Spatial mismatch, wages and unemployment in metropolitan areas in Brazil

Haddad, Eduardo Amaral; Barufi, Ana Maria Bonomi

Veröffentlichungsversion / Published Version
Zeitschriftenartikel / journal article

Empfohlene Zitierung / Suggested Citation:
Haddad, E. A., & Barufi, A. M. B. (2017). Spatial mismatch, wages and unemployment in metropolitan areas in Brazil. *Region: the journal of ERSA*, *4*(3), 175-200. https://doi.org/10.18335/region.v4i3.171

REGION

The Journal of ERSA
Powered by WU

Volume 4, Number 3, 2017, 175–200
DOI: 10.18335/region.v4i3.171

journal homepage: region.ersa.org

Spatial mismatch, wages and unemployment in metropolitan areas in Brazil

Ana Maria Bonomi Barufi[1], **Eduardo Amaral Haddad**[1]

[1] University of São Paulo, São Paulo, Brazil

Received: 25 October 2016/Accepted: 9 August 2017

Abstract. The spatial mismatch hypothesis states that a lack of connection to job opportunities may affect an individual's prospects in the labour market, especially for low-skilled workers. This phenomenon is especially observed in large urban areas, in which low-skilled minorities tend to live far away from jobs and face geographical barriers to finding and keeping jobs. This paper aims to investigate whether this negative relationship between spatial mismatch and labour market outcomes is valid in Brazil after controlling for individual characteristics. Our conclusions indicate that there is no clear relation between different measures of accessibility to jobs and the probability of being unemployed. However, for wages there is a clear correlation, which is stronger in larger metropolitan areas than in the country and has a more detrimental effect for low-skilled workers. This paper contributes to the literature by investigating the spatial mismatch in urban labour markets in Brazil. For the empirical literature in the country, this is an original contribution, as the comparison of intra-urban labour market dynamics of different urban areas provide a more comprehensive perspective of the role city size may play in local labour markets. Given the exploratory nature of this work, our results still rely on strong identification hypotheses to avoid potential bias related to simultaneous location decisions of workers and firms within the city. Even if these conditions do not hold, the results are still meaningful as they provide a better understanding of the conditional distribution of wages and the unemployment rate in the biggest metropolitan areas of Brazil.

JEL classification: R32, J64, J31

Key words: spatial mismatch, labour market, metropolitan areas.

1 Introduction

The spatial landscape of labour market opportunities varies significantly within an urban area. On average, the number of job openings and the wage level tend to decline as distance to the urban centre increases, which is usually modelled as a monocentric city. However, this relationship varies according to specific characteristics of each city, related to geography, amenities' distribution, sector composition and specialisation, transportation policies, the number of business centres (polycentric or monocentric city), among other factors (Capello 2007). Another important source of heterogeneity in the urban shape comes from the locational choices of firms in different sectors (McCann 2013), based on their cost-benefit analysis coming from the interaction between land and transportation costs, and the potential benefits that may arise from a more central location.

In some cases, jobs with better pay in the service sector can be concentrated near the centre, as they benefit more from knowledge spillovers that generate agglomeration externalities (Partridge et al. 2009). On the other hand, manufacturers started moving to the outskirts of the bigger cities in order to avoid high rents, effect that is widely acknowledged in the literature, together with additional impacts on the housing market (Lucas Jr, Rossi-Hansberg 2002). Furthermore, this relationship is said to be stronger for larger and denser areas, because congestion costs and the size of the urban sprawl lead to a higher cost of living in central areas. In this context, the spatial mismatch relates the structure of cities to unemployment and poverty (Gobillon, Selod 2013).

The urbanisation process in Brazil was fast in the second half of the twentieth century, as the urbanisation rate went from less than 50% to more than 80% in forty years. More than 90% of the GDP is created in cities (Da Mata et al. 2007). However, this process was not accompanied by a similar rise in the country's GDP per capita (Chauvin et al. 2016). Urban areas with less than 100,000 inhabitants made up 23% of the Brazilian population in 2010, while in the US, they housed 33% of the population.

Local labour markets are formed by the interaction of firms and workers with heterogeneous skills in various geographical locations, given the strong connection between housing and labour markets. Geographical location gives market power to firms over potential workers, especially over those residing close to them. In their model, Brueckner et al. (2002) define two different spaces (skills spaces and urban spaces), and in equilibrium low-skilled workers will be distant from firms in both of these spaces, providing a rationale for socioeconomic ghettos (Zenou 2009), consistent with the spatial mismatch hypothesis (Kain 1968). The main underlying mechanism of this model is the monopsonistic power of firms in the surroundings close to them, which depends on the elasticity of the firm's labour pool (which itself is negatively related to the costs of commuting and acquiring skills). Brueckner et al. (2002) show that workers will be separated in space by skill type, and firms set wages that exploit this separation in space. Low-skilled workers will then live far away from their jobs.

There are at least two main dimensions through which this intra-urban equilibrium in the labour market can be evaluated: unemployment and wages. According to Zenou (1999), urban efficiency wages may lead to involuntary unemployment, as they are set above the competitive equilibrium wage in order to induce workers not to shirk. Moreover, individuals living far away from jobs have poor information about job opportunities, which decreases their probability of finding a job. As a result, spatial mismatch is observed in large urban areas in which low-skilled minorities live far away from jobs and face geographical barriers to finding and keeping jobs. In addition to the spatial dimension, there is also a social separation faced by low-skilled workers and minorities (Zenou 2013), which reduces their chances of finding a job.

Based on this theoretical perspective, this paper provides a two-fold analysis of the relationship between spatial mismatch and labour market outcomes in large metropolitan areas in Brazil. This effect is calculated through the relationship between the average wage or the probability of being unemployed and distance to jobs (measured as the commuting time from home to work or the distance to the main business centre). This paper therefore contributes to the literature by investigating the spatial mismatch in urban labour markets in Brazil. For the empirical literature in the country, this is an original contribution, as the comparison intra-urban labour market dynamics of different urban areas provide a more comprehensive perspective of the role city size may play in local labour markets.

Moreover, it shows empirically that in the Brazilian case the spatial mismatch is more relevant in relation to individual wages, while the probability of being unemployed is not as regularly distributed in space. The latter result is an interesting contribution to the literature and may indicate that the probability of unemployment may not be the best measure to attest the effect of the spatial mismatch in the labour market. According to the literature, duration of unemployment may be a better fit for this role, a variable that is not available in the database considered in this study.

There is also an emphasis on how the spatial mismatch can be more harmful to low-skilled workers, a result that is in accordance to previous findings in the literature. In sum, city size and the capacity that individuals have to adapt and find job opportunities are relevant aspects to be considered to understand intra-urban labour market dynamics

The paper is structured as follows. Section 2 provides a brief literature review of spatial mismatch and local labour markets focusing on social interactions within the city. In Section 3, we describe the econometric strategy and the database, while in Section 4 we analyse the results. Concluding remarks follow in Section 5.

2 Spatial mismatch and labour market equilibrium

The intra-urban spatial distribution of economic agents and production inputs has been modelled as the result of location decisions made by workers and firms (Roback 1982). A wide range of factors, among whose there are agglomeration economies, may be included in different models, as indicated by the New Economic Geography and Urban Economics literatures (Fujita, Thisse 2012, Krugman 1995, Ottaviano 2004). The locational problem is usually analysed by evaluating how local prices (rents and wages) relate to the distance from the present location to the Central Business District (CBD) of the city (Lucas Jr, Rossi-Hansberg 2002). Distance to multiple tiers of the urban hierarchy within a city can also be relevant for this analysis (Partridge et al. 2009).

