URBAN EMPLOYMENT SUBCENTERS: THE CASE OF CURITIBA, BRAZIL

Eduardo Moreira Thomé [1]
Alexandre Alves Porsse [2]

ABSTRACT: Several studies have been developed to assess the presence of polycentric cities analyzing the spatial distribution of employment into urban space. Most of these studies are concerned to cites in developed countries and few have explored the polycentric urban structure of cities in developing countries, particularly to Latin America. Furthermore, the use of urban planning policies and their influence in promoting subcenters has been left aside in empirical work. In that sense, the city of Curitiba (Brazil) emerges as focal point in this analysis. Since the 1970's Curitiba adopted a TOD oriented urban planning policy searching to promote polycentricity. The core of this policy is the Bus Rapid Transit (BRT) system in combination with land-use patterns which were established over the last decades and are the main force influencing the distribution of people and economic activities across the city. This paper aims to assess the presence of employment subcenters in Curitiba and how it intertwines to the BRT system. In order to identify the employment subcenters we use the semi-parametric approach developed by McMillen (2001) and to assess whether the CBD remains as the main employment attractor in a polycentric city. The employment database contains data on formal employment at the firm level provided by the Ministry of Employment for the year of 2010, representing 50,553 firms and 658.061employees in the city. These data were georeferenced in grid cells of 1 square kilometer over the city. The results show that Curitiba has a polycentric structure composed by nine employment subcenters. The spatial distribution of these subcenters is highly correlated to the BRT network, evidencing the role of TOD urban planning policy in shaping the polycentric structure.

JEL CLASSIFICATION: C14, R12, R30.
KEY WORDS: Semi-parametric regression, Urban Employment Density, Subcenters, Polycentrism.

[1] Student of the Graduate Program of Economic Development at Federal University of Paraná.
[2] Professor of the Department of Economics at Federal University of Paraná.

1. INTRODUCTION

Initial models of economic urban theory developed by Alonso (1964), Mills (1967) and Muth (1969) hypothesized a concentric city that concentrated employment in the Central Business District (CBD). In recent years the urban spatial structure of large cities has seen a trend away the monocentric model.

The reconfiguration process of employment location in a city has been verified through the analysis of employment location and land use patterns. Both point towards a relocation and concentration of employment in different centers other than the CBD. Those secondary centers are known as Subcenter Business District (SBD).

Polycentricity has been empirically verified in several studies in major North American cities (Giuliano and Small, 1991; McMillen, 2001; Yang *et al*, 2019) and across Europe (Veneri, 2013; Krehl, 2016; Mulicek and Maly, 2019). Recent trend has explored the urban configuration in developing countries, predominantly on Chinese municipalities (Yue *et al*, 2010; Lv *et al*, 2017; Xie *et al*, 2019). Nevertheless, little has been explored in other developing countries, particularly in Latin America (Fernandez-Maldonado *et al*, 2014; Belmiro, Rodrigues and Neto, 2016; Campos and Chagas, 2017).

Further, the use of urban planning policies and their influence in promoting subcenters has been left aside in empirical work. In that sense, the city of Curitiba (Brazil) emerges as focal point in this analysis. The Transit Oriented Development (TOD) policies developed since the 1970's that implemented the Bus Rapid Transit (BRT) system in combination with land-use patterns reshaped the city's landscape.

Also, within the country's economy, Curitiba plays an important role. The city is ranked as the fifth municipality that most contributes to the country's GDP with 1.40% overall[3]. Further, is the eighth largest city in Brazil, with 1.917.185 million residents[4].

The primary focus of the present study is to assess whether the city of Curitiba has employment subcenters. Also, to which extent do subcenters influence in the attraction of employment in comparison to the CBD. Further, the TOD policies adopted since the 70's that reshaped the city structure in recent years might be related to those subcenters, locating in close proximity to the main BRT lines. Further, theoretical work acknowledges the commute patterns of employees in accordance to the location of their residences and work. Ultimately, it will be explored the commuting flows of employees to the CBD and the employment subcenters. At this stage it will be investigated to which extent the SBDs are employment attractors within the other municipalities of the RIT.

It also must be stressed that other studies focus in the analysis of polycentric urban structures within a metropolitan region, which accounts for a larger area of analysis (Giuliano and Small, 1991; McMillen, 2001; Krehl, 2016). The greater area of investigation might absorb a municipality's CBD as an employment subcenter when compared to the region's most prominent employment concentration. Further, Curitiba is the main employment municipality of the region, attracting workers of the nearby cities. As it was pointed in the aforementioned paragraph, our focus is in Curitiba and the relationships between the core of the BRT network system and the city's employment subcenters.

The present work contributes to the empirical research of polycentric structures in the Latin American context while examining a major Brazilian city. It also adds to the literature over estimation procedures in order to assess employment subcenters. Further,

[3] According to the Instituto Brasileiro de Geografia e Estatística (IBGE, 2018a) the first four municipalities are: São Paulo (10.85%), Rio de Janeiro (5.85%), Brasília (3.60%) and Belo Horizonte (1,46%).

[4] According to IBGE (2018b) in the 2010 census the first seven cities ranked according to population are: São Paulo (12.176.866), Rio de Janeiro (6.688.927), Brasilia (2.974.703), Salvador (2.857.329), Fortaleza (2.643.247), Belo Horizonte (2.501.576), Manaus (2.145.444).

the analysis of Curitiba, which has a strong urban planning culture, sheds light into the relationship of TOD and polycentricity.

Including this brief introduction, this paper is organized into six sections. The second section explores the city of Curitiba and its TOD policies. The third section is a literature review followed by the identification strategy. The fourth, explores the estimation procedure. The fifth, displays the results and discussion. Lastly, the conclusion.

2. CURITIBA AND THE TRANSIT ORIENTED DEVELOPMENT

This rapid growth spurt experienced by Curitiba between 1950 and 1980 that doubled its population influenced the proposal of a new master plan in 1966. This new urban policy was a result of the problems that were already being faced by the populational growth. This plan focused on public transportation, road system and mixed land use. The main goal was to integrate mixed land use patterns to the road system and the public transportation system[5].

According to Rabinovitch (1996, p. 53) the key principles involving the *Master Plan* were: decongestion of the central area; infrastructure development; economic support to urban development; demographic control and management; imposing a linear urban growth trend. Furthermore, the plan had in view the creation of a specific neighborhood to accommodate industrial venues, the Industrial City of Curitiba (henceforth, CIC). Located approximately ten kilometers from the CBD[6].

Within the core of the plan is the zoning regulations that intertwine between the transit network. As shown by Shertzer *et al* (2018) and Twinam (2018), land use regulations have a stronger impact in determining the location of commercial and industrial activities in the urban space rather than the transportation network. In the case of Curitiba, the zoning restrictions are strongly related to the public transportation system. Therefore, it is expected that there is a relationship between the distribution of employment in the city and it's TOD policy.

According to Dou *et al* (2016) a Transit-Oriented Development (TOD) must include three dimensions: density, diversity and design. Curitiba's urban development contemplates those three aspects.

The transportation system was initially conceptualized in four main axes, north-south and east-west. A fifth axis was later incorporated in the southeast direction. These five axes are the main core of the Bus Rapid Transit (BRT) network. The lack of available resources and investments faced at the time hindered the construction of a subway system. The solution was to build a similar surface transport system, which culminated in the BRT network.

