TRANSPORTES
revistatransportes.org.br
anpet
associação nacional de pesquisa e ensino em transportes
anpet.org.br

Spatial relationship between sociodemographic and retail access data and e-commerce deliveries: the case of Belo Horizonte (Brazil)

Relação espacial entre dados sociodemográficos e de acesso ao varejo e as entregas do comércio eletrônico: o caso de Belo Horizonte (Brasil)

Luísa Tavares Muzzi de Sousa[1], **Isabela Kopperschmidt de Oliveira**[1], **Leise Kelli de Oliveira**[1,2], **Jorge Luiz dos Santos Junior**[2], **Bruno Vieira Bertoncini**[3]

[1]Universidade Federal de Pernambuco, Recife, Pernambuco, Brasil
[2]Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brasil
[3]Universidade Federal do Ceará, Fortaleza, Ceará, Brasil

Contact: luisamuzzi29@gmail.com, (LTMS); isa.kopper@gmail.com, (IKO); leise@etg.ufmg.br, (LKO); jorgelsjunior@gmail.com (JLSJ); bruviber@det.ufc.br, (BVB)

Submitted:
3 August, 2022
Accepted for publication:
4 April, 2023
Published:
25 August, 2023

Associate Editor:
Renato da Silva Lima, Universidade Federal de Itajubá, Brasil

Keywords:
E-commerce deliveries.
Urban freight transport.
Spatial analysis.

Palavras-chave:
Entregas de comércio eletrônico.
Transporte urbano de mercadorias.
Análise espacial.

DOI: 10.58922/transportes.v31i2.2820

ABSTRACT
Many negative externalities are associated with home deliveries, the main e-commerce delivery destination. Despite many solutions that address this problem, the lack of understanding of the spatial pattern of urban deliveries makes it challenging to implement these strategies. This paper analyzed the spatial relationship between sociodemographic, retail and e-commerce deliveries in Belo Horizonte (Brazil) by using official data at the neighborhood level (number of retail shops, gender, income, age, race, and household size), and e-commerce deliveries performed by a transportation company. Global Moran's I indicated the spatial dependence of the e-commerce deliveries. Results of a geographically weighted regression model showed a positive spatial effect of retail, women, Asian population, age from 20 to 29 years old, and income. In addition, a negative spatial effect was identified for the size of the household, 18 to 19 years old, and the black population. Furthermore, the estimated coefficients show small spatial variability, indicating homogeneity in the spatial relation. The uniformity of the parameters indicates that alternative strategies can be implemented throughout the territory to reduce e-commerce deliveries.

RESUMO
As entregas ao domicílio derivadas do comércio eletrônico são uma grande preocupação em áreas urbanas devido às suas externalidades negativas associadas. Apesar de muitas soluções para esse problema, a falta de compreensão do padrão espacial das entregas urbanas torna difícil a implementação dessas estratégias. Neste artigo é analisada a relação espacial entre dados sociodemográficos e de acesso ao varejo e entregas de comércio eletrônico em Belo Horizonte (Brasil) usando dados oficiais em nível de bairro (número de lojas tradicionais de varejo, gênero, renda, idade, raça e tamanho da família) e dados operacionais de uma empresa de transporte. O índice Global Moran's I indicou a dependência espacial das entregas do e-commerce. Os resultados de um modelo de regressão geograficamente ponderada mostraram um efeito espacial positivo do acesso ao varejo tradicional, mulheres, população asiática, idade de 20 a 29 anos e renda. Além disso, foi identificado efeito espacial negativo para o tamanho do domicílio,

idade de 18 a 19 anos e população negra. Além disso, os coeficientes estimados apresentam uma pequena variabilidade espacial, indicando homogeneidade na relação espacial. A uniformidade dos parâmetros permite concluir que estratégias alternativas de entrega em domicílio podem ser implementadas de forma equitativa em todo o território para reduzir as externalidades das entregas de e-commerce e, consequentemente, contribuir para o desenvolvimento sustentável.

1. INTRODUCTION

E-commerce is one consequence of the growth of the digital economy and the experience of digitization (UNCTAD, 2021a). E-commerce means the way to purchase or sell goods or services carried out over computer networks by methods specifically designed to receive or place orders (OECD, 2020). These networks request products and services; the payment and delivery can be online or offline. The share of e-commerce in global retail trade increased from 10.4% in 2017 to 14.1% in 2019 (UNCTAD, 2021a). In the same period, the number of e-consumers (i.e., e-commerce consumers) over 15 years old increased from 1.3 billion in 2017 to 1.48 billion in 2019 (UNCTAD, 2021b). In Brazil, e-commerce grew from 54 million in 2011 to 148 million in 2019. Furthermore, business-to-consumer (B2C) sales increased 41% between 2019 and 2020, surpassing 87 billion R$ in sales (Nielsen, 2022).

The growth of e-commerce impacts urban deliveries, as each order has the potential to generate one or more deliveries (depending on the number of products and the delivery strategy), which are usually destined for the place of residence or work of the consumer. Therefore, understanding the spatial distribution of e-commerce deliveries is also understanding e-commerce, which could provide measures to achieve the Sustainable Development Goals (SDGs) (UNCTAD, 2021a).

Hood et al. (2020) show that there is variation in sociodemographic profiles regarding the choice of distribution channel for grocery products purchased through e-commerce. Therefore, it can be assumed that sociodemographic data is fundamental for understanding e-commerce deliveries. Based on this issue, this paper addresses the following research question: To what extent do sociodemographic factors and traditional retail influence the spatial distribution of e-commerce deliveries? Therefore, this paper aims to analyze the spatial relationship between sociodemographic and retail data and e-commerce deliveries. Data from e-commerce deliveries from a transport company operating in the Belo Horizonte market were used. Geographically weighted regression models were estimated to identify the spatial relationship between the variables considered in this study and e-commerce deliveries.

The contribution of this paper is twofold. First, the spatial effect of e-commerce deliveries as analyzed, as suggested by Song (2021). Second, many scholars report the lack of studies for developing countries (Ren and Kwan, 2009; Loo and Wang, 2018; Kirby-Hawkins et al., 2019; Cheng et al., 2021; Song, 2021). Therefore, this paper addresses the case of Belo Horizonte, a Brazilian city.

This paper has five sections in addition to this introductory section. Section 2 presents the literature concerning e-commerce deliveries and socioeconomic data. Section 3 describes the spatial distribution of B2C deliveries and socioeconomic characteristics of the Belo Horizonte. The research method is detailed in section 4, and its results are presented and discussed in section 5. Finally, conclusions are presented in section 6.

2. RELATION BETWEEN E-COMMERCE DELIVERIES AND SOCIOECONOMIC DATA

The literature suggests some relation between e-commerce deliveries and socioeconomic data. This section reports the main findings reported by the literature on gender, age, race, income, household size, and access to physical stores.

Concerning gender, the proportion of men has a positive effect on e-commerce (Farag et al., 2003, 2005, 2006a, b; Clarke et al., 2015; Beckers et al., 2018), as well as the proportion of women (Ren and Kwan, 2009; Sener and Reeder, 2012; Loo and Wang, 2018; Hood et al., 2020; Jaller and Pahwa, 2020; Saphores and Xu, 2021). However, Farag et al. (2007) identified a negative effect on the proportion of women. Nonetheless, there seems to be a positive effect on the proportion of men or women in e-commerce purchases.