The concept of spatial mismatch dates back to the mid-1960s (Kain 1992). This concept appears as a possible partial explanation to racial conflicts and riots in the United States, with the identification of ghettos and unequal labour market outcomes. Low rates of employment and low wages for Afro-American workers could be related to limitations on residential choice and the distribution of jobs around the city. Among other dimensions, education, housing and employment reflect and reinforce the spatial mismatch in cities. There has been significant discussion on whether this hypothesis does explain inequalities in the city, given the variety of analytical methods, spatial mismatch measures and data aggregation levels.

More than that, there was considerable uncertainty about the magnitude of the effects of spatial mismatch in urban areas (Holzer 1991). Recently, Kain (2004) showed that the public education system of the United States reinforced the spatial mismatch, given that racial segregation resulted in the concentration of Black children in low-achieving schools. More recent developments combine the concept of spatial mismatch with the analysis of local prices within a city and the embedded location decisions of workers and firms. Spatial mismatch in the labour market means that people face spatial frictions when accessing jobs in metropolitan areas (Houston 2005a). This phenomenon relates to the way in which low-skilled minorities are affected by distance to job locations (Zenou 2009). The resulting distributions arise from the equilibria in the labour and the housing markets, which are simultaneously determined by the different decisions made by firms and workers.

The spatial mismatch hypothesis argues that low-skilled minorities face poor labour market outcomes because they are disconnected from job opportunities within the city (Gobillon et al. 2007). Even nowadays, this concept is still commonly used to investigate the case of afro-descendent population or other minorities in US cities, who often live far away from low-skilled jobs that are available in the suburbs of American cities (see for instance Ihlanfeldt 2006, Zenou 2009, Andersson et al. 2014).

The range of mechanisms underlying the theoretical frameworks that generate spatial mismatches are related either to the labour market itself or to the factors that potentially explain why minorities are physically disconnected from jobs (Gobillon, Selod 2013). According to Gobillon et al. (2007), these mechanisms can be analysed separately for workers and firms. From the workers' perspective, they are the following:

(i) long commuting may lead a worker to refuse a job opportunity after carrying out a cost-benefit analysis;

(ii) search efficiency may decrease with distance to jobs;

(iii) search intensity may also be affected by distance to jobs; and

(iv) high search costs may cause workers to restrict their search to a limited area.

From the firms' perspective, the main mechanisms are:

(v) stigma or prejudice may make firms discriminate against workers who live in certain locations;

(vi) employers may pay lower wages or refuse to hire workers who commute for long distances, as the commuting may decrease their productivity; and

(vii) employers may have a prejudice against specific workers because of the expected preferences of their customers.

As mentioned above, the spatial mismatch hypothesis is usually considered in the specific case of low-skilled minorities living in urban areas. However, the concept of 'spatial mismatch' in general terms is broadly used to investigate the disparity between locations of jobs and individuals that lead in an endogenous way to different levels of unemployment and wages across a city.

Among some of the theoretical models devoted to describing spatial mismatches in the urban environment, Zenou (2000) develops a model with endogenous city formation mechanisms that result in jobs concentrating in the CBD, employed individuals residing in the vicinity of the city centre, and the unemployed being further away from jobs. Urban unemployment will then be reinforced in the outskirts of the city, because the further away an individual is from jobs (which are concentrated in the CBD), the harder it is for her to find a job. Within a similar setting generated from a model based on a monocentric city combined with an efficiency wage mechanism and high reallocation costs, wages are expected to decrease with distance to the centre, as demonstrated by Zenou (2006).

It is important to note that the modelling of metropolitan labour markets can be significantly different for low-skilled and high-skilled workers, given the more limited distance that low-income individuals can commute. Thus, low-skilled workers will face a segmented urban labour market, while for high-skilled workers space is less restrictive. Unemployment for low-skilled workers will be associated with the lack of jobs in the areas close to their residence, while high-skilled workers will search for jobs in a wider spatial scale (Morrison 2005). Therefore, for high-skilled individuals, urban landscape is expected to have a smaller impact on their labour market outcomes. These two mechanisms can co-exist within the city to generate the observed distribution of unemployment rates.

One additional remark is that the literature of spatial mismatch is intrinsically related to spatial spillovers, social networks and proximity in different dimensions (Topa, Zenou 2015). Accessibility to jobs captures these effects just in a partial way, as it differentiates individuals by their reach to opportunities, and an interesting extension of research in this area should encompass these neighbouring relations in a more direct way.

Despite the large amount of empirical literature, Houston (2005b) argues that there is no clear consensus on the importance of the spatial mismatch in the explanation of labour market outcomes. Andersson et al. (2014) consider the duration of unemployment as a labour market outcome to measure the effects of spatial mismatch. They use a matched employer–employee database, and build person-specific measures of job accessibility with an empirical model of transport modal choice and network travel-time, finding that better job accessibility helps to decrease the duration of joblessness for lower-paid workers. Moreover, under-privileged groups are more affected by the lack of accessibility. The same dependent variable is employed by Rogers (1997), whose results indicate that unemployment duration in the Pittsburgh labour market area is influenced by residential location relative to employment opportunities, especially for less-educated individuals. According to Johnson (2006), the efficiency of job search is largely related to job accessibility. Then, 40% of the racial disparities in search duration is explained by spatial search-related variables.

The total number of jobs available in each region of the city and the impedance for reaching those regions can be used to define accessibility to jobs in a specific location. The impedance measure is usually defined either by the Euclidean distance or by commuting time between residential location and jobs, which may be derived from transport availability in each area of the city. The latter approach is followed by Vieira, Haddad (2015) for the São Paulo Metropolitan area, and they find indications that accessibility and income are strongly and positively related in the city. Di Paolo et al. (2016) find that car availability is relevant for job–education mismatch and that public transportation has an effect on better matching in the labour market for each schooling level.

Åslund et al. (2010) calculate the accessibility measure by considering the number of jobs and people of working age within a 5 kilometres-radius of the individual's residential location. They consider the exogenous allocation of refugees in Sweden ten years earlier and build an instrument that is based on how accessible jobs are to immigrants in their arrival year, and find a positive correlation between local job proximity and individual outcomes.

Job accessibility, demand and supply in the Chicago metropolitan area are used by Hu (2014) to find that socioeconomic restructuring (an increase in poverty and a reduction in relevant job opportunities) negatively affects poor job seekers, while spatial transformation (when jobs and job seekers move to the outskirts of the city) has a positive effect on their job prospects. The latter effect is caused by poorer individuals following jobs to suburban areas. With a similar empirical strategy, Hu, Giuliano (2014)'s results indicate that there is no relationship between spatial accessibility and the unequal employment status of the poor in the Los Angeles metropolitan area.

According to Tyndall (2015), public transportation has a causal and negative effect on neighbourhood unemployment rates, particularly for groups who are more dependent on this transport mode. The author explores a natural experiment from Hurricane Sandy, which exogenously reduced access to public transport in some neighbourhoods in New York City.