Another key point of the policy is the trinary road system developed around each axe. The main road consists of segregated bus lanes traveling in opposite directions followed by slow speed traffic lanes and parallel parking. One block away from the main road in are high speed car lanes, each road travelling on different ways.

The five main axes were also envisioned to accommodate populational growth through mixed land-use policy while also promoting high-density populational areas.

[5] Curitiba has had international recognition for the Bus Rapit Transit (BRT) and the efforts around its development since the 1970s when the first bus lines started to operate and the incentives towards mixed land-use pattern alongside it. For further references see Cervero and Dai, 2014; Duarte and Ultramari, 2012; and Khayesi and Amekudzi, 2011.

[6] The previous Master Plan developed in 1941 and 1943 differed in the transportation strategy when compared to the 1966 Master Plan, nevertheless the former plan also had contemplated the development of employment subcenters. See Macedo (2004).

Within the main road it was authorized to build high-rise constructions. In the first two floors of those buildings it is encouraged their use in economic activities. Further, the height of the buildings decay within distance from the main road.

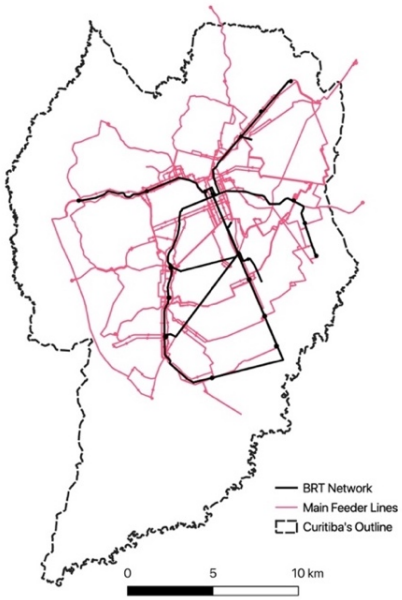
The high populational concentration within the main axes also contributes to make the BRT network self-financing, once it's expected that the residents living nearby will chose the public transportation as their preferred modal[7].

Furthermore, another key issue was the road hierarchy promoted by the BRT network. The arterial roads were the five main axis and core of the system. Followed by the priority roads whose purpose was to link the streets of the arterial axis. Finally, the collector streets accounted for general purposes and to connect the structural roads to neighborhoods without reaching the CBD (RABINOVITCH, 1996).

The design around the transportation system also plays a fundamental role. In the areas that were already developed the plan redesigned the surrounding locations in order to enable the full potential of the area[8]. The undeveloped areas were envisioned to accommodate the expansion of the BRT network and populational growth taking into consideration the design around neighboring areas [9]. The design around the main road is key to the success of the mixed-land use and expansion of the CBD economic activities to the remainder of the city through the transport axes.

Throughout the years the *Master Plan* suffered several modifications, but the plan's core established in 1966 still serves as the main guideline in the urban planning and development of Curitiba. Therefore, the 1966 *Master Plan* was not only conceived to promote an efficient public system transportation. It is a plan that has been implemented over the last fifty years that redesigned the shape of the city through land-use policy.

Figure 1 – BRT Network and Main Feeder Lines



[7] The work developed by Duarte and Ultramari (2012) show evidence of the contrary.

[8] See Rabinovitch (1996) for more details.

[9] Cervero and Dai (2014, p. 132) cite the problems faced in the implementation of the BRT system in Bogotá (Colombia) where "aligning corridors in mostly economically stagnant zones that were largely built out has suppressed land development". They also point that "siting BRT in busy roadway medians, which limited land supplies for leveraging TOD and resulted in mostly unattractive pedestrian environments immediate to stations". Those problems were not faced in Curitiba once the policy was developed to accommodate future populational expansion.

Nowadays, the BRT network expanded to other municipalities around Curitiba. The adjacent cities embraced by the BRT including Curitiba are known as the Integrated Transport Region (henceforth, RIT). As it is exposed in Table 1, Curitiba concentrates most of the formal employment and economically active population in the RIT.

Moreover, Curitiba is the main economic driving force within the region, as it can be seen by the difference between the number of formal employees and the economically active population in the other municipalities of the RIT. Therefore, it should be expected that employees that reside in neighboring locations commute to Curitiba to work.

3. FROM THE MONOCENTRIC TO POLYCENTRIC URBAN PATTERN

The Alonso-Mills-Muth Model analyses a concentric city where a predetermined single location concentrates the firms, the Central Business District (CBD). Notwithstanding, Alonso (1964), Mills (1967) and Muth (1969) acknowledged the possibility that cities could incorporate other concentrations of employment within its limits besides the CBD[10].

A theoretical microeconomic model proposed by Fujita and Ogawa (1982) argue in favor of different urban pattern equilibria that comprises a polycentric city. Further, the literature review by Berry and Kim (1993) and Anas, Arnot and Small (1998)[11] sheds light to the polycentric urban pattern. Both researches display several works that corroborate that polycentricism is not a phenomenon, rather a recurring urban pattern in major American cities.

Table 1 - Total Employment and Economically Employed Population Comprising the RIT in 2010

Municipality	Employment (formal)		Economically Active Population	
	Total	%	Total	%
Bocaiuva Do Sul	1,509	0.13	9,134	0.35
Campo Magro	3,688	0.32	20,852	0.81
Rio Branco do Sul	4,559	0.39	25,480	0.99
Mandirituba	5,137	0.44	18,519	0.72
Campina Grande do Sul	7,547	0.65	32,401	1.26
Piraquara	7,664	0.66	77,085	2.99
Quatro Barras	8,848	0.77	16,754	0.65
Fazenda Rio Grande	9,671	0.84	67,119	2.60
Almirante Tamandaré	11,272	0.98	85,510	3.31
Campo Largo	25,275	2.19	96,196	3.73
Colombo	37,529	3.25	178,212	6.91
Pinhais	41,461	3.59	99,181	3.84
Araucária	51,802	4.48	100,428	3.89
São José dos Pinhais	90,277	7.82	222,039	8.60
Curitiba	848,850	73.49	1,531,838	59.36
Total	1,155,089	100	2,580,748	100

Source: RAIS 2010; IBGE, 2019.

[10] Moses and Golstein (1973) indicate the modification that Muth (1969) made to his model in order to incorporate smaller centers of employment located in the residential area. However, they are considered to small and have no effect on prices or land area occupied to exert any influence over space. Likewise, Mills (1967) cites an example of shopping centers and their possibility to exhibit increasing returns to scale sufficient enough to hinder housing from an area. Although, this force is not enough to be defined as a central location. But considering this possibility in the model would complicate its solution.

[11] The work of Anas, Arnot and Small (1998) comprises a section describing the historic conditions of the urban structure in the United States.

The work empirical research of Shukla and Wadell (1991) and Wadell and Shukla (1993) focus on the analysis of land-use patterns to assess the presence of employment subcenters within the urban space for the Dallas-Fort Worth region. Even so, they were able to determine that the region has several subcenters.

Several studies developed a methodological approach to determine employment subcenters within the urban context. Their focus shifted towards the analysis of employment rather than land-use patterns in order to determine SBDs.

The study of McDonald (1987) proposes an empirical method to identify employment subcenters. In order to address the issue, he determines that local peaks in either gross employment density or employment-population ratio can determine subcenters. Simplifying, a site can only be considered as an employment subcenter if the employment-population ratio or gross employment density doesn't decline the further a location is to the CBD.