Regarding age, the literature points out that the effect of age on e-commerce depends on the age group. People between 25 and 35 years are more likely to buy online (Farag et al., 2003). Moreover, there is a positive influence of people between 25-45 years (Farag et al., 2005) and 26-45 years (Clarke et al., 2015). Similarly, Sener and Reeder (2012) identified a positive influence of the age group 22 to 36 years, while Farag et al. (2006a) concluded that people up to 33 years are more likely to use e-commerce, and Loo and Wang (2018) verified the positive influence of people between 21 and 30 years. Song (2021) observed that people under 40 years are more likely to buy using e-commerce.

However, Lee et al. (2015) concluded that people under 35 years are 1.7 times more likely to buy online than those over 50 years, while people between 35 and 50 years were twice as likely to use e-commerce than older age groups. Beckers et al. (2018) found that the group of people between 30 and 39 years old, 40 and 49 years old, and 60 and 69 years old have a positive influence on e-commerce use, while people between 50 and 59 years old and over 70 years old have a negative influence. Hood et al. (2020) verified the positive influence of people between 25 and 44 years old and over 55 years old.

Few scholars have analyzed the influence of race on e-commerce and the white people are the most frequent users of online shopping (Ren and Kwan, 2009; Saphores and Xu, 2021). Regarding income, a positive effect was identified in all studies identified in the literature (Farag et al., 2005, 2006a, b; Krizek et al., 2005; Blasio, 2008; Sener and Reeder, 2012; Cao et al., 2013; Zhou and Wang, 2014; Clarke et al., 2015; Lee et al., 2015; Lee et al., 2017; Motte-Baumvol et al., 2017; Beckers et al., 2018; Hood et al., 2020; Cheng et al., 2021; Saphores and Xu, 2021; Song, 2021).

Regarding the size of the household (number of inhabitants per household), this factor has a positive effect on e-commerce purchases in studies developed by Sener and Reeder (2012) and Cheng et al. (2021) and a negative effect in the study by Zhou and Wang (2014). However, only Song (2021) analyzed the spatial effect of these variables. The authors suggest more studies considering the spatial effects of e-commerce.

Furthermore, the literature indicates that the influence of traditional retail stores on e-commerce depends on the characteristics of the study area. However, the literature has no consensus on the analyses. Most scholars found a positive influence of access to physical stores (Farag et al., 2006a, b, 2007; Blasio, 2008; Cao et al., 2013; Kirby-Hawkins et al., 2019; Hood et al., 2020). However, some scholars also found that people with lower access to traditional retail shopped more online (Ren and Kwan, 2009; Clarke et al., 2015; Motte-Baumvol et al., 2017; Loo and Wang, 2018; Cheng et al., 2021).

Table 1 summarizes the literature on the relationship between sociodemographic factors and e-commerce. Most studies relate to age and income as factors that explain e-commerce. Additionally, most of these studies considered the context of developed countries, such as European countries and the United States, except for the studies developed by Loo and Wang (2018) and Song (2021) carried out in China. This concentration of studies in developed countries is most likely due to the low availability of data in developing countries (UNCTAD, 2021a). Although e-commerce is data-driven, there are challenges in collecting and analyzing data regarding the e-commerce experience. Moreover, few countries have collected data with sufficient detail for formulating public policies (UNCTAD, 2021a).

Table 1: Synthesis of the literature

	Socioeconomic factors						Place		Data		Technique				
Reference (Chronological order)	**Gender**	**Age**	**Race**	**Income**	**Household size**	**Access to physical shops**	**Developed countries**	**Developing countries**	**Real data**	**Self-declared survey data**	**Logistic regression**	**Structural equation modeling**	**Copula modeling**	**Spatial regression**	**Other**
Farag et al. (2003)	+	•					•			•	•				
Farag et al. (2005)	+	•		+			•			•					•
Krizek et al. (2005)		•		+			•			•	•				
Farag et al. (2006a)	+	•		+		+	•			•	•				
Farag et al. (2006b)	+	•		+		+	•			•	•				
Farag et al. (2007)	-	•		+		+	•			•		•			
Blasio (2008)		•		+		+	•			•	•				
Ren and Kwan (2009)	+		•			-	•		•	•	•				
Sener and Reeder (2012)	+	•		+	+		•			•			•		
Cao et al. (2013)				+		+	•			•		•			
Zhou and Wang (2014)		•		+	-		•			•		•			
Clarke et al. (2015)	+	•		+		-	•			•	•				
Lee et al. (2015)		•		+			•			•	•				
Lee et al. (2017)				+			•			•			•		
Loo and Wang (2018)	+	•				-		•		•	•				
Motte-Baumvol et al. (2017)		•		+		-	•			•	•				
Beckers et al. (2018)	+	•		+			•			•	•				
Kirby-Hawkins et al. (2019)		•				+	•		•						•
Hood et al. (2020)	+	•		+		+	•			•					•
Jaller and Pahwa (2020)	+						•			•	•				
Saphores and Xu (2021)	+	•	•	+			•			•	•				
Cheng et al. (2021)				+	+	-	•		•						•
Song (2021)		•		+				•	•					•	

+: positive effect; - negative effect; •: paper has analyzed the factor.

In addition, most studies used self-reported data. Among the analysis methods, logistic regression was the most used in the literature. Furthermore, studies also used structural equation modeling and path analysis (Farag et al., 2005). Other techniques used are spatial interaction models (Kirby-Hawkins et al., 2019), statistical analyzes (Hood et al., 2020), linear regression (Cheng et al., 2021), and copula-based models formulated from multivariate distributions to represent different types of dependence (Sener and Reeder, 2012; Lee et al., 2017). However, since e-commerce deliveries do not present a homogeneous distribution in the territory, spatial models are a research opportunity to understand the spatial distribution of e-commerce (Song, 2021).

Specifically considering e-commerce deliveries, few scholars have considered e-commerce delivery data. Ren and Kwan (2009) analyzed the influence of socioeconomic data and traditional retail on the adoption of e-commerce using data from delivered packages extracted from a public database. Cheng et al. (2021) analyzed factors influencing home deliveries in Singapore using delivery data.

Based on the literature review, this section supports two main contributions of this paper. First, few scholars have used spatial techniques and delivery data to analyze e-commerce deliveries. Second, no studies were carried in Latin-American, where Belo Horizonte is located. Based on these contributions, the following section presents the data and the hypotheses developed for this paper.

3. DATA AND DEVELOPMENT OF THE HYPOTHESES

Belo Horizonte is the capital of Minas Gerais State and has an estimated population of over 2.5 million people spread over in 331 km^2 (IBGE, 2020). Socioeconomic data was obtained through the 2010 census and extrapolated to the year 2019. Figure 1 shows the spatial distribution of the retail shops, gender (men and women), household size (inhabitants per household), and income in Belo Horizonte. The spatial pattern of retail shops, gender, and household size are heterogeneous in the territory. Retail stores are concentrated in a few neighborhoods, mainly in the central region. The spatial pattern of gender is similar, with minor differences in the neighborhoods. On the other hand, the income is concentrated in the Center-South region and isolated neighborhoods in the West and Pampulha regions. Finally, the household size is less than 2.58 people/household in most neighborhoods.

Figure 2 shows the population distribution by race in the neighborhoods of Belo Horizonte. The concentration of white population is observed in the neighborhoods of the Center-South region. At the same time, the black, mixed, and Asian people are concentrated mainly in the Venda Nova region and more heterogeneously in the rest of the territory. In general, no race distribution pattern is observed in Belo Horizonte.