The empirical literature on spatial mismatch can be subdivided into two main strands: the first aims to understand the causes, while the second discusses the consequences of a spatial mismatch (Gobillon, Selod 2013). Houston (2005b) states that the consequences of a spatial mismatch are usually evaluated through an analysis of (i) residential segregation, (ii) comparisons of commuting times, (iii) comparisons of earnings, and (iv) measures of job proximity. Accordingly, Ihlanfeldt (2006) highlights the fact that the effects of spatial mismatch have been investigated on lower earnings, longer commutes and higher unemployment, especially in the case of black workers in the United States. Usually, employment and earnings equations include measures of local job opportunities, with a strategy based on a gravity model with a distance-decay function to take account of being further away from job opportunities.

Among the main econometric problems arising from this strategy there is the fact that residential location and the measurement of job opportunities are potentially endogenous (Ihlanfeldt 2006). Such endogeneity may appear through self-selection of more or less productive workers to specific areas, by the potential reverse causality of job opportunities and the probability of being unemployed, or through the simultaneous location decisions of firms and workers in a general equilibrium setting. One can deal with the simultaneity issue by including historical or geographical instruments that influenced the location of transportation infrastructure within a city without directly determining the location of workers and firms. This approach is explored by Haddad, Barufi (2017) for São Paulo Metropolitan Region with river shore access as an instrument, but is not replicable for the whole country as such detailed geographical information is not available yet in a larger scale.

Our identification strategy will be based in more restrictive hypotheses. In the short run, prices in the labour market are assumed to be close to the equilibrium level, and workers and firms are relatively immobile (Gibb et al. 2014). This endogeneity issue is then expected to be less relevant in the case of labour market outcomes. In addition, the measurement of local job opportunities can be indirect (using the assumption that there is a geographical centre in the city or by considering commuting time as a possible measure of the distance to jobs). The specific location of job opportunities is not included in the analysis, meaning that this endogeneity issue can be less relevant. In this study, we will assume that these aspects are able to soften such concerns. In any case, the potential direction of an endogeneity bias will be discussed in the following sections.

Usual measures of spatial mismatch may be problematic (Houston 2005b). On the one hand, long commutes may be a sign of either high mobility (highly paid workers) or a spatial mismatch between workers and jobs. On the other hand, different groups have specific propensities to commute, which means that studies usually measure commuting patterns of employed individuals, while spatial mismatch is generally concerned with the unemployed, who may behave differently. Houston (2005b) also suggests that job accessibility should take into account not only distance but also the amount of competition for the accessible jobs. Finally, total travel burden should take into account time, pecuniary costs and inconvenience (Bruzelius 1979). Commuting time, cost or distance are therefore, by themselves, incomplete measures. However, data availability restricts this analysis to such incomplete measures. We acknowledge this limitation and try to assess its potential impact in our results.

In summary, the empirical literature finds some mixed results, especially regarding the relationship between different measures of spatial mismatch and the unemployment rate. However, an increase in accessibility to jobs seems to improve labour market outcomes, especially for low-skilled minorities for whom the spatial mismatch is more relevant. There are significant empirical issues related to the estimation of this effect, whose consequences will be further discussed.

The next section presents our empirical strategy, which deals with comparisons of earnings and measures of job proximity (items (iii) and (iv) discussed above and listed by Houston (2005b)). In addition, we focus on the probability, for each economically active individual, of being unemployed, according to her residential location. To compare earnings, the unavailability of data means that we measure wages from a residential location perspective instead of a workplace basis, even if the latter would be a more appropriate approach (Houston 2005b).

3 Empirical strategy and data

The empirical strategy developed here is based on the estimation of the relationship between different measures of distance to jobs and labour market outcomes (earnings and the probability of being unemployed). All dependent variables are residence-based, due to data availability. Such strategy aims at exploring different dimensions of the spatial mismatch hypothesis in Brazilian metropolitan areas.

Estimations are conducted for individuals residing in a specific metropolitan area in order to capture the effect of each variable in relative terms within a specific urban structure. We assume that the wage equation can be written as follows:

$$w_i = \alpha + \beta X_i + \gamma_1 \text{inv\_dist}_r + \gamma_2 \text{inv\_dist}_r^2 + \epsilon_i \tag{1}$$

where w_i is the logarithm of the hourly wage measured for employed individuals who do not work at home, and X_i includes age, age squared, sector of activity, occupation, formalization status of the job, colour or race, education level, whether the individual is married, whether he or she has at least one child younger than fifteen living in the house, whether the house is owned by the family and whether the person is or is not the head of the household. In addition, inv_dist_r refers to the inverse of the Euclidean distance from the centroid of the weighting area to the main business centre[1]. This strategy is adopted since there is no data available to measure distance over each city's road infrastructure.

An alternative formulation for the reduced form presented in (1) is given by:

$$\begin{aligned} w_i \quad = \quad & \alpha + \beta X_i + \theta_1 \text{time\_commut\_6\_30}_i + \theta_2 \text{time\_commut\_31\_60}_i \\ & + \theta_3 \text{time\_commut\_61\_120}_i + \theta_4 \text{time\_commut\_121\_more}_i + \epsilon_i \end{aligned} \tag{2}$$

In this case, instead of the inverse distance to the centre, commuting time from home to work is used to evaluate the relationship between wages/productivity and the urban

[1]Under the simplifying assumption of a monocentric city, we consider the inverse distance from the weighting area where the individual lives to the main business centre of the metropolitan area, to calculate an approximate measure of distance to jobs.

landscape[2]. All these models are estimated with a simple OLS. Another dimension of spatial mismatch is the heterogeneity in the unemployment rates within the urban area. This dimension will be assessed by estimating the probability of being unemployed for each economically active individual, given her relative location to the main centre of the city:

$$h_i = P[U_i = 1] = F[\beta X_i + \gamma_1 \text{inv\_dist}_r + \gamma_2 \text{inv\_dist}_r^2] \tag{3}$$

In this specification, U_i refers to the employment status (it equals 1 when a person is unemployed) and F is a logistic cumulative probability function. Here, X_i is the set of observed characteristics for the individual (age, age squared, colour or race, education level, whether the individual is married, whether he or she has at least one child younger than fifteen living in the house, whether the house is owned by the family and whether the person is or is not the head of the household). Finally, β is a vector of parameters, and inv_dist_r is measured as before. An alternative formulation is the following:

$$\begin{aligned} h_i &= P[U_i = 1] \\ &= F[\beta X_i + \theta_1 \%\text{\_time\_commut\_6\_30}_r + \theta_2 \%\text{\_time\_commut\_31\_60}_r + \\ &\quad \theta_3 \%\text{\_time\_commut\_61\_more}_r] \end{aligned} \tag{4}$$

In this case, the spatial mismatch is approximated by the percentage of individuals in the neighbourhood whose time spent in commuting belongs to a particular time span.

In sum, two different measures of accessibility to jobs are considered here. Individuals in the Demographic Census are located in weighting areas, as it will be better explained below. Then, the first measure is based on the Euclidean distance of the centroid of the weighting area of residence to the business centre of each metropolitan area. This centre is equivalent to the geographic coordinates of the administrative centre of the municipality with the largest employment level of each metropolitan area.

The second accessibility measure is calculated through the commuting time spent from home to work. As a limitation, this variable is only available in categories (up to five minutes, from six to thirty minutes, thirty minutes to one hour, more than one hour to two hours, more than two hours). In the case of wage models, this variable is obtained through the individual's own reported commuting time. For the probability of unemployment, it is calculated by the percentage of workers who reside in each weighting area that are classified in each category and used in the regressions for the individuals living in that specific weighting area.

