific region or city, related to the socioeconomic features of the area, the choice for a preemption strategy to get a first mover advantage, related to the network age, and the location choice related to the ownership strategy of the chain.

The empirical analysis is conducted on a new, recent, and unique panel dataset compiling information from three complementary sources, and containing geographical data, for the Brazilian case. With a two-step cluster analysis, we compile information related to location decisions to distinguish several spatial strategies of franchising chains in Brazil. These strategies are then submitted to econometric estimations in order to derive performance outcomes.

Our results show that the choice for agglomeration leads to higher chain performances, which is consistent with the Hotelling's model prediction. In addition, our estimation results provide evidence for the gravity model, as the chains located in area with high population and human development index reach higher performances.

References

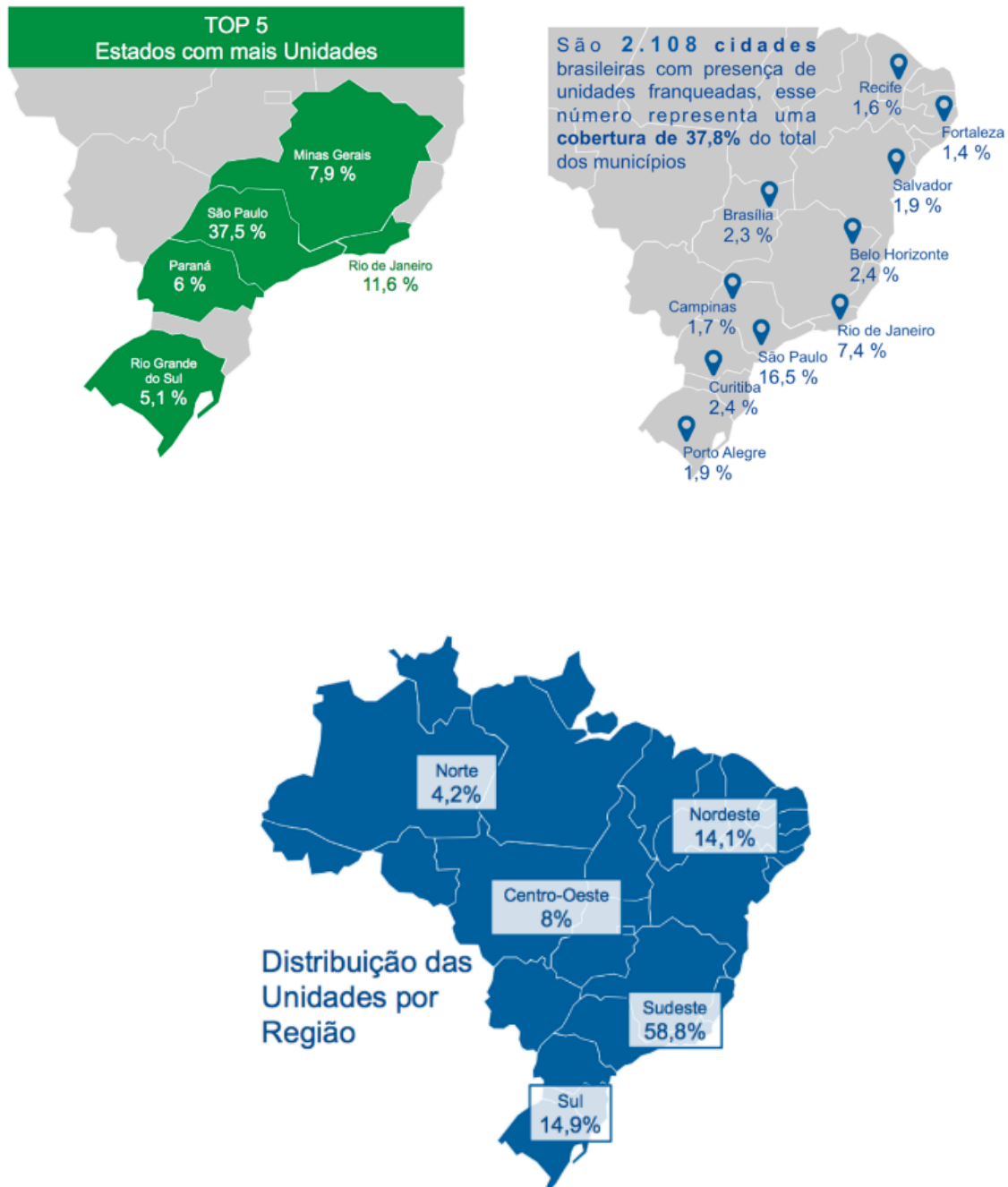
- Applebaum, W., 1966, "Methods for determining store trade areas, market penetration, and potential sales", *Journal of Marketing Research*, Vol. 3 Issue 2, 127-141.
- Azevedo, P., Silva, V. 2005. Governance inseparability in franchising: Evidences from case-studies in France and Brazil, *2nd EMNet Conference*, Budapest, Hungary.
- Bitti E.J.S, Neto J A, Aquino A.C.B, 2010, "Vectors of influence on the contractual mix: geographical dispersion vs. automation in Brazilian franchised chains", *Mimeo*.
- Brickley, JA, Dark, FH. 1987, "The choice of organizational form: the case of franchising", *Journal of Financial Economics*, 18, 401-420.
- Brown, S., 1989, "Retail location theory: the legacy of Harold Hotelling", *Journal of Retailing*, Vol. 65 Issue 4, 450-470.
- Carney, M., Gedajlovic, E., 1991, "Vertical integration in franchise systems: agency theory and resource explanations", *Strategic Management Journal*, Vol. 12, No. 8, 607-629.
- Cataluña F.J.R., García A.N., Díez de Castro, E.C., 2007, "Proposing new variables for the identification of strategic groups in franchising", *International Entrepreneurship and Management Journal*, 3, 355-377.
- Chaudhuri A, Ghosh P, Spell C., 2001, A location based theory of franchising, *Journal of Business and Economic Studies*, 7, 54-67.
- Cheong, T.S., Wu, Y., 2014, "The impacts of structural transformation and industrial upgrading on regional inequality in China", *China Economic Review*, 31, 339-350.
- Christaller, W., 1933, *Die zentralen Orte in Süddeutschland*, Iéna, Fischer.
- Christensen J., Drejer I., 2005, "The Strategic Importance of Location: Location Decisions and the Effects of Firm Location on Innovation and Knowledge Acquisition", *European Planning Studies*, 13, 807-14.