Figure 3 shows the distribution of the population by age group in the Belo Horizonte neighborhoods. There is a heterogeneous distribution of all age groups in the territory. However, the group between 30-49 years does not have the highest concentration in the central area, which differs from the other groups.

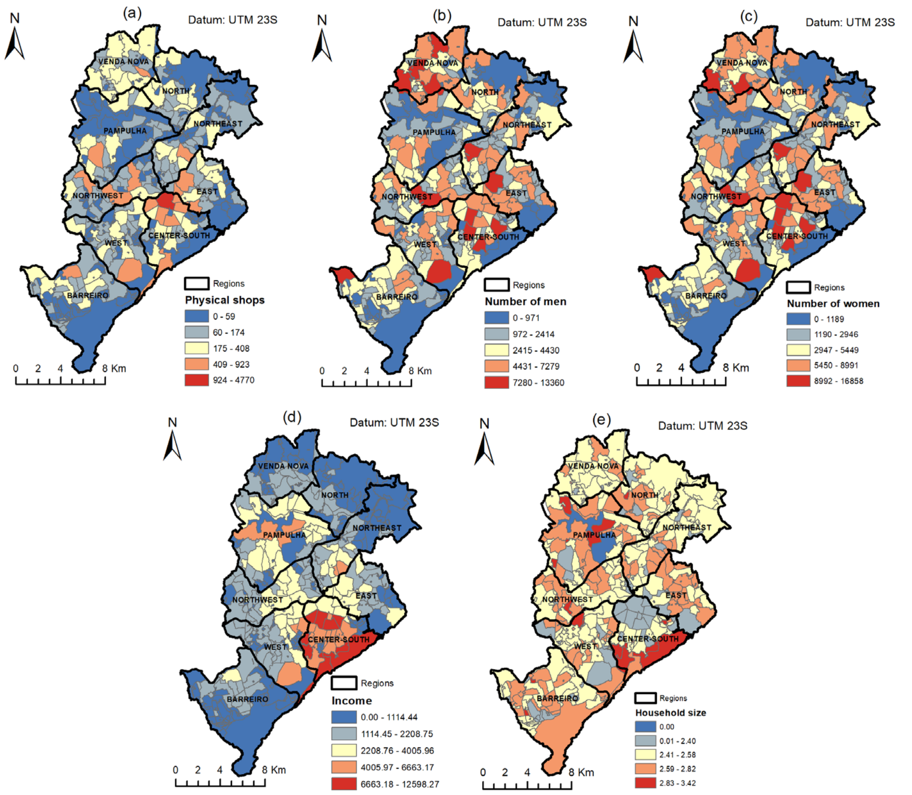

Figure 1. (a) Distribution of the retail shops; (b) number of men; (c) number of women; (d) average income by neighborhood; and (e) household size, in Belo Horizonte.

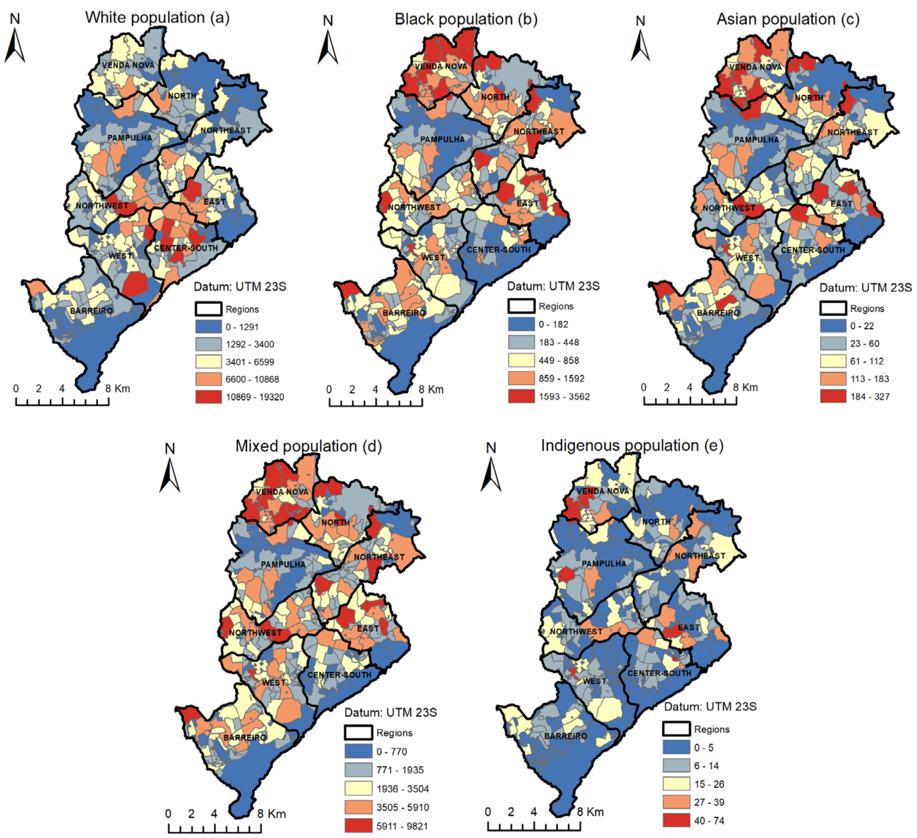

Figure 2. Distribution of races by neighborhood in Belo Horizonte: (a) white population; (b) black population; (c) Asian population; (d) mixed population; and (e) indigenous population.