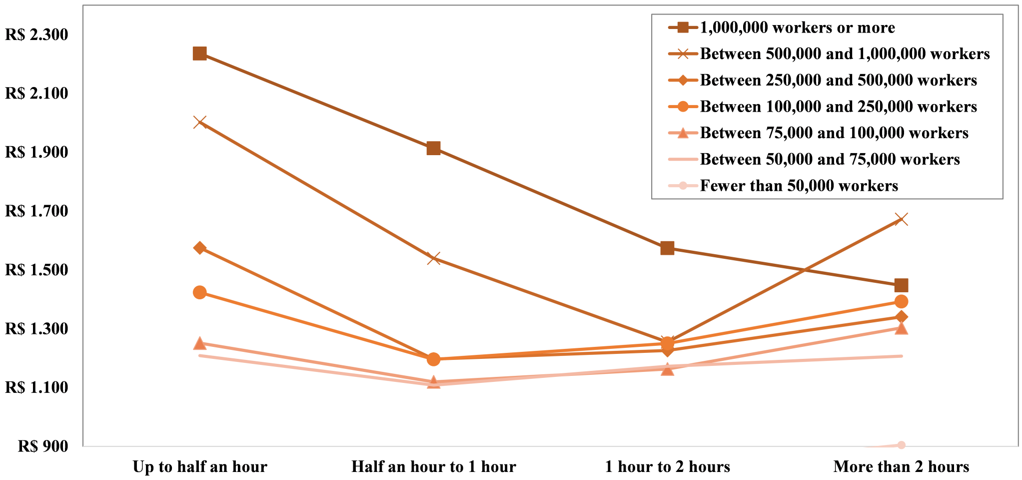
Apart from the whole database, these four models will be estimated for each metropolitan area and for three separate groups: (i) individuals who did not complete primary school[3], (ii) up to high school graduates without a college degree, and (iii) individuals who completed college education. In a country such as Brazil, inequality derived from the spatial mismatch can be more or less pronounced depending on the city size and the distance to the main concentration of job opportunities, and it may affect distinct skilled groups in different ways.

3.1 Database

The Brazilian Institute of Geography and Statistics (Instituto Brasileiro de Geografia e Estatística – IBGE) conducts a Demographic Census every ten years, with regional disaggregation at the municipal level (or at the census area level for bigger municipalities). The Demographic Census collects information on the main characteristics of individuals and households, providing details on the living conditions of the population in each municipality, and serving as a very important policy instrument in a country with a land area the size of Brazil. A shorter questionnaire applies to the whole population at the

[2]This impedance measure is the commuting time from home to work, calculated at the individual level for the wage equation or for the neighbourhood in the case of the estimation of unemployment probability. This second approach may be associated with a multicentric city structure. Census data only provides commuting time in categories, what represents an additional limitation of this analysis.

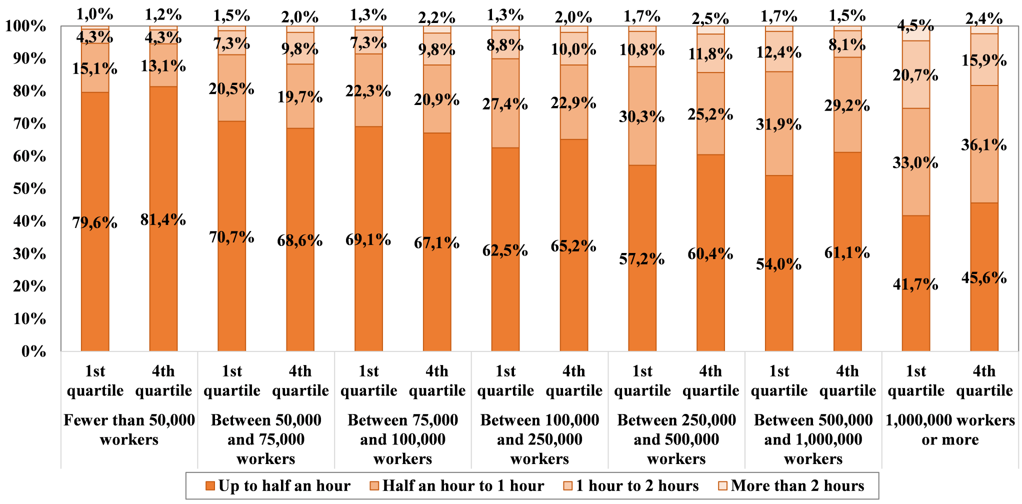
[3]Eight years of education.



Source: IBGE

Figure 1: Average wage of workers according to their commuting time from home to work and the size of the municipality of residence, 2010

census tract level, while specific individual characteristics are investigated in a longer set of questions that are given to a sample and are representative at the weighting areas level (conglomerates of census tracts with at least 400 households). Microdata at the individual level are available for this sample. We will use weighting areas as our definition of neighbourhood.

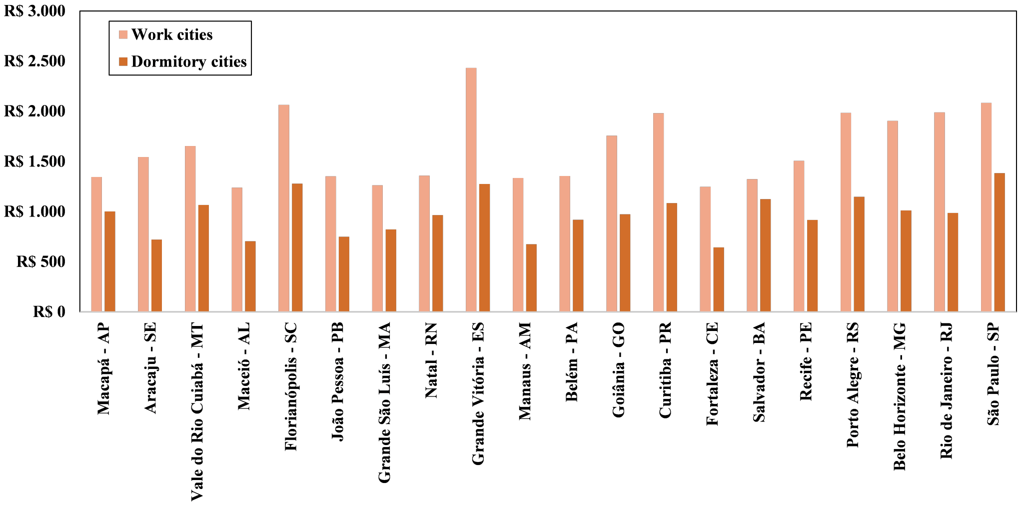


Source: IBGE

Figure 2: Distribution of workers who commute from home to work and belong to the 1st and the 4th quartile of the wage distribution, according to their commuting time and the size of the workforce in the municipality of residence, 2010