- Cliquet G., 1998, "Integration and territory coverage of the hypermarket industry in France : a relative entropy measure", *International Review of Retail, Distribution and Consumer Research*, Volume 8, Issue 2, 205-224.
- Combs J., Ketchen D. and Hoover V., 2004, A strategic groups approach to the franchising-performance relationship, *Journal of Business Venturing*, 19, 877-897.
- Combs J.G, Ketchen D.J., 1999. Can capital scarcity help agency theory explain franchising? Revisiting the capital scarcity hypothesis, *Academy of Management Journal*, 42, 196-207.
- Combs J.G., Ketchen Jr. D.J, Hoover V.L., 2004, "A strategic groups approach to the franchising-performance relationship", *Journal of Business Venturing*, 877-897.
- Combs, J.G., Ketchen, D.J. Jr., 2003, "Why do firms use franchising as an entrepreneurial strategy? A metaanalysis", *Journal of Management*, 29, 443-365.
- Combs, J.G., Michael S.C., Castrogiovanni G.J., 2004, "Franchising: a review and avenues to greater theoretical diversity", *Journal of Management*, 30, 907-931.
- Craig C.S., Ghosh A., McLafferty S., 1984, "Models of the Retail Location Process: A Review", *Journal of Retailing*, 60, 5-36.
- Dant, R., Perrigot, R., Cliquet G., 2008. "A cross-cultural comparison of the plural forms in franchise networks: United States, France and Brazil", *Journal of Small Business Management*, 46:286-311.
- Ehrmann T., Meiseberg B., 2011, "Inner strength against competitive forces: successful site selection for franchise network expansion" *New Developments in the Theory of Networks*, 93-116.
- Fadairo, M., Lanchimba C., 2013, "Signaling the quality of a business concept; Evidence from a structural equations model with Brazilian data", *40th Annual Conference of the European Association for Research in Industrial Economics*, Évora, Portugal, August-September 30-1.
- Falbe, C., Welsh, D., 1998. Nafta and franchising: A comparison of franchisor perceptions of characteristics associated with franchisee success and failure in Canada, Mexico, and the United States, *Journal of Business Venturing*, 13:151-171.
- Fally, T., Paillacar, R., Terra, C., 2010, "Economic geography and wages in Brazil: Evidence from micro-data", *Journal of Development Economics*, 91, 155-168.
- Ghosh A., McLafferty S., 1982, "Locating stores in uncertain environments: a scenario planning approach", *Journal of Retailing*, 58, 14, 5-22.
- Ghosh A., McLafferty S.L., 1987, *Location strategies for retail and service firms*, Lexington Books.
- Gonzalez-Diaz M., Solis-Rodriguez V., 2015, Differences in contract design between successful and less successful franchises, *European Journal of Law and Economics*, Published online: 01 May.