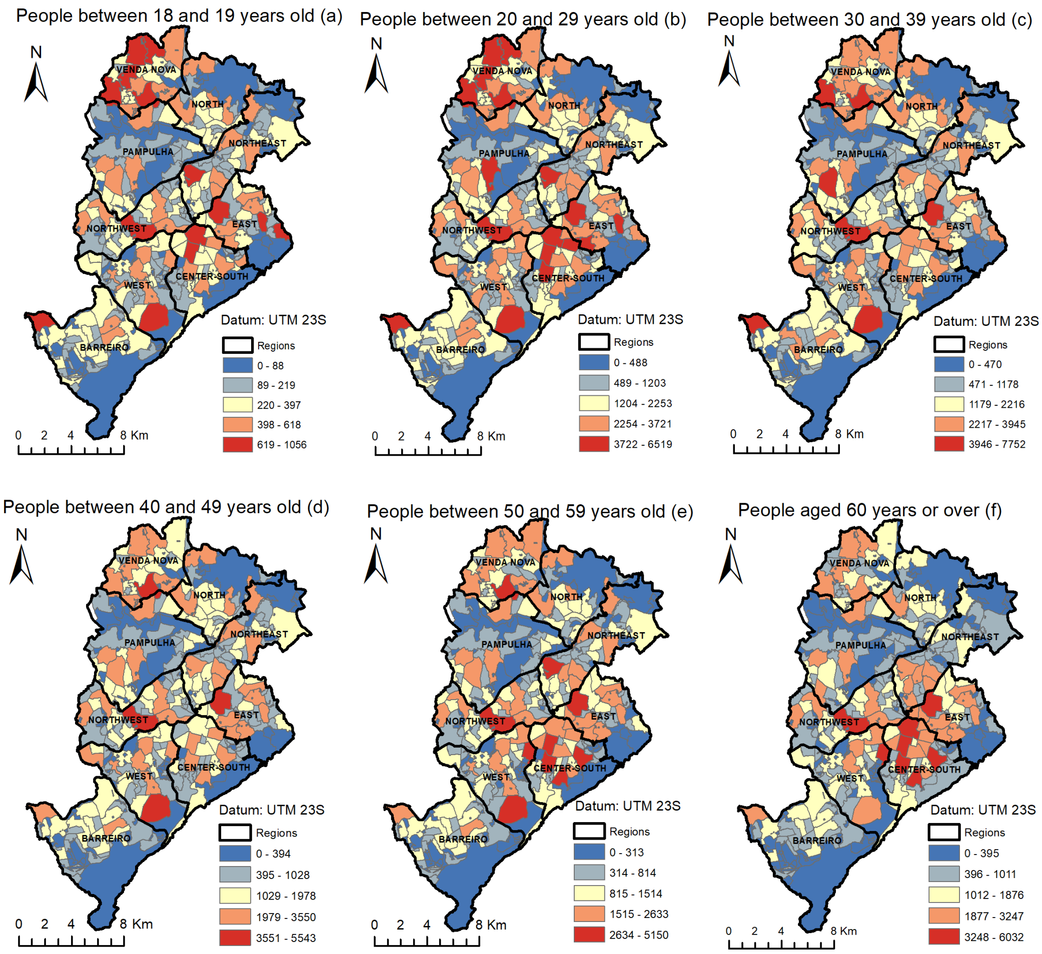

Figure 3. Distribution of the population by age group in the Belo Horizonte neighborhoods: (a) people between 18 and 19 years; (b) between 20 and 29 years; (c) between 30 and 39 years; (d) 40 and 49 years; (e) 50 and 59 years; and (f) 60 years or over (f).

The representation of e-commerce deliveries in 2019 is illustrated in Figure 4. These data were obtained with a carrier that, in 2019, performed more than 395 thousand deliveries in Belo Horizonte. Deliveries were more concentrated in the Center-South region. On the contrary, the lowest delivery volumes occurred in the south of the Barreiro region, the north portion of the North region, east of the East region, and some neighborhoods of the Pampulha region.

Based on these data and considering the literature, the following research hypotheses are proposed:

H_A: E-commerce deliveries are spatially dependent.

H_B: Gender has a positive spatial effect on e-commerce deliveries.

H_C: Age stratification has a heterogeneous spatial effect on e-commerce deliveries.

H_D: The stratification of race has a heterogeneous effect on e-commerce deliveries.

H_E: Income has a positive spatial effect on e-commerce deliveries.

H_F: The size of the household has a negative spatial effect on e-commerce deliveries.

H_G: The retail shops have a positive spatial effect on e-commerce deliveries.

These hypotheses outlined the research method detailed in the next section.

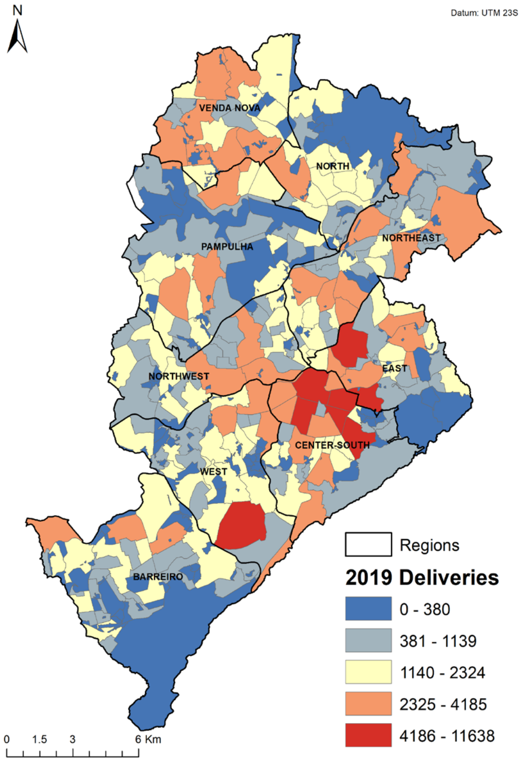

Figure 4. E-commerce deliveries by neighborhood in Belo Horizonte in 2019.

4. RESEARCH METHOD

The research method was based in two spatial techniques: (i) Global Moran's I index and (ii) geographically weighted regression. Global Moran's I index was used to verify the spatial autocorrelation of e-commerce deliveries in Belo Horizonte, addressing the hypothesis H_A. The spatial autocorrelation of a variable is the degree of spatial dependence between the value of an observation and the neighboring observations. Spatial dependence can be positive when similar values tend to be closer to each other in space or negative when different values are more relative to each other. When spatial autocorrelation is non-existent, the distribution of values is spatially random (Grekousis, 2020).

Determining spatial autocorrelation is necessary when investigating the spatial relationship and spatial patterns between values of a particular variable. When spatial autocorrelation is detected for a variable, spatial statistics techniques are suitable for evaluating the phenomena associated with that variable to obtain more accurate results (Grekousis, 2020). Global Moran's I was chosen because it is the most used index when dealing with data with spatial variation. It computes the spatial autocorrelation considering the local values of the factors and their locational attributes simultaneously (Moran, 1950).

The Global Moran index I obtains the spatial dependence from the relationship between the mean, the deviation from the mean, and the data variance. Therefore, when its p-value is less than 0.05, the index must be compared to the absolute value of its expected value. On the other hand, spatial dependence is observed when the index is greater in terms of absolute value (Grekousis, 2020). This index was calculated with the "spdep" package (Bivand et al., 2013) in the R environment.

Hypotheses H_B to H_G were tested by estimating a Geographically Weighted Regression (GWR) models. GWR is a type of local analysis; its technique is intuitive and is based on the structure of linear regression. However, unlike a linear regression model (global analysis), a local analysis allows the relationship between variables to be evaluated considering the spatial variation of the data, bringing valuable insights about those relationships that otherwise could be ignored (Fotheringham et al., 2002).

GWR defines spatial subunits that contain data called regression points. In this work, the regression points are the polygons that represent the neighbourhoods. The model also establishes optimal regions (bandwidths) around each regression point and gives different weights to data according to their distance from a regression point; data closer to a regression point have a higher weight than data farther from it. The bandwidths was identify by using a spatial kernel function (Fotheringham et al., 2002).

The GWR was estimating by using a Gaussian kernel function that allows continuous weighing. The estimated models assumed error distribution follows a Gaussian distribution. The cross-validation technique calibrated the spatial matrix. The GWR was estimated with the "spgwr" package (Bivand et al., 2013) in the R environment. See Fotheringham et al. (2002) for more information on GWR models.

Two GWR models were estimated: The model GWR1 considered all variables, i.e., retail shops, men, women, household size, income, population by age groups (18 and 19 years old, 20 to 29 years old, 30 to 39 years old, 40 to 49 years old, 50 to 59 years old, and 60 years or older), and population by race (white, mixed, black, Asian, and indigenous). The model GWR2 was estimated considering the minimization of the corrected Akaike Information Criterion (AICc) using the function "gwr.model.selection" from the "GWmodel" package (Gollini et al., 2015) in the R environment. This function uses an algorithm that determines all possible models with only one independent variable; then, the one with the lowest AICc is selected and the corresponding independent variable is permanently included in the model. After that, new models are estimated, containing all the possible combinations using the prior selected variable and the remaining independent variables; the model with the lowest AICc is selected. This process is repeated for all independent variables. The best and final model is the one with the lowest AICc (Lu et al., 2014).