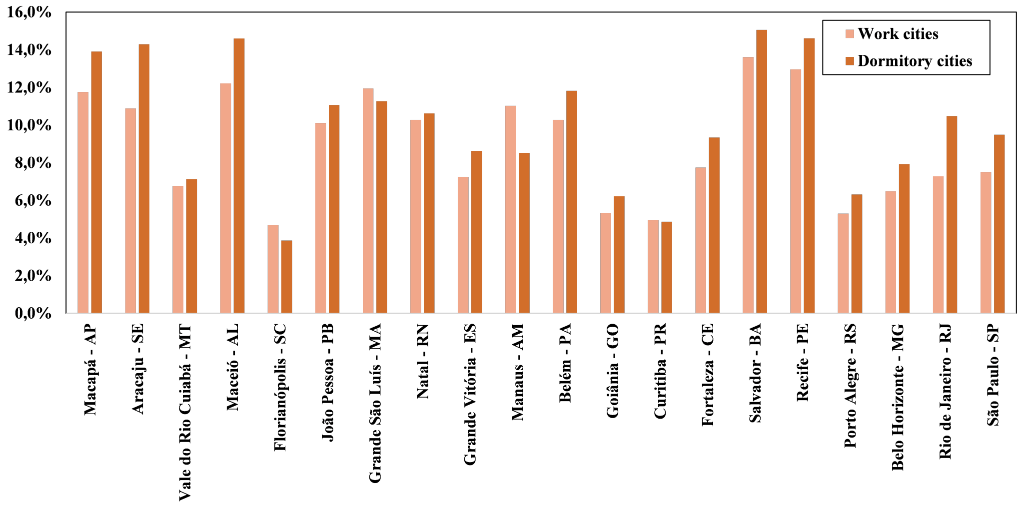
3.2 Descriptive statistics

The problem at hand is fundamentally related to metropolitan areas, as commuting costs and agglomeration economies become more relevant at a larger urban scale (Partridge et al. 2009). In fact, if one considers the average wage received by workers according to their commuting time from home to work, it is noticeable that the negative relationship between these two variables is clearer when cities with at least 500,000 workers are taken into account (Figure 1).



Source: IBGE

Figure 3: Average monthly wage for workers who live in work or dormitory cities inside each metropolitan area, (ordered by the size of working population), 2010



Source: IBGE

Figure 4: Average unemployment rate for people who live in work or dormitory cities inside each metropolitan area, (ordered by the size of working population), 2010

This difference between cities of different sizes is made clear in the analysis presented in Figure 2. In fact, the biggest differences in commuting times faced by workers in the richest (4^{th}) and the poorest (1^{st}) quartiles of the wage distribution in each municipality is seen in places with at least 500,000 workers. Furthermore, the decreasing relationship between wages and commuting time is stronger for those who commute for up to two hours.

For this reason, only 20 metropolitan areas containing state capitals were included in the study. In addition, only male workers aged 25 to 64 years old were kept in the database, in order to homogenise their decisions to participate in the labour market. For the wage regression, the database contained only workers who commuted to work and returned home every day.

It is also possible to show how wages and the unemployment rate vary according to the distance between the residential location of a worker and the centre of the city. Considering the daily commuting flows from home to work obtained from the Demographic

Table 1: Descriptive characteristics of each metropolitan area (ordered by the size of working age population), 2010

Metropolitan region	Macro region	Average hourly wage (R$ 2010)	Unem-ployment rate	Individuals commuting >1 hour (in %)	Working age population (men aged 25-64)
Macapá - AP	North	R$ 10.44	7.7%	5.3%	85,494
Aracaju - SE	Northeast	R$ 10.87	7.4%	10.7%	159,838
Vale do Rio Cuiabá - MT	Centre-West	R$ 13.58	4.3%	7.7%	160,638
Maceió - AL	Northeast	R$ 9.27	8.2%	13.3%	216,904
Florianópolis - SC	South	R$ 13.77	2.6%	6.6%	217,208
João Pessoa - PB	Northeast	R$ 9.72	6.5%	7.7%	230,930
Grande São Luís - MA	Northeast	R$ 10.96	7.5%	16.1%	244,017
Natal - RN	Northeast	R$ 9.85	7.1%	8.4%	258,207
Grande Vitória - ES	Southeast	R$ 11.94	4.9%	14.6%	353,561
Manaus - AM	North	R$ 11.19	7.1%	16.7%	378,496
Belém - PA	North	R$ 10.85	7.0%	14.4%	402,170
Goiânia - GO	Centre-West	R$ 12.32	3.4%	11.2%	415,541
Curitiba - PR	South	R$ 13.51	3.0%	13.1%	623,103
Fortaleza - CE	Northeast	R$ 9.41	5.6%	12.4%	666,504
Salvador - BA	Northeast	R$ 11.01	9.2%	20.0%	723,297
Recife - PE	Northeast	R$ 10.00	9.5%	17.2%	745,952
Porto Alegre - RS	South	R$ 12.38	3.7%	11.4%	807,268
Belo Horizonte - MG	Southeast	R$ 11.82	4.2%	18.7%	1,115,715
Rio de Janeiro - RJ	Southeast	R$ 12.92	5.8%	30.5%	2,402,075
São Paulo - SP	Southeast	R$ 15.37	5.7%	28.8%	3,953,270

Source: IBGE

Census of 2010, it is possible to define work and dormitory cities in each metropolitan area. The former are characterized by a higher inflow of people going there to work than an outflow of those who live there and go somewhere else to work, while the latter present a higher daily worker outflow than an inflow.

Figure 3 shows that average wages are much higher for people who live in work cities than for those who live in dormitory cities. However, in the case of the unemployment rate, there are mixed signs (Figure 4). In some metropolitan areas (Manaus, Grande São Luís, Florianópolis and Curitiba), dormitory cities show a lower unemployment rate than work cities. This pattern is unexpected under the hypothesis of a monocentric metropolitan area, but may be associated to the fact that these specific metropolitan areas are less dense than other more developed metropolitan areas in Brazil, for which the unemployment rate is larger in dormitory cities.

The econometric discussion outlined above explains the need to calculate the distance of each weighting area to the relevant business centre. This should be done on the basis of the main location of jobs around the city. In Brazil, however, there is no consolidated database covering all metropolitan areas and showing the location of jobs. Therefore, we consider a different approach, in which the centre of the metropolitan area is given by the administrative centre of the largest municipality (defined according to the number of employed individuals in 2010[4]).

Focusing more specifically on the models, the main descriptive characteristics are presented in Tables 1, 2 and 3, and Tables A.1, A.2 and A.3 in the Appendix. Table 1 indicates that the metropolitan areas considered in this study are significantly heterogeneous and should be treated separately, as each of them has a specific distribution of jobs and wages. Furthermore, areas with a bigger labour market have a higher average wage and a higher percentage of workers who commute for more than one hour to reach their jobs. This characteristic is clearer for metropolitan areas with more than a million male workers aged 25 to 64. For the unemployment rate, there seems to be more of a

[4]Data obtained from the Ministry of Labour and available at http://pdet.mte.gov.br/acesso-online-as-bases-de-dados.