Further, gross employment density and the employment-population ratio are calculated considering the employment in manufacturing and the total employment in each location. The local peaks occur when the indicator for a particular site exceeds the value of the contiguous observations on at least one of the two categories.

The proposed identification of subcenters by McDonald (1987) has several issues. In the case of the gross densities his calculations are subject to tract sensitivity once the total areas of each zone and the contiguous zones, which the densities are compared to, might vary. Also, if the choice of contiguity zones is modified, the results might differ as well. Further, his procedure is not subject to statistical testing in order to assess the subcenters significance.

The procedure of Giuliano and Small (1991) to determine subcenter candidates is quite straightforward and replicable in other study regions. They define an employment center as a set of contiguous zones that share a common boundary of least a quarter of a mile. Each zone has to have a minimum total gross employment density of ten employees per acre. Also, all the zones combined to form the subcenter must have at least 10.000 employees. The zones that do not meet the minimum density per acre or the minimum number of total employees are not classified as an employment center.

Further, gross employment density and the employment-population ratio are calculated considering the employment in manufacturing and the total employment in each location. The local peaks occur when the indicator for a particular site exceeds the value of the contiguous observations on at least one of the two categories.

The proposed identification of subcenters by McDonald (1987) has several issues. In the case of the gross densities his calculations are subject to tract sensitivity once the total areas of each zone and the contiguous zones, which the densities are compared to, might vary. Also, if the choice of contiguity zones is modified, the results might differ as well. Further, his procedure is not subject to statistical testing in order to assess the subcenters significance.

The procedure of Giuliano and Small (1991) to determine subcenter candidates is quite straightforward and replicable in other study regions. They define an employment center as a set of contiguous zones that share a common boundary of least a quarter of a mile. Each zone has to have a minimum total gross employment density of ten employees per acre. Also, all the zones combined to form the subcenter must have at least 10.000 employees. The zones that do not meet the minimum density per acre or the minimum number of total employees are not classified as an employment center.

Nevertheless, their choice for the cutoffs is discretionary. As it is mentioned in their study, the total employment in a cluster is reduced in order to keep the CBD of other

three counties significant. Further, the study of Cervero and Wu (1998) also use density cutoffs of seven workers per acre and 9.500 workers in contiguous tracts for the San Francisco Bay Area. According to McMillen (2001) and Redfearn (2007), the arbitrariness involved in the choice of values for the thresholds might be due to previous knowledge of the area studied.

Another problem that arises from the proposed method is that the further a site locates from the CBD the lower its employment density. Thus, the minimum density per acre and total employment per center are difficult to achieve. This difficulty is further enhanced by their use of census tracts. Therefore, their methodology fails to detect employment concentrations in distant areas of the CBD and those sites that represent a local peak in comparison to the surrounding area. Taking into consideration that most of the employment in an urban area is not concentrated, rather dispersed, choosing such criteria favors the identification of SBDs closer to the CBD. The comparison made by Redfearn (2007) between results of the procedures of Giuliano and Small (1991) and McMillen (2001) confirms the previous affirmation.

The work developed by Craig and Ng (2001) tries to remove the high degree of subjectivity pertained in the studies of McDonald (1987) and Giuliano and Small (1991). They use a non-parametric employment density quantile splines regression.

The focus of their procedure relies on the analysis on the probability distribution of the observations on the top quintile of the density function. According to the authors this is an advantage since they aim at the median and not the mean of the distribution. Another gain by their procedure is the use of a non-parametric approach when compared to the linear estimation of parametric models that would possibly disregard subcenters.

The employment density function estimates the logarithm of the employment density of a census tract conditioned to its distance from the CBD. Since the distance to the CBD is being used as the explanatory variable this yields concentric circles around the CBD. Those rings that present an employment gradient greater than what would be expected[12] have a subcenter. Further, they choose the sites with the highest employment[13] density within that ring as the SBD.

The procedure proposed by Craig and Ng (2001) allows for statistical testing and since it is a non-parametric approach it captures the high employment density areas in comparison to their neighbors. Allowing the parameters to vary also takes into account possible geographical restrictions. Nevertheless, their method can only assess rings with high employment density gradients, it cannot determine specific sites as subcenter candidates. In order to so, within those rings that have gradient peaks they choose the census tracts with the highest employment density.

The methodology proposed by McMillen (2001) tries to identify subcenters that can be applied to different cities without the use of threshold values and prior knowledge of the specificities of the city. His two-stage semi-parametric procedure identifies statistically significant sharp local rises in the employment density while conditioning to the distance from the CBD. His methodology is also suited for different units of analysis and takes into consideration the possibility that the distance from the CBD may influence variations in the employment density.

[12] According to Craig and Ng (2001, p. 102) the ring that present an "employment gradient greater than what would be expected from a smoothly declining density quantile spline and use the variation in the gradient to define the location of an employment center".

[13] Craig and Ng (2001, p.102) mention that "using the upper quantile spline, however, greatly increases the chances the chosen point is one that influences the top quantile spline, and so is consistent with finding an employment concentration".

The methodology proposed by Redfearn (2007) determines employment subcenters through nonparametric regression. His main goal was to determine employment concentrations that are significantly higher than the neighboring areas. In his work he also considers the extent of subcenters.

While analyzing the census tracts of the greater Los Angeles area he aggregates them in small subsamples that do not alter the employment density structure of the region. In the first step, he estimates two Locally Weighted Regressions (LWR) for employment density using a 24% window size and a tricube kernel weight[14] considering as the independent variables the North-South and East-West directions.

The highest points in the first step are considered to be candidate subcenters. In the second step he analyzes the tracts that are inside the local maxima previously estimated. Redfearn (2007) considers that local employment density around each maximum to determine the contour lines of the candidate subcenter in a way to approximate the resemblance of the shape of the area estimated in the first step. They are then analyzed through an iterative selection of a group of partitions of neighboring tracts located around the local maximum.

For each maxima another LWR is estimated and those tracts that are considered to be part of the subcenter receive the values estimated by the LWR and, those outside, the average value of the employment density in the area. The partitions that are considered to be the best fit are those that have the lowest sum of squared residuals. Further, for each partition there are other partitions within that minimize the average between the Akaike and Schwartz Information Criterion. The test for significance for the subcenters is taken through a bootstrapped difference-in means test.

Redfearn (2007) compares the results in his model with the ones of McMillen (2001) for the greater Los Angeles area. He states that McMillen's (2001) methodology finds employment subcenters of any given significance throughout the entire region the results of McMillen's (2001) estimations leads to the most significant employment subcenters in the highly urbanized areas of Los Angeles.

Nevertheless, Redfearn's (2007) methodology has different discretionary choices, as in the aggregation of tracts prior to the first step. Further, in the second step the choice of employment density to determine the contour lines of the subcenters estimated in the first step are also discretionary. Lastly, other problems in the estimations such as multicollinearity are not mentioned to be solved for.

The works of McMillen (2001) and Redfearn (2007) sheds light in the problems faced while using density cutoff values and the traditional parametric negative exponential estimation in order to identify employment subcenter. Both agree in the sense the model specification should be nonparametric, accounting for variations in the urban area that influence employment location. Thus, the literature searching to identify employment subcenters has reached a consensus in the model specification, the use of nonparametric models.