The residuals of each model were tested for spatial autocorrelation using the Global Moran I index to verify the statistical significance of the GWR models. The residuals are recommended to present a random distribution in space (Fotheringham et al., 2002). The multicollinearity between the independent variables was measured using the Variance Inflation Factor (VIF), calculated with the "car" package (Fox and Weisberg, 2019) in the R environment. A VIF smaller than ten is desirable (James et al., 2015). Multicollinearity in GWR models is dealt with in the same way that in global models (Fotheringham and Oshan, 2016).

5. RESULTS AND DISCUSSION

The Global Moran's I index indicated the spatial dependence of the e-commerce deliveries, confirming hypothesis H_A. The index value was 0.080, higher in absolute value than the expected value (−0.002), and with a p-value of 0.004. Therefore, the use of spatial techniques to analyze the phenomenon becomes fundamental.

Figure 5 illustrates the estimated coefficients of the GWR1 model. Darker shades depict the highest estimated values of the coefficients, while lighter shades portray smaller estimated values of the coefficients in absolute values for both cases. In addition, blue shades depict negative spatial effects while red shades depict positive spatial effects. In general, the estimated coefficients show little spatial variability.

Results show that the retail shops positively affect the e-commerce deliveries, indicating that the neighborhoods with more shops receive more deliveries. Moreover, income positively affects e-commerce deliveries, i.e., high-income neighborhoods have more purchasing power. On the other hand, household size negatively influences deliveries; neighborhoods with smaller households receive more deliveries. The number of men and the number of women have a negative influence, with the coefficient for men being slightly higher, indicating that men tend to receive fewer deliveries than women. The variables in the age group showed a positive effect on e-commerce deliveries, with the coefficient for people between 20 and 29 years being slightly higher, indicating that this age group receives more deliveries than the others. Race-related variables positively influenced deliveries, with the coefficient for the Asian population being higher, indicating that this group tends to receive more deliveries than the others. The residuals of GWR1 have a random spatial distribution that validates the model.

Figure 6 shows the coefficients estimated for model GWR2, which included the retail shops, income, women, the population between 18 and 19 years old, the population between 20 and 29 years old, black population, and Asian people. The variability of the estimated coefficients is small, similar to that of GWR1. In model GWR2, the number of shops and the income positively influenced the e-commerce deliveries. On the other hand, women positively affected e-commerce deliveries. The age group 18 and 19 years negatively affected deliveries, while the population between 20 and 29 years had a positive effect. Regarding race-related variables, black people had a negative effect on deliveries, while the Asian population had a positive influence. The residuals of GWR2 show a random spatial distribution, validating the model.

5.1. Discussion

The results show the spatial dependence of urban deliveries, confirming the hypothesis H_A. Since spatial dependence was identified, analyzing e-commerce deliveries with techniques other than spatial techniques could result in errors. Furthermore, the estimated coefficients of the geographically weighted models show small spatial variability, which indicates that even with the space being heterogeneous, the combination of these factors presents a homogeneous pattern in the neighborhoods.

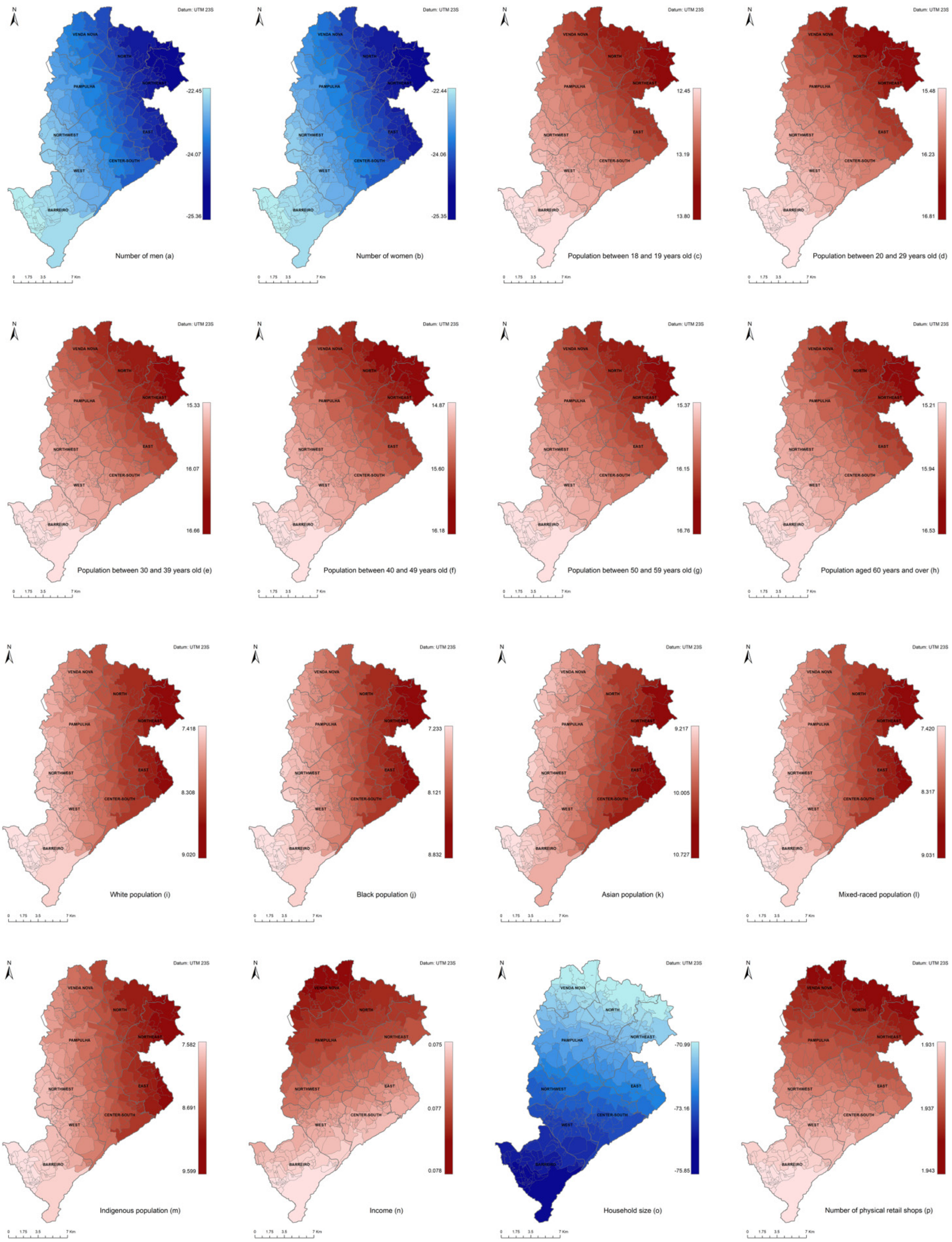

Figure 5. Spatial coefficients estimated for model GWR1: (a) men; (b) women; (c) population between 18 and 19 years; (d) population between 20 and 29 years; (e) population between 30 and 39 years; (f) population between 40 and 49 years; (g) population between 50 and 59 years; (h) population aged 60 years or over; (i) white population; (j) black population; (k) Asian population; (l) mixed population; (m) indigenous population; (n) income; (o) household size; and (p) physical retail shops.