Table 2: Percentage of workers who spend more than one hour commuting from home to work, according to the distance the worker lives from the centre, (ordered by the size of working population), 2010

	Distance from centre (in kilometer)							
	<2.5	2.5 to <5	5 to <10	10 to <20	20 to <30	30 to <40	40 to <50	50 or more
Macapá - AP	4.4%	4.9%	5.5%	6.3%		8.1%		4.0%
Aracaju - SE	7.0%	8.1%	11.4%	15.3%	15.9%			
Vale do Rio Cuiabá - MT	6.0%	5.0%	7.2%	10.2%	15.8%			8.7%
Maceió - AL	5.6%	6.9%	10.2%	20.4%	14.2%	11.9%		
Florianópolis - SC	3.5%	3.0%	5.2%	9.4%	12.5%	4.1%	4.9%	2.5%
João Pessoa - PB	6.7%	6.8%	8.0%	8.7%	6.0%	7.0%	10.7%	7.7%
Grande São Luís - MA	6.0%	11.2%	12.0%	23.2%	21.4%	14.5%		
Natal - RN	16.3%	13.4%	5.8%	5.4%	9.6%	8.3%	6.2%	
Grande Vitória - ES	7.4%	9.9%	16.9%	16.3%	22.2%	11.4%	6.1%	
Manaus - AM	10.6%	9.4%	12.0%	22.7%			20.0%	11.4%
Belém - PA	5.8%	6.7%	7.8%	19.4%	23.0%	11.2%	15.2%	
Goiânia - GO	3.6%	4.5%	6.6%	15.7%	21.2%	12.2%	11.5%	8.2%
Curitiba - PR	3.8%	4.2%	9.8%	16.5%	21.1%	6.9%	17.7%	8.7%
Fortaleza - CE	14.7%	13.2%	14.1%	12.6%	8.7%	8.0%	6.4%	5.3%
Salvador - BA	15.2%	13.8%	19.8%	29.4%	18.8%	9.0%	9.4%	10.7%
Recife - PE	7.4%	8.1%	13.1%	24.1%	20.1%	15.9%	7.2%	12.6%
Porto Alegre - RS	2.5%	5.1%	8.4%	15.7%	17.6%	7.8%	4.3%	4.7%
Belo Horizonte - MG	7.3%	10.9%	13.0%	24.6%	25.8%	17.5%	13.4%	6.3%
Rio de Janeiro - RJ	14.1%	13.5%	18.3%	28.1%	37.8%	39.7%	39.4%	22.8%
São Paulo - SP	23.9%	20.0%	21.1%	28.5%	34.7%	29.9%	21.5%	16.9%

Source: IBGE

regional aspect to the level observed in each metropolitan area, as regions located in the Northeast, for example, show a much higher level of unemployment than other regions.

There is a strong relationship between commuting time and distance to the centre, as can be seen in Table 2. In São Paulo and Rio de Janeiro, the largest metropolitan areas in Brazil, the percentage of individuals who commute for more than one hour is significantly higher for people who live more than 10km from the centre than for those living less than this distance away. However, this percentage decreases when the distance to the centre is greater than 30km in São Paulo or 40km in Rio de Janeiro. Since our objective is to investigate labour market characteristics related to the main business centre of each metropolitan area, we will focus on individuals living within a circle with a radius of 30km.

Table 3: Descriptive statistics by individual characteristics, 2010

	Unemp. Rate	Average hourly wage (R$ 2010)
Age		
25 to 34 years old	7.4%	9.75
35 to 44 years old	4.9%	12.31
45 to 54 years old	4.7%	15.48
55 to 64 years old	4.7%	19.06
Education level		
Less than 7 years of schooling	6.9%	6.62
8 to 10 years of schooling	6.3%	8.29
11 to 14 years of schooling	5.7%	11.21
15 years of schooling or more	3.0%	33.37

Continued on next page

Table 3 – continued from previous page

	Unemp. Rate	Average hourly wage (R$ 2010)
Colour		
White	4.8%	16.91
Black	6.8%	8.43
Yellow	5.1%	18.20
Brown	6.7%	8.84
Indigenous	6.4%	8.97
Marital status		
Single	7.7%	10.08
Married	3.7%	15.48
Children		
No children up to 15 years old	6.9%	13.07
Has at least one child up to 15 years old	4.1%	12.29
Home ownership		
Tenant	5.4%	11.39
Owned home	5.9%	13.21
Household position		
Another member of the household	8.0%	10.36
Head of the household	4.2%	14.33
Formality status		
Informal sector		9.46
Formal sector		13.93
Sector		
Agriculture		7.51
Manufacture and construction		9.59
Other industrial activities		14.26
Commerce		10.26
Services		10.53
Auxiliary services		17.83
Transport and communication		9.86
Health and social services		24.84
Education		17.85
Public sector		22.01
Other activities		15.79
Occupation		
Non-applicable		16.80
Leaders		30.45
Scientific, artistic or similar		30.88
Technical level		14.93
Administrative service		9.61
Commerce and service		7.14
Agriculture, livestock, extractive activities		4.14
Manufacture		7.26
Military		23.73
Commuting time to work		
Up to 5 minutes		13.66
6 to 30 minutes		13.47
31 minutes to 1 hour		12.68
More than 1 hour to 2 houres		11.27
More than 2 hours		11.15

Source: IBGE
Notes: The unemployment rate is calculated for the weighting area in which the individual resides

In Table 3, we can note that the wage level is higher for older individuals, those who are better educated, married people, those who are Indians, from Asiatic ancestry or white, those who are the head of a household, people employed in the formal sector and those who work in health and social services or leaders, scientists or artists. In addition, workers who commute for a longer time have a lower salary, on average. On the other hand, the unemployment rate is higher for younger individuals, those who are less educated, those who are black or brown, single people, people with no children, and those who are not heads of households.

The theory of spatial mismatch states that a lack of connection to job opportunities may affect an individual's prospects in the labour market, especially for low-skilled workers. Complementing the results presented in Table 3, Tables A.1, A.2 and A.3 provide wage levels and unemployment rates using different impedance measures. Distance to jobs can be calculated in many ways: (i) distance from the centroid of the weighting area to the business centre of the metropolitan area; (ii) individual commuting time from home to work; or (iii) percentage of workers in the weighting area whose commuting time falls within each time span. For the wage equation, we consider alternatively (i) and (ii) for employed individuals. On the other side, for the estimation of the probability of unemployment, (i) and (iii) are used, calculated at the weighting area level.

With these considerations in mind, Tables A.1, A.2 and A.3 show that wages seem to be higher near the centre of each metropolitan area, and that this effect is stronger in larger areas. However, for the unemployment rate, the expected positive relationship with distance to jobs is not clear. The main results will be presented in the next section.

4 Results

The first set of results refers to the estimation of wage equations that control for individual characteristics and uses two different measures of relative distance in the city: the distance to a unique centre (a monocentric city) and the distance to each worker's job (a multicentric city).

Table 4 shows that wages have a positive relationship with the inverse distance to the main centre of each metropolitan area (and, as a consequence, a negative relationship with distance itself). This effect is more significant for larger metropolitan areas, and it seems to be stronger for individuals with a higher education level. Therefore, wages are lower for individuals who live further away from the main business centre. However, this result demonstrates more of a correlation than a causal effect, especially because individuals are analysed with reference to their residential location. There may be inverse causality in this case, as an individual's choice of location may be affected by the wage previously received, and this can affect current labour market prospects and productivity.