Therefore, the estimation procedure proposed by McMillen (2001) appears to be the most suited to assess subcenters[15] within the highly urbanized area of Curitiba. The proposed methodology is readily reproductible and also provides statistical criteria to determine that an SBD is a site that has a significant impact in the estimated employment density. Further, his procedure overcomes *ad hoc* definitions to estimate subcenters.

---

[14] The tricube weight is given by: $(1 - (dist_{ij}/dist_i^{max})^3)^3$. Where $dist_{ij}$ is the distance from the point of interest *i* to tract *j* located within the window; $dist_i^{max}$ represents the distance largest distance from point *i* to any point witihin the window.

[15] Other studies employ McMillen's (2001 procedure, see Krehl (2016) and Lv *et al* (2017).

4. DATA AND ESTIMATION PROCEDURE

In the present study, the presence of employment subcenter within Curitiba's urban space will be determined through the two-stage semi-parametric approach developed by McMillen (2001).

Considering that the analysis in this study relies on spatial distribution of the employment density, first is necessary to georeference the employment through the city. Using the data provided for the year of 2010 by the Brazilian Ministry of Employment (MTE) which conducts an annual report at the firm[16] level, known as RAIS, gathering socioeconomic[17] data of their formal employees[18].

The data was narrowed to the number of formal employees registered in December 31st and the firms' ZIP codes, which enables the georefencing of the employees to determine employment density. The latter was possible through the streets network shapefile provided by IPPUC (*Instituto de Pesquisa e Planejamento Urbano de Curitiba*) that contains the ZIP codes of all streets in Curitiba.

The initial data provided 138.094 firms. After the removal of companies that had no formal employees, that closed throughout the year and those that provided human resources to other firms[19] it remained 50.497 firms. Further, public employees were not taken into consideration. Nevertheless, formal employees at public offices were considered. The initial data at the firm level provided 658.937 formal employees, given the aforementioned considerations, 586.021 employees were georeferenced.

The georeferencing of the data was enabled by the street-level data of Curitiba provided it had the ZIP codes of each street. The firm-level data also had the ZIP codes of each companies located in the city. The merge between the two datasets allowed the georeferencing of the firms.

The spatial location of each firm allowed the georeferencing of the employees since key variables facilitated the merge of the firm- and employee-level data, thus matching each individual to their workplace.

The firm's employment was georeferenced using grid cells of one kilometer by one kilometer[20]. This removes possible tract sensitivity when using the census tracts as McDonald (1987), Giuliano and Small (1991) and McMillen (2001). The study of Krehl (2016) analyzes employment subcenters through McMillen's (2001) procedure also employing grid cells. The methodology proposed in McMillen (2001) allows for different units of analysis and the use of grid cells should not cause any impact in future results. Therefore, the employment density is the number of employees per square kilometers.

The first stage of the procedure developed by McMillen (2001) serves as benchmark for the number of subcenter candidates. In this step, the procedure uses a non-parametric Locally Weighted Regression (LWR) to smooth the natural logarithm of

[16] As mentioned, the data is collected at the firm level. In the case of a firm with multiple production plants, franchises or offices that are in a different address or facility the data is recorded separately even though they belong to the same company.

[17] RAIS collects data for the following information: number of employees in December 31st; employees admitted and dismissed according to their gender, years of education, time employed, their wages and nationality; the data is also available according to the business occupation and sector; the firms' data is also disaggregated according to country region, state and municipality (IBGE, 2019).

[18] According to the methodology proposed in RAIS formal employees are considered those that have their job information in their employment/work card, public employees, temporary workers with a predetermined contract expiration and single workers contracted through syndicates (IBGE, 2019).

[19] The firms classified in the ISIC (2008) at the three-digit level as 783 descripted as "Other human resources provision" were removed provided that the employees provide services to other firms. Therefore, their workplace is not at the provided address by the firm.

[20] Using a grid cell of $1km^2$ is not expected to influence the results of the estimation since the city of Curitiba has a north-south extension of 35km and east-west extension of 20 km (PMC, 2019).

employment density (y_i) over the distance to the CBD ($DCBD_i$). The use of the LWR contributes to identify rises in the employment density function considering only observations that are close to the unit of analysis. Also, it accounts for the difference in the employment density gradients through the area studied.

This initial smooth is also efficient to remove possible sensitivity of the results caused by tract size. The method developed uses a large window size, the nearest 50% of observations of a given site *i* are attributed a weight. The further an observation is from the site *i*, smaller its weight. This large window size leads to a smoother employment density estimate, allowing to identify the sites that have significantly positive residuals, that is, outliers. These sites are possible subcenter candidates.

In this step, McMillen (2001) uses the distance between the centroid of grid cell *i* to the CBD[21] as the explanatory variable. The LWR regression in the first step is:

$$y_i = g(DCDB_i) \quad (1)$$

The subcenter candidates are sites that exceed the smoothed employment with residuals that are significantly greater than zero at a five percent significance level[22]. In order to avoid multicollinearity when several sites are in close proximity, McMillen (2001) narrows the candidates to those that have the highest residuals within a three-mile radius, approximately five kilometers. As the author mentions, this radius is arbitrary, in the present work it will be used a three-kilometer radius[23].

Until this moment it has been identified possible subcenter candidates through local rises in the employment density function. However, this first step does not identify if those sites have a statistically significant effect over the outline of the employment density function. The second step uses a semiparametric regression to determine the significance of the subcenter candidates obtained in the first step, as it follows:

$$y_i = g(DCBD_i) + \sum_{j=1}^{S}\left(\delta_{1j} D_{ij}^{-1} + \delta_{2j} D_{ij}\right) + u_i \quad (2)$$

The third term in equation (2) estimates the distance between a subcenter candidate site *i* and a site *j* in level (D_{ij}) and inverse (D_{ij}^{-1}) form. The distance in level form displays those subcenters that have an effect over the region analyzed, while the inverse form displays subcenters that have a local impact in the employment density function. Hence, if δ_{1j} is significant, a subcenter is called a local subcenter, and if δ_{2j} is significant, a subcenter can be considered a global subcenter. The last term, u_i, is the error. The $DCDB_i$ continues to represent the distance between site *i* and the CBD.

In the second step McMillen (2001) uses a highly flexible function to estimate $DCDB_i$. As the author indicates, it can be used the LWR, cubic splines or a kernel procedure. In his work it was chosen a Fourier function. In our estimations we considered an LWR, the same as in the initial smooth. As Krehl (2016) points out, in the first step the focus was to estimate a smooth employment density, though, at this moment the focus is to approximate the employment density function the best way possible.

---

[21] The nearest cells centroids that contain employees are 1km distant from the CBD and the furthest cell centroid is 18.79km apart.

[22] The significance level is: $(y_i - \hat{y}_i)/(\hat{\sigma}_i) > 1.96$. Where $\hat{y}_i$ is the estimated ln-density of *y* at site *i* and $\hat{\sigma}_i$ is the estimated standard error for the prediction in site *i*.

[23] In McMillen's (2001) work he analyzes metropolitan areas with large area (kilometers squared): Chicago (12,450.07), Dallas (15,415.61), Houston (16,871.183), Los Angeles (13,750.25), New Orleans (7,270.097) and San Francisco (11,046.3). In this study is only analyzed the data for the city of Curitiba, not taking into consideration nearby cities nor the metropolitan area.