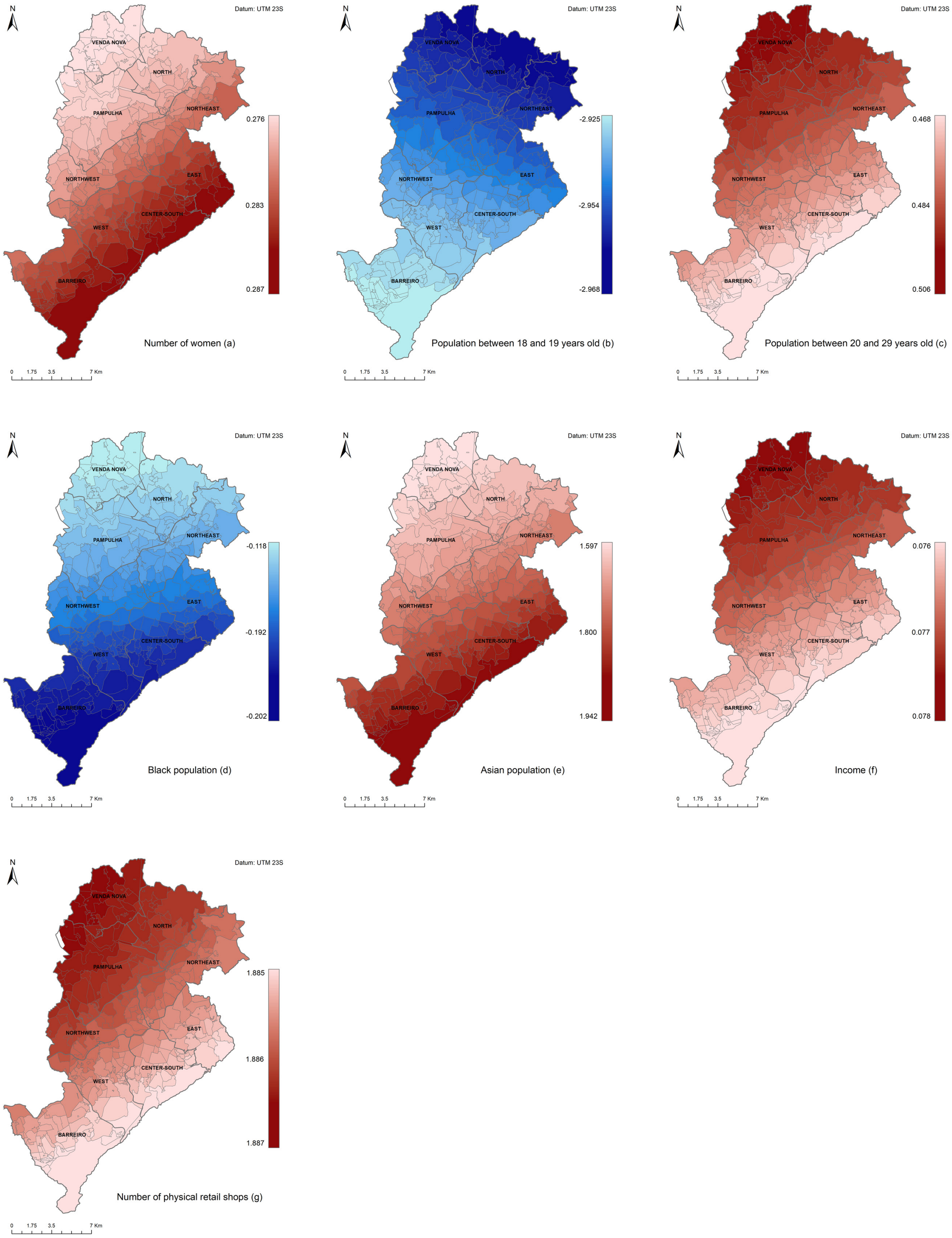

Figure 6. Spatial coefficients estimated for model GWR2: (a) women; (b) population between 18 and 19 years; (c) population between 20 and 29 years; (d) black population; (e) Asian population; (f) income; and (g) physical retail shops.

Table 2 summarizes the effects of models GWR1 and GWR2. Different effects results (positive and negative) for the same variable between models might be explained by the presence of multicollinearity (Hair et al., 2019). For example, in model GWR1, only shops, income, and household size had a VIF <10. In GWR 2, VIF > 10 for women, people between 18 and 19 years of age, and between 20 and 29 years of age. Gender was found to have a negative impact on e-commerce deliveries for model GWR1. On the other hand, a positive influence of the number of women was obtained in model GWR2, partially confirming hypothesis H_B. Furthermore, some scholars found a positive influence between women and e-commerce (Ren and Kwan, 2009; Sener and Reeder, 2012; Cao et al., 2013; Loo and Wang, 2018; Jaller and Pahwa, 2020; Saphores and Xu, 2021).

Table 2: Summary of the results

Variable	GWR1	GWR2
Men	Negative	
Women	Negative	Positive
Population between 18 and 19 years old	Positive	Negative
Population between 20 and 29 years of age	Positive	Positive
Population between 30 and 39 years of age	Positive	
Population between 40 and 49 years of age	Positive	
Population between 50 and 59 years old	Positive	
Population older than 60 years	Positive	
White population	Positive	
Black population	Positive	Negative
Asian population	Positive	Positive
Mixed-race population	Positive	
Indigenous population	Positive	
Income	Positive	Positive
Household Size	Negative	
Retail shops	Positive	Positive

Age is an important factor for e-commerce and is intrinsically linked to the type of product sold and the marketing channel. In the US, the 25-34 age group concentrates most on e-commerce consumers (Statista, 2021). In Europe, the 16-24 and 25-54 age groups have an equal share of e-commerce usage (Eurostat, 2022). Furthermore, those born after 1985 are people who use the Internet naturally and are more likely to use e-Commerce (Jongen, 2018). In Brazil, the age groups that most use e-commerce are people between 26 and 35 years of age and between 36 and 50 years of age (Statista, 2022). The positive effect for the 20 to 29 age group converges with the studies by Farag et al. (2006b) and Loo and Wang (2018). However, the two age groups of GWR2 had a heterogeneous influence on deliveries, partially confirming the hypothesis H_C.

Race is a recurrent variable in the literature. Brazil has a more egalitarian labor market, despite observing heterogeneity with respect to gender and remuneration (Silveira and Leão, 2020). Findings showed the Asian population had a positive influence on deliveries while the black people had a negative influence. Therefore, race has a heterogeneous effect on urban deliveries (different races have different results), which confirms the hypothesis H_D.

Income was found to have a positive influence on deliveries; neighborhoods with higher income tend to receive more e-commerce deliveries, confirming the hypothesis H_E. This finding converges with the literature (Farag et al., 2003, 2005, 2006a, b, 2007; Blasio, 2008; Sener and Reeder, 2012; Cao et al., 2013; Zhou and Wang, 2014; Clarke et al., 2015; Lee et al., 2015, 2017;

Motte-Baumvol et al., 2017; Beckers et al., 2018; Hood et al., 2020; Cheng et al., 2021; Saphores and Xu, 2021; Song, 2021). Although the spatial variation of the coefficients associated with income was significantly low, considering model GWR2, the influence of this variable is higher in the Venda Nova region. This region has a lower-income and other variables associated with income can influence the spatial distribution of e-commerce deliveries.