- Granot, D., Granot, F., Raviv T., 2010, "On competitive sequential location in a network with a decreasing demand intensity", *European Journal of Operational Research*, Vol. 205 Issue 2, 301-312.
- Greene, W., 2000, *Econometric Analysis*, Prentice-Hall International, 3th edn edition.
- Hadfield G., 1991, "Credible spatial preemption through franchising", *RAND Journal of Economics*, Vol. 22, No. 4, pp. 531-543.
- Hausman, J., 1978, "Specification tests in econometrics", *Econometrica*, 46:, 251-1271.
- Hotelling H., 1929, Stability in Competition, *Economic Journal*, 39, 41-57.
- Huff D., Batsell R., 1977, "Delimiting the areal extent of a market area", *Journal of Marketing Research*, 14, 4, 581-585.
- Huff, D.1964, "Defining and estimating a trading area", *Journal of Marketing*, 28,34-38.
- Ingene C.A., Yu E., 1982, "Environment determinants of total per capita retail sales in SMSAs", *Regional Science Perspectives*, 12, 52-61.
- Jones, K., Simmons, J., 1990, The retail environment, Routledge, London.
- Jones, K., Simmons, J., 1987, *Location, Location, Location: Analysing the retail environment*, Methuen, London.
- Journal of Economics, 23,263-283.
- Julian, S.D., Castrogiovanni, G.J., 1995, "Franchisor geographic expansion", *Journal of Small Business Management*, 33, 1-11.
- Kaufmann, P.J., Rangan, V.K., 1990, "System conflict during franchise expansion", *Journal of Retailing*, Vol. 66 Issue 2, 155-173.
- Kelly J.P., Freeman D.C., Emlen J.M., 1993, "Competitive impact model for site selection: the impact of competition, sales generators, and own store cannibalization", *International Review of Retail, Distribution, and Consumer Research*, 3, 237-59.
- Khan M.A., 1999, *Restaurant Franchising*, 2nd ed. John Wiley & Sons, Inc., New York.
- Kumar, V., Karande, K., 2000, "The effect of retail store environment on retailer performance", *Journal of Business Research*, Vol. 49 Issue 2, 167-181.
- Lafontaine F., 1992, "Agency theory and franchising some empirical results", *RAND Journal of Economics*, Vol. 23, No. 2, pp. 263-283.
- Lafontaine F., Slade M., 2001, "Incentive contracting and the franchise decision", *Advances in Business Application of Game Theory*, Kluwer Academic Press.
- Lafontaine, F., Oxley, J., 2004. International franchising practices in Mexico: Do franchisors customize their contracts. *Journal of Economics and Management Strategy*, 13:95-123.
- Lal, R. 1990, "Improving channel coordination through franchising", *Marketing Science*, 9,299-318.
- Lee, K.H., McCracken, M., 1982, "Spatial adjustment of retail activity: a spatial analysis of supermarkets in metropolitan Denver", 1960-1980, *Reg Sci Perspect*, 12, 62-75.
- Lessmann, C., 2014, "Spatial inequality and development ; Is there an inverted-U relationship? ", *Journal of Development Economics*, 106, 35–51.
- Lillis, C.M., Narayana C.L., Gilman J.L., 1976, "Competitive advantage variation over the life cycle of a franchise", *Journal of Marketing*, 40, 77-80.
- Lösch, A., 1954, *The economics of location*, New Haven, Yale University Press.
- Maness, R., 1996, "Incomplete contracts and the choice between vertical integration and franchising", *Journal of Economic Behavior and Organization*, 31, 1, 1001-115.
- Martin, R.E. and Justis, R.T., 1993, "Franchising, liquidity constraints and entry", *Applied Economics*, 25, 9, 1269–77.
- Mathewson, G.F., Winter, R.A. (1983a), "The economics of vertical restraints in distribution", in J.E. Stiglitz, G.F. Mathewson, *New developments in the analysis of market structure*, 211-239.
- Mathewson, G.F., Winter, R.A. (1983b), "Vertical integration by contractual restraints in spatial markets", *Journal of Business*, 56, 4, 497-517.
- Mathewson, G.F., Winter, R.A. (1984), "An economic theory of vertical restraints", *RAND Journal of Economics*, 15, 27-38.
- Mathewson, G.F., Winter, R.A. (1985), "The Economics of Franchise Contracts ", *Journal of Law and Economics*, 28, pp. 503-526.
- Mendes, A.B., Themido, I.H., 2004, "Multi-outlet retail site location assessment", *International Transactions in Operational Research*, Vol. 11, Issue 1, 1-18.
- Meyer, A.D., Tsui, A.S., Hinings, C.R. 1993, Configurational approaches to organizational analysis, *Academy of Management Journal*, 36, 1175-1195.