A problem that might appear is multicollinearity when multiple distance variables are incorporated in the regression caused by a large amount of candidate subcenter sites. To solve this issue McMillen (2001) uses a stepwise procedure in order to determine the subcenter candidates. The estimated coefficients of both distance variables (level or inverse) are positive when the site is close to a subcenter. The subcenter distance variable that scores the lowest *t-value* is excluded, even if the variable is statistically significant. The smaller regression is then estimated. This procedure is repeated until all subcenter distance variables are statistically significant at the 20% confidence level. Then, the subcenter candidates are those sites that have positive coefficients on either one or both D_j and D_j^{-1} after the stepwise procedure is finished.

Therefore, the procedure proposed by McMillen (2001) provides statistical criteria to analyze the presence of employment subcenters within a region. Nonetheless, this methodology overcomes *ad hoc* definitions to determine the presence of SBD is a site that has a significant impact in the estimated employment density.

McMilllen (2001) also searches to determine the extent to which a variable related to the proximity of a subcenter improves the goodness of fit of the logarithm of the employment density function related to the distance of the CBD. Therefore, he uses a gravity variable to represent the possible outcomes linked to the proximity of a subcenter.

The gravity variable for the *i-th* observation is:

$$Gravity_i = \sum_{j=1}^{S} \frac{\hat{f}(x_j)}{D_{ij}^{\alpha}} \quad (3)$$

Where the term S is the number of subcenters in the studied region. The D_{ij} variable measures the distance between a subcenter j and the *i-th* observation. The parameter α represents a decay rate. This parameter is estimated starting at 0.25 adding 0.25 until it reaches the value of 3.0. The chosen value of the parameter is the one that contributes to the highest explanation power.

Finally, the $\hat{f}(x_j)$ represents the already estimated density of the subcenter observations. In this context, the definition of density is referred to as the number of observations that are nearby. When applying the weight of the terms in their estimated densities, therefore placing less weight on remote subcenters, which tend to be surrounded by grids with little or no employment.

The estimated regression with the inclusion of the gravity variable is:

$$y_i = \beta_0 + \beta_1 DCBD_i + \beta_2 Gravity_i + u_i \quad (4)$$

The inclusion of the gravity variable contributes to assess whether the CBD is the main contributor to the spatial distribution of the densities in the studied area. Hence, the inclusion of this variable contributes to acknowledge whether the CBD is the main driving force in attracting employment in the urban context.

Furthermore, the data available for the Origin-Destination (OD) survey for 2017 will be used to analyze commuting flows of workers. The weight variable is not available and, therefore, each individual is attributed the same weight. This might hinder possible inferences. Nevertheless, at this moment the analysis will be restricted to individuals that commute from other municipalities to Curitiba's CBD or SBDs. In order to do so, we restrict our analysis only for the morning commuters that are travelling from their residence to their workplace.

In order to match the zoning system of the OD survey and the employment grid cells we considered in our analysis all the individuals that worked in the zones inside (entirely or not) a cell that had an employment subcenter. To analyze the commute between the zone of an individual's residence and the SBD we created a straight line between the respective centroids.

5. RESULTS AND DISCUSSION

The grid cell considered as the CBD for Curitiba was the one that had the highest employment density. As it was expected, the CBD is located in the downtown area and in close proximity to the five main axes of the BRT network, as it is shown in Figure 3.

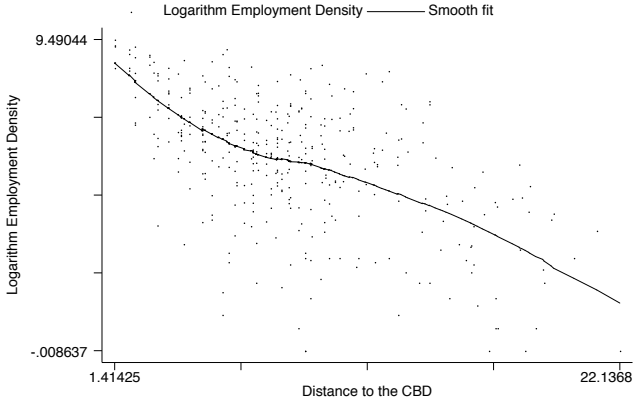
In order to estimate the first step of the procedure the distance between the *i-th* observation to the CBD was measured. The results LWR estimation provided 180 subcenter candidates significant at the five percent level. The results for the initial smooth can be seen in Figure 2. Considering the large number of candidates, in order to avoid multicollinearity, it was applied the three-kilometer radius to each observation with the highest residuals. After this removal, the number of subcenter candidates lowered to 25.

Further, the stepwise procedure eliminated more 15 observations, resulting in the final list of ten subcenters. Out of those SBDs, six were considered as Global and four as Local Subcenters. The location within the urban space to the main axes of the BRT system and feeder lines are shown in Figure 3.

It worth analyzing separately subcenter 402. As it can be seen in Figure 3, the SBD is surrounded by low density tracts, which consequently turns his density significantly higher than the neighbors. Also, his density is significantly lower when compared to the other global subcenters. As it can also be seen in Figure 3, center 402 has limited interaction to the BRT network. Therefore, despite being statistically significant, this subcenters appears to be spurious.

The results for the two-step estimation show that the SBDs concentrate a considerable amount of employment. The firm's specialization in each grid cell were classified considering a three-digit level firm classification of the International Standard Industrial Classification (ISIC) shown in Table 2. As it would be expected, the CBD embraces several firms. The largest firm of the center employs 1.391workers and only concentrates 0,04% of the employment in the region.

Figure 2 – Locally Weighted Regression Estimates



Nevertheless, most of the global subcenters are specialized in the service sector. Only one of those centers is in the industrial sector. Also, two SBDs out of the five relate to transportation activities. And the highest concentration of firm's employees within the grids are in those two transportation firms.

It should be mentioned that the procedure of McMillen (2001) does not capture the extent of the subcenters, whether, high employment density nodes. When comparing his procedure with the ones proposed by Giuliano and Small (1998) and Redfearn (2007), both estimate the size of the SBDs, accounting for a higher proportion of employment. Hence, it is expected that the share of employment from the McMillen (2001) procedure within the subcenters is significantly lower.

Table 2 - Descriptive Analysis at the firm level

	Employment (total)	Highest Employer		Economic Activity *	
		Total	%	Level	Discription
CBD	37.574	1.391	3,70	853	Higher education
Global					
153	6.194	5.048	81,50	492	Other land transport
260	5.112	1.068	20,89	293	Manufacture of parts and accessories for motor vehicles
316	9.823	2.054	20,91	861	Hospital activities
391	5.873	687	11,70	812	Cleaning Activities
402	295	99	33,56	492	Other land transport
624	2.640	1.924	72,88	491	Transport via railways
Local					
79	1.643	599	36,46	429	Construction of other civil engineering projects
180	9.931	6.657	67,03	351	Electric power generation, transmission and distribution
432	1.071	459	42,86	471	Retail sale in non-specialized stores
618	3.783	1.011	26,72	429	Construction of other civil engineering projects

* Three digit-level classification of the International Standard Industrial Classification

Source: RAIS (2010)