Estimated coefficients of household size negatively affect deliveries for model GWR1, confirming hypothesis H_F and converging with Zhou and Wang (2014). Finally, the positive impact of retail shops shows that neighborhoods with more stores tend to receive more deliveries, confirming hypothesis H_G. Farag et al. (2005) found similar results and stated that online and traditional shopping are complementary activities for urban regions. Moreover, those who live in urban areas have more access to traditional retail and tend to shop online more than residents of neighborhoods with lower access to shops (Farag et al., 2007; Blasio, 2008). In addition, Blasio (2008) suggested that this positive influence is related to the desire to see the product physically before buying it. The influence of this variable is more considerable in the Venda Nova region, which has retail stores, and in the Pampulha region. However, the spatial variation of the coefficients related to this variable is significantly low.

The findings can support freight policies to reduce the externalities of e-commerce deliveries. The spatial pattern of e-commerce deliveries shows that some neighborhoods have more deliveries than others. In addition, some socioeconomic characteristics of the neighborhoods have more influence than others. As e-commerce deliveries have an impact on urban mobility (cargo vehicles use the same road infrastructure as people), alternatives become essential to minimize activity-related externalities.

Identifying the pattern of e-commerce deliveries allows for finding the best strategies for locations with more deliveries, dedicated to the characteristics of consumers who live there. For example, delivery and collection systems are an alternative to home deliveries. However, this strategy has generally been implemented in central or residential regions (Rai et al., 2019). Beckers and Verhetsel (2021) state that delivery and collection systems are promising and more sustainable solutions for e-commerce deliveries. However, adopting this solution involves a paradigm shift concerning home deliveries (the traditional e-commerce delivery type) and developing an environmental awareness among e-consumers. In this way, knowing the neighborhoods with the highest concentration of e-commerce deliveries and the characteristics that are related to these deliveries allows a better allocation of delivery and collection systems.

A positive effect was identified to traditional retail, women, income, the population between 20 and 29 years old, and the Asian population. Therefore, neighborhoods with these characteristics could be potential places for delivery and collection points as an alternative to home deliveries. However, delivery and collection points in neighborhoods with predominance of large household size, population between 18 and 19 years old, and black population will not be largely used since these variables have a negative effect on e-commerce deliveries. However, more research is needed to assess the potential use of this alternative service by the population with the identified characteristics.

Additionally, the small variation indicates homogeneity of the effects of the variables in the territory. Thus, alternatives to reduce home deliveries can be implemented more seamlessly. These alternatives must be evaluated on a case-by-case basis, as they require investment in infrastructure or in partnership.

6. CONCLUSION

This article analyzed the spatial relationship between sociodemographic and retail access data and e-commerce deliveries in Belo Horizonte. The Global Moran's I index indicated spatial dependence of e-commerce deliveries. The estimated GWR models identified the variables that contribute most to spatial dependence.

The results allowed for answer to the research question proposed for the study: “To what extent do sociodemographic factors and access to traditional retail influence the spatial distribution of e-commerce deliveries?”. The positive effect of traditional retail, women, income, the population between 20 and 29 years old, and the Asian population was identified. On the other hand, household size, population between 18 and 19 years old, and black people negatively influence deliveries. Determining these effects made it possible to prove the hypotheses proposed for this research.

The results show that the spatial pattern of e-commerce deliveries in Belo Horizonte, located in a developing and Latin American country, is similar to that in China or even developed countries. Spatial analysis shows the homogeneity in the estimated parameters throughout the studied territory. However, other factors may still have spatial influence in addition to those considered in this study. Therefore, stated preference surveys could be used to understand the influence of those variables on consumer behavior. Furthermore, it is suggested to expand the analysis to 2020 and to analyze the effect of the Covid-19 pandemic on the spatial pattern of e-commerce deliveries.

ACKNOWLEDGEMENTS

The authors thank CNPq, CAPES, FAPEMIG, and FACEPE for supporting this research.

REFERENCES

Beckers, J.; I. Cárdenas and A. Verhetsel (2018) Identifying the geography of online shopping adoption in Belgium. *Journal of Retailing and Consumer Services*, v. 45, p. 33-41. DOI: 10.1016/j.jretconser.2018.08.006.

Beckers, J.; A. Verhetsel (2021) The sustainability of the urban layer of e-commerce deliveries: the Belgian collection and delivery point networks. *European Planning Studies*, v. 29, p. 2300-2319. DOI: 10.1080/09654313.2021.1921118.

Bivand, R.S.; E. Pebesma and V. Gomez-Rubio (2013) *Applied Spatial Data Analysis with R*. New York: Springer. DOI: 10.1007/978-1-4614-7618-4.

Blasio, G. (2008) Urban-rural differences in internet usage, e-commerce, and e-banking: evidence from Italy. *Growth and Change*, v. 39, n. 2, p. 341-367. DOI: 10.1111/j.1468-2257.2008.00422.x.

Cao, X.; Q. Chen and S. Choo (2013) Geographic distribution of e-shopping: application of structural equation models in the Twin Cities of Minnesota. *Transportation Research Record*, v. 2383, n. 1, p. 18-26. DOI: 10.3141/2383-03.

Cheng, C.; T. Sakai; A. Alho et al. (2021) Exploring the relationship between locational and household characteristics and e-commerce home delivery demand. *Logistics*, v. 5, n. 2, p. 29. DOI: 10.3390/logistics5020029.

Clarke, G.; C. Thompson and M. Birkin (2015) The emerging geography of e-commerce in British retailing. *Regional Studies, Regional Science*, v. 2, n. 1, p. 371-391. DOI: 10.1080/21681376.2015.1054420.

Eurostat (2022) E-commerce statistics for individuals. Available at: <https://ec.europa.eu/eurostat/statistics-explained/index.php?title=E-commerce_statistics_for_individuals> (accessed 04/04/2023).

Farag, S.; M. Dijst and M. Lanzendorf (2003) Exploring the use of e-shopping and its impact on personal travel behavior in The Netherlands. *Transportation Research Record*, v. 1858, n. 1, p. 47-54. DOI: 10.3141/1858-07.

Farag, S.; T. Schwanen and M. Dijst (2005) Empirical investigation of online searching and buying and their relationship to shopping trips. *Transportation Research Record*, v. 1926, n. 1, p. 242-251. DOI: 10.1177/0361198105192600128.

Farag, S.; T. Schwanen; M. Dijst et al. (2007) Shopping online and/or in-store? A structural equation model of the relationships between e-shopping and in-store shopping. *Transportation Research Part A, Policy and Practice*, v. 41, n. 2, p. 125-141. DOI: 10.1016/j.tra.2006.02.003.

Farag, S.; K.J. Krizek and M. Dijst (2006a) E-shopping and its relationship with in-store shopping: empirical evidence from The Netherlands and the USA. *Transport Reviews*, v. 26, n. 1, p. 43-61. DOI: 10.1080/01441640500158496.

Farag, S.; J. Weltevreden; T. van Rietbergen et al. (2006b) E-shopping in the Netherlands: does geography matter?. *Environment and Planning B: Planning and Design*, v. 33, n. 1, p. 59-74. DOI: 10.1068/b31083.

Fotheringham, A.S. and T.M. Oshan (2016) Geographically weighted regression and multicollinearity: dispelling the myth. *Journal of Geographical Systems*, v. 18, n. 4, p. 303-329. DOI: 10.1007/s10109-016-0239-5.