- Michael, S., 2014, Can franchising be an economic development strategy for developing nations? An empirical investigation. *Small Business Economics*, 42, 611–620.
- Michael, S.C., 1999, “The Elasticity of Franchising”, *Small Business Economics*, 12, 313 – 320.
- Minkler, AP., 1990, “An empirical analysis of a firm’s decision to franchise”, *Economic Letters*, 34, 77–82.
- Norton, SW. 1988, “Franchising, brand name capital, and the entrepreneurial capacity problem”, *Strategic Management Journal*, Summer Special Issue 9, 105–114.
- Pancras J., Sriram S., Kumar, V., 2012, “Empirical investigation of retail expansion and cannibalization in a dynamic environment”, *Management Science*, Vol. 58 Issue 11, 2001-2012.
- Park K., Khan M., 2006, “An exploratory study to identify the site selection factors for U.S. franchise restaurants”, *Journal of Food service Business Research*, 8, 97-114.
- Perryman, A.A., Combs J.G., 2012, “Who should own it? An agency-based explanation for multi-outlet ownership and co-location in plural form franchising”, *Strategic Management Journal*, 33, 368–386.
- Peterson K., 2003, “Enterprise-wide analytical solutions for distribution planning”, *Database Marketing & Customer Strategy Management*, 11, 13-25.
- Porter M.E., 1980. *Competitive Strategy*. Free Press, New York.
- Reilly, W., 1931, *The Law of Retail Gravitation*, Putham, New York.
- Resende Melo, P.L., Felipe Mendes Borini, F., Miranda Oliveira M., Couto Parente, R., 2015, “Internationalization of Brazilian franchise chains: a comparative study”, *Revista de Administração de Empresas*, São Paulo, 55, 3, 258-272.
- Rey, P., Tirole, J., 1986a, “The logic of vertical restraints “, *American Economic Review*, 76, 921-939.
- Rey, P., Tirole, J., 1986b, “Vertical restraints from a principal-agent viewpoint “, in L. Pellegrini, S.K. Reddy, *Marketing Channels, Relationships and Performance*, Lexington books, 3-30.
- Silva,V., Corvacho, T., Franco, N., Sanches, I., 2009. “Franchising and the acceleration of the food service sector in Brazil: Opportunities for the food industry?”, *EMNet Conference*, Sarajevo, Bosnia and Herzegovina, September 3-5.
- Simons R.A., 1992, “Site attributes in retail leasing: an analysis of a fast-food restaurant market”, *The Appraisal Journal*, 60, 4, 521-531.
- Streed, O., Cliquet, G., 2013. “International market entry expansion of retail networks: Determinants of market entry failures”, *Network Governance: Alliances, Cooperatives and Franchise Chains*, pages 87-106. Springer Verlag.
- Vance, P., Madeira, A., Brashear, T., 2011, “Internationalization of Brazilian franchisors: profiles and strategies”, *25th Annual International Society of Franchising Conference*, Boston, June 16-18.
- Windsperger J., 2004, “Centralization of franchising networks: Evidence from the Austrian franchise sector”, *Journal of Business Research*, 57, 1361-69.

Appendix 1. Brazilian franchising around the country

Source: Brazilian Franchising Association (2015)



Appendix 2. Summary of clustering analyzes

Variables	<i>Clustering 1</i> (Main clustering)		<i>Clustering 2</i> (Control clustering)		<i>Clustering 3</i> (Control clustering)	
Dispersion	Yes		Yes		Yes	
Mean population of the network location areas	No		No		Yes	
Percentage of network outlets in shopping malls	Yes		No		Yes	
Mean HDI of the network location areas	Yes		Yes		Yes	
Network age	Yes		Yes		Yes	
Percentage of company owned-unit	Yes		No		Yes	
Brand value	Yes		Yes		Yes	
	Clusters	Chains	Clusters	Chains	Clusters	Chains
Number of clusters & franchising chains	1	1020	1	229	1	176
	2	852	2	388	2	155
	3	138	3	506	3	229
			4	412	4	318
			5	251	5	371
			6	200	6	397
			7	24	7	151
					8	175
					9	38
Quality of the clustering						