Table 3 - Descriptive Analysis of Employees

	Employment (total)	Proportion of Men	Age (mean)	Years of Education (proportion)			Income (mean) *
				Abandoned High School (%)	High School Graduate (%)	Bachelors Degree (%)	
CBD	37.574	40,94	33,07	7,76	44,02	25,23	1871,71
Global							
153	6.194	56,15	32,30	10,76	61,22	7,52	1.897,18
260	5.112	89,71	33,97	6,84	54,65	6,99	2.033,04
316	9.823	43,51	35,28	7,11	40,06	18,38	1.696,74
391	5.873	60,59	34,34	10,37	49,29	9,77	1.267,76
402	295	84,40	36,22	10,50	52,20	2,71	1.252,49
624	2.640	86,47	33,44	4,74	63,78	13,47	1.999,58
Local							
79	1.643	87,82	33,74	10,34	38,58	6,39	1.663,22
180	9.931	79,20	38,79	1,95	45,85	39,46	3.887,14
432	1.071	57,51	32,82	18,86	36,78	5,69	902,95
618	3.783	69,01	35,13	9,99	48,69	6,71	1.593,01

Source: RAIS (2010)

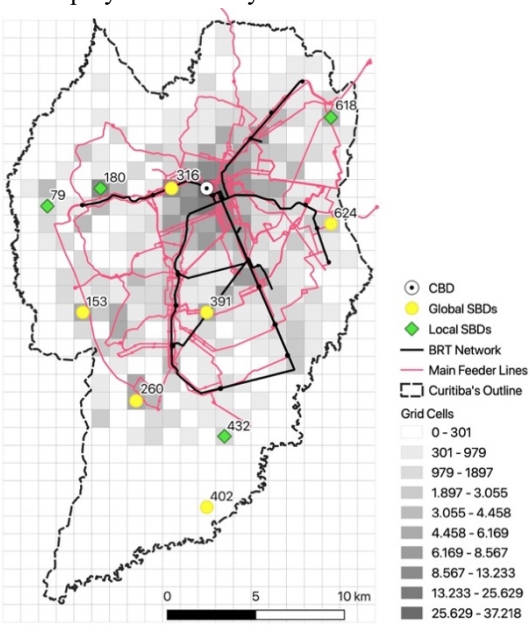
In the case of the local SBDs, the highest employers in the grid concentrates more employment when compared to global subcenters. Further, two SBDs are specialized in the construction sector, one in retail and the other on the provision of public utilities. As it can be seen in Figure 3, three subcenters locate near one another, but within the 3-kilometer radius to avoid multicollinearity.

Focusing on the individual level data it can be observed a strong correlation between years of education and income within the SBDs. The subcenter 432 has the

highest proportion of employees that abandoned high school and the lowest that did graduate high school. Also, it is the second to last in the proportion of workers that have a bachelor's degree. Therefore, as it can be seen in Table 3, this SBD has the lowest income of all subcenters.

The SBD with the highest income not only has the most employees with bachelor's degree but also the highest age average within all subcenters. Given these two statements this result should be expected, confirming the relationship between years of education and experience. It is worth mentioning that the average income in this grid is two times greater than the average for the CBD. Further, it should be taken into consideration the level of education within all SBDs, revealing that only two of those subcenters do not have a proportion of high school graduates higher than 40%.

Figure 3 – Employment Density and the SBDs at the Grid Level



Moreover, when considering the high school graduates and those with bachelor's degree, the proportion of employees with high level of human capital within the SBDs are over 55%. Except in the two cases mentioned in the previous paragraph.

Taking into consideration only the mains axes of the BRT network, four subcenters are located within a close range of the network. While also considering the main feeder lines of the network only two subcenters are not located adjacent to the transportation network. Nevertheless, local subcenters 79 and 432 are located the farthest from the network, as it can be seen in Table 4.

Therefore, as it was expected, given the results in Shertzer *et al* (2018), Twinam (2018) and Redfearn (2007) there is a strong relationship between the TOD policy and the land use regulations which lie behind it in the promotion of employment subcenters within Curitiba's urban space and their close proximity to the BRT network.

Analyzing the commute flows towards the CBD, it would be expected that since Curitiba and its CBD concentrates most of the employment within the region, that individuals that reside in surrounding cities commute daily to the region. The evidence shown in Figure 4 support this view.

Table 4 - Distance between SBDs to the CBD and BRT system (kilometers)

	CBD	BRT	Feeder Network
Global			
153	9,90	5,00	0,34
260	12,65	2,77	0,35
316	2,00	0,05	0,05
391	7,00	0,07	0,01
624	7,28	0,76	0,07
Local			
79	9,06	1,90	1,09
180	6,00	0,71	0,50
432	14,04	2,63	1,27
618	6,06	2,53	0,11

The pattern for the Global SBDs is similar to those observed in the CBD, with a strong outreach to nearby and outer municipalities. Furthermore, it can be seen that the majority of the interviewed individuals residing in different cities, do locate near the SBD. In the case of the Local SBDs it can be observed that there is an outreach to other municipalities of the RIT but not as strong as the Global ones.

Furthermore, the use different years across the databases in the study enables to infer if the subcenters estimated for 2010 are still relevant in the influx of workers. In this case, the only subcenter that draws a low number of workers in 2017 was node 402, as seen in Figure 5. This was the only subcenter with the predominant employer in the retail sector. This suggests that this SBD no longer exists.

Finally, the inclusion of the gravity variable in the model searches to determine whether the CBD is the main employment attractor within the urban context. At this moment, the analysis returns its focus exclusively to Curitiba.

The reported coefficients for the estimation including gravity leads us to infer that the CBD is no longer determines the spatial trends in the employment densities within Curitiba. Further, the estimated coefficients for this case follow the same signals as those reported by McMillen (2001) for the cities whose CBD no longer are the main driving force in employment location.

Table 5 - Coefficients for the log employment regression including the gravity variable

	Coefficient		Standard Error		t-value		p-value	
	Not Including	Including	Not Including	Including	Not Including	Including	Not Including	Including
Constant	8.2242	-5.2209	0.1897	1.579	43.33	-3.31	0.000	0.000
DCBD	0.2745	0.2024	0.02011	0.0593	-13.65	3.31	0.000	0.001
Gravity	-	0.2722	-	0.0318	-	8.55	-	0.001
R^2	0.3266	0.4534						
Observations	383	373						
α	-	0.25						

Figure 4 – Commute Flows to the CBD.