Table 4: OLS regressions of the logarithm of the hourly wage, for all individuals and by education group

	Macapá - AP	Aracaju - SE	Vale do Rio Cuiabá - MT	Maceió - AL	Florianó-polis - SC	João Pessoa - PB	Grande São Luís - MA
All individuals							
Inverse of distance	0.244‡	0.935‡	1.323‡	0.439‡	0.281‡	0.004	0.122
Inverse of distance squared	-0.072	-0.855‡	-2.057‡	-0.064	-0.053	-0.004	-0.007
N	5,559	7,736	8,121	9,068	15,481	10,490	10,421
Adjusted R squared	0.429	0.463	0.364	0.455	0.421	0.44	0.354
Up to incomplete primary school							
Inverse of distance	0.349†	1.482‡	0.627	0.096	-0.308†	0.428‡	0.546‡
Inverse of distance squared	-0.126	-2.011‡	-0.878	0.234	0.213*	-0.118‡	-0.030‡
N	1,754	2,889	2,777	3,918	4,158	4,551	2,916
Adjusted R squared	0.134	0.121	0.088	0.103	0.081	0.116	0.081
Complete primary school to incomplete tertiary school							
Inverse of distance	0.215*	0.534*	0.902†	0.568‡	0.151*	0.008	-0.046
Inverse of distance squared	-0.058	-0.199	-1.604†	-0.244*	-0.033	-0.001	0.002

Continued on next page

Table 4 – continued from previous page

N	3,015	3,979	4,187	4,091	8,071	4,674	6,403
Adjusted R squared	0.321	0.32	0.2	0.276	0.23	0.243	0.197
Complete tertiary school							
Inverse of distance	0.246	1.458†	3.897‡	0.740‡	0.725‡	-1.237‡	-0.152
Inverse of distance squared	-0.083	-1.601†	-5.515‡	-0.257	-0.274‡	0.316‡	0.007
N	790	868	1,157	1,059	3,252	1,265	1,102
Adjusted R squared	0.293	0.283	0.196	0.296	0.3	0.298	0.237

	Natal - RN	Grande Vitória - ES	Manaus - AM	Belém - PA	Goiânia - GO	Curitiba - PR	Fortaleza - CE
All individuals							
Inverse of distance	-0.179	0.300‡	0.633‡	0.594‡	1.283‡	1.069‡	-0.355‡
Inverse of distance squared	-0.211	-0.344‡	-0.347‡	-0.258‡	-0.739‡	-0.302‡	0.224*
N	12,056	24,887	11,912	15,523	17,317	32,745	26,254
Adjusted R squared	0.46	0.431	0.34	0.374	0.361	0.389	0.408
Up to incomplete primary school							
Inverse of distance	0.365*	-0.143	1.001‡	0.461‡	1.252‡	0.764‡	-0.341†
Inverse of distance squared	-0.714†	0.253	-0.492‡	-0.214*	-0.812‡	-0.174‡	0.364*
N	4,512	7,780	3,541	4,933	6,466	10,605	9,248
Adjusted R squared	0.124	0.094	0.097	0.077	0.071	0.092	0.076
Complete primary school to incomplete tertiary school							
Inverse of distance	-0.613‡	0.202†	0.471‡	0.439‡	1.175‡	1.056‡	-0.023
Inverse of distance squared	0.278	-0.290†	-0.238‡	-0.148	-0.674‡	-0.293‡	-0.161
N	6,041	13,138	6,989	8,728	8,448	16,835	14,345
Adjusted R squared	0.268	0.216	0.182	0.19	0.215	0.19	0.213
Complete tertiary school							
Inverse of distance	0.334	1.460‡	0.445	1.134‡	1.255‡	1.038‡	-1.937‡
Inverse of distance squared	-1.546	-1.576‡	-0.333*	-0.646‡	-0.716‡	-0.295‡	1.361†
N	1,503	3,969	1,382	1,862	2,403	5,305	2,661
Adjusted R squared	0.268	0.283	0.216	0.256	0.249	0.254	0.269

	Recife - PE	Salvador - BA	Porto Alegre - RS	Belo Horizonte - MG	Rio de Janeiro - RJ	São Paulo - SP
All individuals						
Inverse of distance	0.758‡	0.538‡	1.251‡	1.267‡	0.617‡	1.399‡
Inverse of distance squared	-0.639‡	-0.202‡	-0.616‡	-0.681‡	-1.005‡	-1.107‡
N	33,635	25,865	35,715	47,034	79,277	154,088
Adjusted R squared	0.409	0.416	0.435	0.427	0.404	0.369
Up to incomplete primary school						
Inverse of distance	0.07	0.177	0.411*	0.362†	-0.013	1.359‡
Inverse of distance squared	-0.001	0.12	-0.118	0.045	0.036	-1.011‡
N	11,393	7,617	10,425	17,149	20,809	45,808
Adjusted R squared	0.093	0.088	0.091	0.079	0.073	0.08
Complete primary school to incomplete tertiary school						
Inverse of distance	0.607‡	0.379‡	1.234‡	1.101‡	1.066‡	1.650‡
Inverse of distance squared	-0.549‡	-0.112†	-0.575‡	-0.600‡	-1.686‡	-1.271‡
N	18,298	14,563	20,476	23,695	43,703	79,875
Adjusted R squared	0.203	0.217	0.233	0.216	0.186	0.18
Complete tertiary school						
Inverse of distance	1.776‡	1.202‡	1.392‡	1.821‡	1.165‡	0.705‡
Inverse of distance squared	-1.734‡	-0.552‡	-0.717‡	-1.079‡	-4.248‡	-0.641‡
N	3,944	3,685	4,814	6,190	14,765	28,405
Adjusted R squared	0.254	0.246	0.244	0.291	0.228	0.202

Source: Authors' calculations
Notes: Controls: age, age squared, colour or race, household head, with children up to 15 years old, married, sector of activity, occupation, existence of a formal contract. For the regressions with all individuals, the education attainment of the individual was included as an additional control. Significance levels: * $p<0.10$, †$p<0.05$, ‡$p<0.01$. Only male individuals aged 25 to 64 years old living within a distance of 30km from the centre are considered in the analysis. Sampling weights are taken into account with Stata command pweight. Complete tables can be requested from the authors.

This issue may also be present when the spatial mismatch is captured by each individual's commuting time from home to work (Table 5). The estimated coefficients are then likely to be underestimating the real effect. Therefore, if this reverse causality issue is correctly dealt with, distance to jobs should be even more relevant in determining wage levels, as it would be possible to discount the effect of relocation by looking at job opportunities over the city.

In any case, Table 5 shows that the negative effect of commuting time on wages is significant for workers commuting for 30 minutes or more, and is higher the longer the time spent in this activity. For low-skilled workers in smaller metropolitan areas, wages are not significantly correlated to this measure of spatial mismatch. Moreover, for most metropolitan areas, workers who commute for two hours or more do not see any significant effect on their wages, which may result from the fact that there are only a few workers belonging to this group, and no clear wage pattern.

The second set of results refers to the probability of being unemployed. Coefficients are presented as odds-ratios, with values greater than one indicating a positive effect of the variable of interest on the probability of unemployment. Tables 6 and 7 present the estimated coefficients related to specific distance measures. Metropolitan areas are ranked from left to right according to the size of their labour market. There is an indication in Table 6 that the probability of unemployment is not significantly correlated with the inverse distance to the centre. This result is consistent for most metropolitan areas, and there is no specific pattern for groups with different levels of schooling. The same result is found when distance to jobs is measured by the time spent by workers in the neighbourhood commuting from home to work (Table 7). Once again, for most metropolitan areas this relationship is not significant, and it does not show any pattern regarding education level, labour market size, or the sign of the correlation itself in cases when it is in fact significant.