Fotheringham, A.S.; C. Brunsdon and M. Charlton (2002) *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*. West Sussex: John Wiley and Sons.

Fox, J. and S. Weisberg (2019) *An R Companion to Applied Regression*. Thousand Oaks: Sage.

Gollini, I.; B. Lu; M. Charlton et al. (2015) GWmodel: an R package for exploring spatial heterogeneity using geographically weighted models. *Journal of Statistical Software*, v. 63, n. 17, p. 1-50. DOI: 10.18637/jss.v063.i17.

Grekousis, G. (2020) *Spatial Analysis Methods and Practice: Describe – Explore – Explain through GIS*. Cambridge: Cambridge University Press. DOI: 10.1017/9781108614528.

Hair, J.F.; W.C. Black; B.J. Babin et al. (2019) *Multivariate Data Analysis*. Hampshire: Cengage.

Hood, N.; R. Urquhart; A. Newing et al. (2020) Sociodemographic and spatial disaggregation of e-commerce channel use in the grocery market in Great Britain. *Journal of Retailing and Consumer Services*, v. 55, p. 102076. DOI: 10.1016/j.jretconser.2020.102076.

IBGE (2020) IBGE Cidades – Belo Horizonte. Available at: <https://cidades.ibge.gov.br/brasil/mg/belo-horizonte/panorama> (accessed 06/22/2023).

Jaller, M. and A. Pahwa (2020) Evaluating the environmental impacts of online shopping: a behavioral and transportation approach. *Transportation Research Part D, Transport and Environment*, v. 80, p. 102223. DOI: 10.1016/j.trd.2020.102223.

James, G; D. Witten; T. Hastie et al. (2015) *An Introduction to Statistical Learning: With Applications in R*. New York: Springer.

Jongen, W. (2018) *The End of Online Shopping: The Future of Retail in an Always Connected World*. New Jersey: World Scientific Publishing Company. DOI: 10.1142/11106.

Kirby-Hawkins, E.; M. Birkin and G. Clarke (2019) An investigation into the geography of corporate e-commerce sales in the UK grocery market. *Environment and Planning. B, Urban Analytics and City Science*, v. 46, n. 6, p. 1148-1164. DOI: 10.1177/2399808318755147.

Krizek, K.J.; Y. Li and S.L. Handy (2005) Spatial attributes and patterns of use in household-related information and communications technology activity. *Transportation Research Record*, v. 1926, n. 1, p. 252-259. DOI: 10.1177/0361198105192600129.

Lee, R.J.; I.N. Sener and S.L. Handy (2015) Picture of online shoppers: specific focus on Davis, California. *Transportation Research Record*, v. 2496, n. 1, p. 55-63. DOI: 10.3141/2496-07.

Lee, R.J.; I.N. Sener; P.L. Mokhtarian et al. (2017) Relationships between the online and in-store shopping frequency of Davis, California residents. *Transportation Research Part A, Policy and Practice*, v. 100, p. 40-52. DOI: 10.1016/j.tra.2017.03.001.

Loo, B.P.Y. and B. Wang (2018) Factors associated with home-based e-working and e-shopping in Nanjing, China. *Transportation*, v. 45, n. 2, p. 365-384. DOI: 10.1007/s11116-017-9792-0.

Lu, B.; M. Charlton; P. Harris et al. (2014) Geographically weighted regression with a non-Euclidean distance metric: a case study using hedonic house price data. *International Journal of Geographical Information Science*, v. 28, n. 4, p. 660-691. DOI: 10.1080/13658816.2013.865739.

Moran, P.A.P. (1950) Notes on continuous stochastic phenomena. *Biometrika*, v. 37, n. 1-2, p. 17-23. DOI: 10.1093/biomet/37.1-2.17. PMid:15420245.

Motte-Baumvol, B.; L. Belton-Chevallier; L. Dablanc et al. (2017) Spatial dimensions of e-shopping in France. *Asian Transport Studies*, v. 4, n. 3, p. 585-600. DOI: 10.11175/eastsats.4.585.

Nielsen (2022) Webshoppers. Available at: <https://nielseniq.com/global/pt/landing-page/ebit/nielseniq-ebit-brasil/webshoppers/> (accessed 04/04/2023).

OECD (2020) E-commerce in the time of COVID-19. Available at: <https://www.oecd.org/coronavirus/policy-responses/e-commerce-in-the-time-of-covid-19-3a2b78e8/> (accessed 04/04/2023).

Rai, H.B.; S. Verlinde and C. Macharis (2019) The “next day, free delivery” myth unravelled. *International Journal of Retail & Distribution Management*, v. 47, n. 1, p. 39-54. DOI: 10.1108/IJRDM-06-2018-0104.

Ren, F. and M.-P. Kwan (2009) The impact of geographic context on e-shopping behavior. *Environment and Planning B: Urban Analytics and City Science*, v. 36, n. 2, p. 262-278. DOI: 10.1068/b34014t.

Saphores, J.-D. and L. Xu (2021) E-shopping changes and the state of e-grocery shopping in the US-evidence from national travel and time use surveys. *Research in Transportation Economics*, v. 87, p. 100864. DOI: 10.1016/j.retrec.2020.100864.

Sener, I.N. and P.R. Reeder (2012) An examination of behavioral linkages across ICT choice dimensions: copula modeling of telecommuting and teleshopping choice behavior. *Environment and Planning A: Economy and Space*, v. 44, n. 6, p. 1459-1478. DOI: 10.1068/a44436.

Silveira, L.S. and S.N. Leão (2020) O impacto da segregação ocupacional por gênero e raça na desigualdade de renda no Brasil em três décadas (1986-2015). *Revista Latinoamericana de Población*, v. 14, n. 27, p. 41-76. DOI: 10.31406/relap2020.v14.i12.n27.2.

Song, Z. (2021) The geography of online shopping in China and its key drivers. *Environment and Planning. B, Urban Analytics and City Science*, v. 49, n. 1, p. 259-274. DOI: 10.1177/23998083211002189.

Statista (2021) Distribution of digital buyers in the United States as of February 2020, by age group. Available at: <https://www.statista.com/statistics/469184/us-digital-buyer-share-age-group/> (accessed 04/04/2023).

Statista (2022) Distribution of digital buyers in Brazil in 2022, by age group. Available at: <https://www.statista.com/statistics/519746/age-distribution-of-online-shoppers-in-brazil/> (accessed 04/04/2023).

UNCTAD (2021a) *COVID-19 and E-commerce: a Global Review*. New York: United Nations. Available at: <https://unctad.org/system/files/official-document/dtlstict2020d13_en_0.pdf> (accessed 04/04/2023).

UNCTAD (2021b) *Estimates of Global E-commerce 2019 and Preliminary Assessment of COVID-19 Impac on Online Retail 2020. UNCTAD Technical Notes on ICT for Developmen Nº 118*. New York: United Nations. Available at: <https://unctad.org/system/files/official-document/tn_unctad_ict4d18_en.pdf> (accessed 04/04/2023).

Zhou, Y. and X.C. Wang (2014) Explore the relationship between online shopping and shopping trips: an analysis with the 2009 NHTS data. *Transportation Research Part A: Policy and Practice*, v. 70, p. 1-9. DOI: 10.1016/j.tra.2014.09.014.