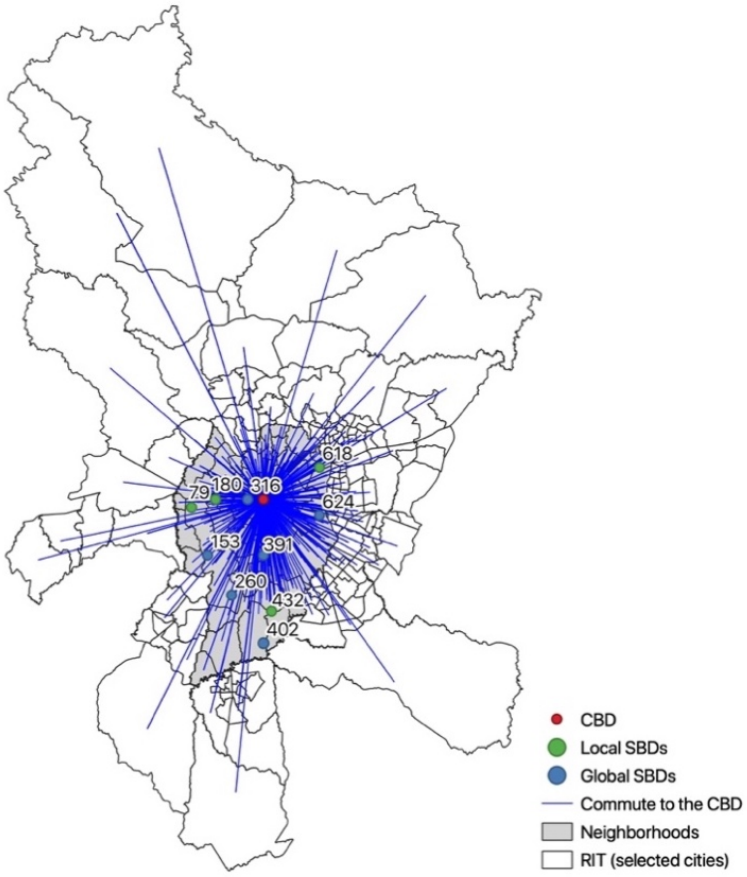


Figure 5 – Commute Flows to the Global SBDs

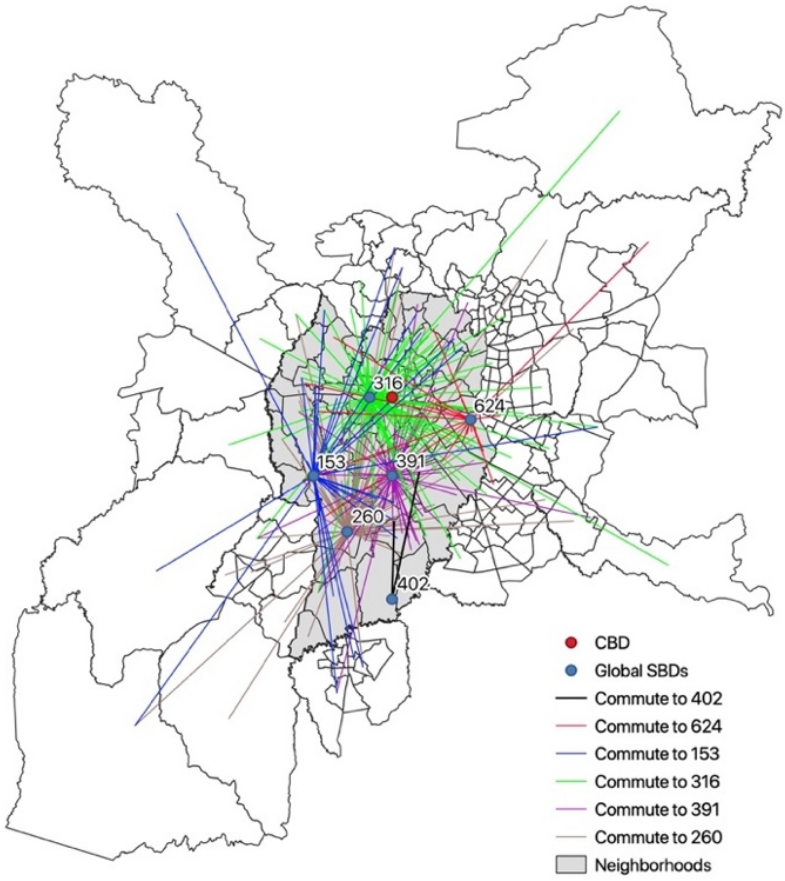
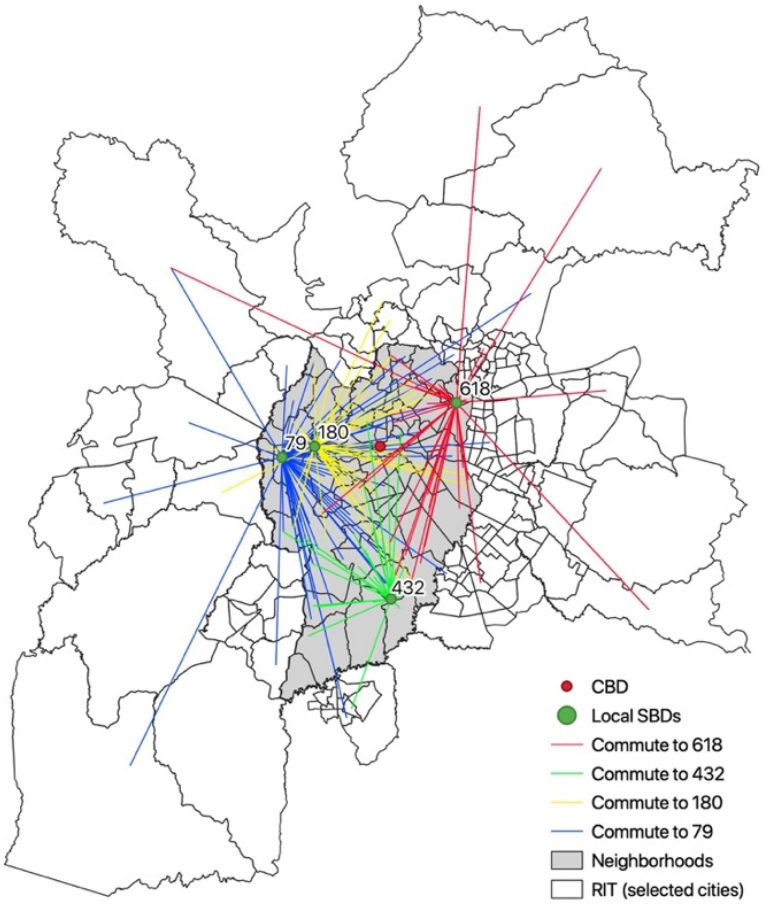


Figure 6 – Commute Flows to the Local SBDs



6. CONCLUSION

This paper intended to determine whether Curitiba (Brazil) has employment subcenters within its urban context and how do they relate to the TOD policies that have been developed and implemented since the 1970's. Further, it was also explored whether the CBD still is the most influential location in the spatial trends of employment density. Finally, it was also investigated the outreach of the subcenters to other municipalities comprising the RIT.

In order to do so, the chosen estimation procedure was the one proposed by McMillen (2001). This methodology was preferred because it relies mostly in statistical criteria in order to determine employment subcenters, removing possible *ad hoc* definitions. The opposite of the proposed procedure by Giuliano and Small (1991).

The results confirm the initial hypothesis. Therefore, Curitiba is indeed a polycentric city with the presence of nine subcenters. Out of those, four are considered to be local subcenters and the six remainder subcenters are global.

Further, the inclusion of the gravity variable sheds light to the fact that Curitiba's CBD is no longer the main determinant of the employment location in the urban context. Therefore, the employment subcenters can be considered to be influential in the attraction of employment in the city. Further, the spatial distribution of these subcenters is highly correlated to the BRT network, evidencing the role of TOD urban planning policy in shaping the polycentric structure.

The analysis of the OD survey for the year of 2017 show that the CBD has more capillarity to other municipalities within the RIT when comparing to the SBDs.

Nevertheless, the Global SBDs have a higher outreach to other outer locations in comparison to the Local subcenters. Nevertheless, the unavailability of the weights hinders a deeper comprehension of the commuting flows within the RIT.

Furthermore, the policies developed through the BRT network, which promoted the development of subcenters, in consonance with the results in the model that included the gravity variable, show that the presence of SBDs in Curitiba is not a phenomenon, but rather a trend.

Considering the results when including the gravity variable, further research should investigate the preferred transportation modal according to wage and other socioeconomic variables. In order to do so, estimating adequate weights for the OD survey should be necessary.

REFERENCES

ALONSO, W. **Location and land use**: Toward a general theory of land rent. Cambridge, Harvard University Press, 1964.