Appendix 3. Additional results for the performance outcome of spatial strategies

Random effects models for the average monthly income per outlet

	(7)	(8)	(9)	(10)
Network Geo Dispersion	-0.321* (0.269)	-0.450* (0.277)	-0.141* (0.0846)	-0.078** (0.00349)
% of network outlets in shopping malls	8.096*** (2.296)	7.634*** (1.955)		
Dispersion x Age			-0.0170 (0.221)	
Dispersion x Owned units				0.194 (0.124)
% of company-owned units				-2.075 (1.742)
Mean GDP of the network location areas	-0.276*** (0.00492)			
Mean HDI of the network location areas		241.2*** (39.79)	190.6*** (36.00)	175.8*** (49.76)
Mean population of the network location areas	0.139*** (0.0210)	0.0172*** (0.00301)	0.0259*** (0.00304)	0.0268*** (0.00451)
Network age	0.448*** (0.0551)	0.465*** (0.0517)	0.406*** (0.0471)	0.443*** (0.0461)
Brand value	0.563** (0.238)	0.645** (0.215)	0.279 (0.232)	0.519** (0.236)
Foreign	-26.24*** (4.196)	-28.63*** (4.319)	-24.63*** (3.776)	-26.50*** (4.107)
<i>Sector dummies</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>
<i>Year dummies</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>
Constant	48.43*** (3.631)	-127.9*** (31.44)	81.59*** (1.051)	83.73*** (0.925)
<i>Number of networks</i>	335	335	335	335
<i>Wald chi2</i>	712.49***	877.87***	790.55***	996.18***

Standard errors in brackets

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Appendix 4. Specification tests

(Econometric models 1-10)

Models	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
<i>Hausman test</i>										
χ^2	1.44	9.81	0.23	11.55	0.76	1.02	7.19	0.00	0.01	0.01
(p-value)	0.487	0.1995	0.998	0.1164	0.685	0.599	0.3033	0.9467	0.9306	0.9308
<i>Lagrange multiplier test</i>										
χ^2	3228.64	3149.50	3227.81	3137.45	3221.05	3127.55	3182.14	3172.48	3006.71	3189.98
(p-value)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
<i>Likelihood-ratio test for heteroskedasticity</i>										
χ^2	3307.8	4787.50	3307.85	4781.52	438.44	3536.06	3707	3724.46	4686.16	4786.44
(p-value)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
<i>Autocorrelation test</i>										
χ^2	13.23	13.231	13.052	13.056	13.25	13.251	11.531	11.530	12.844	11.633
(p-value)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
<i>Test to include fixed effects</i>										
<i>Year dummies</i>	215.11	158.74	223.50	213.37	224.04	207.26	110.22	88.78	144.66	85.15
	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
<i>Sector dummies</i>	606.22	763.24	493.58	620.25	532.31	668.14	217.85	169.38	450.31	344.88
	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Tested variable ⁴	<i>Test for Endogeneity (Hausman test)⁵</i>									
Location	0.33	0.32	3.11	4.15	4.09	3.40				
Strategy	1.00	1.00	0.9597	0.9652	0.943	0.992				
Dispersion							0.00	0.00	0.00	0.00
							1.00	1.00	1.00	1.00
% of company-owned units										
Network age							0.21	0.21		0.00
							1.00	1.00		1.00
Dispersion x Owned units									0.00	

⁴ Regarding the two variables “Foreign” and “ABF label”, the instrumental variables methodology cannot be used. Indeed, our dataset does not provide relevant instruments for these two variables. However, removing them from the models (step-by-step approach), we observe that the estimation results remain qualitatively similar. For this reason, we conclude that there is no endogeneity problem with these variables.

⁵ We perform Hausman tests to compare the random effects model with a model including instrumental variables for the suspected endogenous variables.

			1.00	
Dispersion x Age				0.00
				1.00
% of network outlets in shopping centre	0.00	0.00		
	1.00	1.00		
Mean GDP of the network location areas		1.51		
		1.00		
Mean population of the network location areas		0.00	0.00	0.00
		1.00	1.00	1.00
Mean HDI of the network location areas	0.00		0.00	0.00
	1.00		1.00	1.00
Brand value	6.73	0.02	0.00	0.00
	0.965	1.00	1.00	1.00