ANAS, A.; ARNOTT, R.; SMALL, K. A. Urban spatial structure. **Journal of Economic Literature**, v. 36, p. 1426–1464, 1998.

BAUMONT, C.; ERTUR, C.; Le GALLO, J. Spatial Analysis of Employment and Population Density: The Case of the Agglomeration of Dijon 1999. **Geographical Analysis**, v. 36, p. 146-176, 2004.

BELMIRO, C. P.; RODRIGUES, F.; SILVEIRA NETO, R. Monocentrismo e Estrutura Urbana: uma Análise Empírica para a cidade do Recife. In: **Encontro Nacional de Economia**, Anais, 46, 2018, Rio de Janeiro: ANPEC, 2018.

BERRY, B. J. L.; KIM, H.-M. Challenges to the monocentric model. **Geographical Analysis**, v. 25, p. 1–4, 1993.

CAMPOS, R. B. A.; CHAGAS, A. S. Identificação e caracterização das subcentralidades de emprego no município de São Paulo. In: **Encontro Nacional de Economia**, Anais, 45, 2017, Natal: ANPEC, 2017.

CERVERO, R. WU, K-L. Sub-centring and Commuting: Evidence from the San Francisco Bay Area, 1980-90. **Urban Studies**, v. 35, p. 1059-1076, 1998.

CERVERO, R.; Dai, D. BRT TOD: Leveraging transit-oriented development with bus rapid transit investments. **Transport Policy**, v. 36, p. 127–138, 2014.

CRAIG, S. G.; NG, P. Using quantile smoothing splines to identify employment subcenters in a multicentric urban area, **Journal of Urban Economics**, v. 49, p. 100–120, 2001.

DOU, Y.; LUO, X.; DONG, L; WU, C; LIANG, H.; REN, J. An Empirical Study on Transit-Oriented Low-Carbon Urban Land Use Planning: Exploratory Spatial Data Analysis (ESDA) on Shanghai, China. **Habitat International**, v. 53, p. 379-389, 2015.

DUARTE, F.; ULTRAMARI, C. Making public transport and housing match: accomplishments and failures of Curitiba's BRT. **Journal of Urban Planning and Development**, v. 138, n. 2, p. 183–194, 2012

FERNÁNDEZ-MALDONADO, A. M.; ROMAIN, A.; Otto VERKOREN, O.; PAULA-PESSOA, R. P. Polycentric Structures in Latin American Metropolitan Areas: Identifying Employment Sub-centres, **Regional Studies**, v. 48, n. 12, p. 1954-1971, 2014

FUJITA, M.; OGAWA, H. Multiple equilibria and structural transition of non-monocentric urban configurations. **Regional Science and Urban Economics**, v. 12, n. 2, p. 161–196, 1982.

G. GIULIANO, G.; SMALL, K. A. Subcenters in the Los Angeles region, **Regional Science and Urban Economics**, v. 21, p. 163–182, 1991.

IBGE (2018a). Produto Interno Bruto a preços correntes segundo os Municípios - 2015. Disponível em https://www.ibge.gov.br/estatisticas-novoportal/economicas/contas-nacionais/9088-produto-interno-bruto-dos-municipios.html?=&t=downloads. Consultado em 14 de outubro de 2018.

IBGE (2018b). Estimativas da População segundo as os Municípios - 2018. Disponível em https://www.ibge.gov.br/estatisticas-novoportal/sociais/populacao/9103-estimativas-de-populacao.html?=&t=downloads. Consultado em 14 de outubro de 2018.

Instituto de Pesquisa e Planejamento Urbano de Curitiba (IPPUC). Dados Geográficos 2018, disponível em http://ippuc.org.br/geodownloads/geo.htm.

KHAYESI, M.; AMEKUDZI, A. A. Kingdon's multiple streams model and automobile dependence reversal path: the case of Curitiba, Brazil. **Journal of Transport Geography**, v. 19, n. 6, p. 1547-1552, 2011.

KREHL, A. Urban subcentres in German city regions: Identification, understanding, comparison, **Papers in Regional Science**, v. 97, p. S79-S104, 2016.

LV, Y.; ZHENG, X.; ZHOU, L.; ZHANG, L. Decentralization and Polycentricity: Spatial Changes of Employment in Beijing Metropolitan Area, China. **Sustainability**, vol. 9, n. 10, pages 1-17, 2017.

MACEDO, J. Curitiba, **Cities**, v. 21, n. 6, p. 537-549, 2004.

McDONALD, J. F. The identification of urban employment subcenters. **Journal of Urban Economics**, v. 21, p. 242–258, p. 1987.

McMILLEN, D. P. Employment subcenters in Chicago: past, present, and future. **Economic Perspectives**, Federal Reserve Bank of Chicago, v. 27, n. 2, pages 2-14, 2003.

McMILLEN, D. P. Nonparametric employment subcenter identification. **Journal of Urban Economics**, v. 50, p. 448–473, 2001.

MILLS, E. S. An aggregative model of resource allocation in a metropolitan area. **The American Economic Review**, v. 57, p. 197–210, 1967.

MULICEK, O; MALY, J. Moving towards more cohesive and polycentric spatial patterns? Evidence from the Czech Republic. **Papers in Regional Science**, v. 98, p. 1177-1194, 2019

MUTH, R. **Cities and housing**: The spatial pattern of urban residential land use. Chicago: University of Chicago, 1969.

RABINOVITCH, J. Innovative land use and public transport policy: The case of Curitiba, Brazil. **Land Use Policy**, v. 13, p. 51-67, 1996.

REDFEARN, C. L. The topography of metropolitan employment: Identifying centers of employment in a polycentric urban area. **Journal of Urban Economics**, v. 61, p. 519–541, 2007.

SHERTZER, A.; TWINAM, T.; WALSH, R. P. Zoning and the economic geography of cities. **Journal of Urban Economics**, v. 105, p. 20-39, 2018.

SHUKLA, V.; WADDELL, P. Firm location and land use in discrete urban space: A study of the spatial structure of Dallas-Fort Worth. **Regional Science and Urban Economics**, v. 21, p. 225–253, 1991.

TWINAM, T. The long-run impact of zoning: Institutional hysteresis and durable capital in Seattle, 1920–2015. **Regional Science and Urban Economics**, v. 73, p. 155-169, 2018.

VENERI, P. The identification of sub-centres in two Italian metropolitan areas: A functional approach. **Cities**, v. 31, p. 177-185, 2013.

WADDELL, P.; SHUKLA, V. Manufacturing location in a polycentric urban area: A study in the composition and attractiveness of employment Subcenters. **Urban Geography**, v. 14, p. 277–296, 1993.

XIE, Z.; YE, X.; ZHENG, Z.; LI, D.; SUN, L.; LI, R.; BENYA, S. Modeling Polycentric Urbanization Using Multisource Big Geospatial Data. **Remote Sensing**, v. 11, n. 310, 2019.

YANG, T.; PAN, H.; HEWINGS, G.; JIN, Y. Understanding urban sub-centers with heterogeneity in agglomeration economies — Where do emerging commercial establishments locate?. **Cities**, v. 86, p. 25-36, 2019.

YUE, W.; LIU, Y.; FAN, P. Polycentric Urban Development: The Case of Hangzhou. **Environment and Planning A**: Economy and Space, v. 42, n. 3, p. 563–577, 2010.