Table 5: OLS regressions of the logarithm of the hourly wage, for all individuals and by education group

	Macapá - AP	Aracaju - SE	Vale do Rio Cuiabá - MT	Maceió - AL	Florianó-polis - SC	João Pessoa - PB	Grande São Luís - MA
All individuals							
Workers commuting 6'-30'	-0.093‡	0.002	-0.038	-0.029	-0.008	0.021	-0.013
Workers commuting >30'-1 hour	-0.100†	-0.058	-0.137‡	-0.017	-0.040*	-0.015	-0.022
Workers commuting >1-2 hours	-0.194‡	-0.069	-0.216‡	-0.085†	-0.098‡	-0.043	-0.106†
Workers commuting >2 hours	0.086	0.121	-0.046	0.007	0.081	0.002	-0.048
N	5,559	7,736	8,121	9,068	15,481	10,828	10,680
Adjusted R squared	0.429	0.461	0.366	0.446	0.418	0.442	0.356
Up to incomplete primary school							
Workers commuting 6'-30'	-0.035	0.018	0.029	0.008	0.001	0.048	-0.028
Workers commuting >30'-1 hour	-0.017	0.015	-0.033	0.050	0.033	0.084	0.008
Workers commuting >1-2 hours	-0.089	0.017	-0.129*	0.038	0.032	0.048	-0.054
Workers commuting >2 hours	0.062	0.016	-0.085	0.034	-0.136	-0.003	-0.038
N	1,754	2,889	2,777	3,918	4,158	4,804	3,017
Adjusted R squared	0.130	0.114	0.090	0.097	0.080	0.123	0.079
Complete primary school to high school graduates without college degree							
Workers commuting 6'-30'	-0.109†	-0.016	-0.056	-0.056	-0.016	-0.018	-0.029
Workers commuting >30'-1 hour	-0.084	-0.110*	-0.176‡	-0.078	-0.042	-0.105†	-0.014
Workers commuting >1-2 hours	-0.214‡	-0.163†	-0.211‡	-0.192‡	-0.096†	-0.113*	-0.121†
Workers commuting >2 hours	0.009	0.069	0.153	0.039	0.069	0.080	-0.022
N	3,015	3,979	4,187	4,091	8,071	4,749	6,550
Adjusted R squared	0.320	0.320	0.206	0.271	0.230	0.246	0.199
College degree							
Workers commuting 6'-30'	-0.171†	0.099	-0.058	0.041	-0.020	-0.051	0.063
Workers commuting >30'-1 hour	-0.367‡	0.057	-0.186	0.087	-0.116†	-0.115	-0.129
Workers commuting >1-2 hours	-0.452†	0.028	-0.293	-0.037	-0.286‡	-0.252	-0.192
Workers commuting >2 hours	0.104	0.579	-0.259	0.181	0.444†	-0.232	-0.128
N	790	868	1,157	1,059	3,252	1,275	1,113

Continued on next page

Table 5 – continued from previous page

Adjusted R squared	0.301	0.279	0.181	0.263	0.287	0.285	0.241

	Natal - RN	Grande Vitória - ES	Manaus - AM	Belém - PA	Goiânia - GO	Curitiba - PR	Forta-leza - CE
All individuals							
Workers commuting 6'-30'	-0.014	-0.027	0.029	-0.044	-0.018	0.011	-0.013
Workers commuting >30'-1 hour	-0.095‡	-0.075‡	-0.042	-0.066†	-0.089‡	-0.033*	-0.024
Workers commuting >1-2 hours	-0.109‡	-0.153‡	-0.121‡	-0.132‡	-0.193‡	-0.123‡	-0.112‡
Workers commuting >2 hours	0.078	-0.069*	-0.039	-0.015	-0.050	-0.102‡	-0.086*
N	12,056	24,887	11,912	15,523	16,951	32,523	27,034
Adjusted R squared	0.458	0.433	0.340	0.369	0.355	0.379	0.408
Up to incomplete primary school							
Workers commuting 6'-30'	-0.035	-0.095‡	0.045	-0.001	-0.005	0.006	0.031
Workers commuting >30'-1 hour	-0.023	-0.083†	0.048	-0.021	-0.028	0.014	0.091‡
Workers commuting >1-2 hours	-0.013	-0.125‡	-0.032	-0.106*	-0.122†	-0.018	0.038
Workers commuting >2 hours	-0.063	-0.054	-0.018	-0.077	0.017	-0.119†	-0.083
N	4,512	7,780	3,541	4,933	6,284	10,494	9,662
Adjusted R squared	0.122	0.096	0.089	0.074	0.062	0.088	0.078
Complete primary school to high school graduates without college degree							
Workers commuting 6'-30'	-0.002	0.004	-0.013	-0.061	-0.017	-0.032	-0.032
Workers commuting >30'-1 hour	-0.083†	-0.044	-0.119†	-0.049	-0.091†	-0.065‡	-0.052*
Workers commuting >1-2 hours	-0.162‡	-0.118‡	-0.172‡	-0.095†	-0.239‡	-0.176‡	-0.143‡
Workers commuting >2 hours	0.168	-0.041	-0.036	0.045	-0.105	-0.137†	-0.034
N	6,041	13,138	6,989	8,728	8,284	16,737	14,691
Adjusted R squared	0.262	0.219	0.186	0.184	0.210	0.180	0.214
College degree							
Workers commuting 6'-30'	-0.053	-0.059	0.143	-0.054	-0.060	0.128†	-0.137
Workers commuting >30'-1 hour	-0.359‡	-0.189‡	0.058	-0.223†	-0.227‡	-0.003	-0.289‡
Workers commuting >1-2 hours	-0.305†	-0.373‡	-0.217*	-0.360‡	-0.178	-0.249‡	-0.612‡
Workers commuting >2 hours	0.218	-0.187	0.004	-0.061	-0.135	0.323†	-0.323
N	1,503	3,969	1,382	1,862	2,383	5,292	2,681
Adjusted R squared	0.281	0.287	0.222	0.247	0.240	0.241	0.270

	Recife - PE	Salvador - BA	Porto Alegre - RS	Belo Horizonte - MG	Rio de Janeiro - RJ	São Paulo - SP
All individuals						
Workers commuting 6'-30'	-0.016	0.040	0.005	-0.013	-0.011	-0.018
Workers commuting >30'-1 hour	0.010	0.051†	-0.009	-0.048‡	-0.002	-0.014
Workers commuting >1-2 hours	-0.051†	0.042	-0.044†	-0.125‡	-0.031†	-0.067‡
Workers commuting >2 hours	-0.022	0.085†	-0.062	-0.134‡	-0.049†	-0.095‡
N	33,852	27,923	42,000	48,518	83,302	154,584
Adjusted R squared	0.406	0.409	0.424	0.419	0.397	0.367
Up to incomplete primary school						
Workers commuting 6'-30'	-0.050	0.020	-0.037	-0.068†	-0.082‡	-0.046†
Workers commuting >30'-1 hour	0.001	0.067	0.004	-0.035	-0.025	-0.012
Workers commuting >1-2 hours	-0.026	0.022	-0.028	-0.091‡	-0.030	-0.041†
Workers commuting >2 hours	-0.063	0.091	-0.110†	-0.092†	0.005	-0.047†
N	11,485	8,451	13,073	17,989	22,455	45,652
Adjusted R squared	0.094	0.084	0.092	0.078	0.072	0.077
Complete primary school to high school graduates without college degree						
Workers commuting 6'-30'	-0.010	0.018	0.004	-0.010	-0.017	-0.022
Workers commuting >30'-1 hour	0.023	0.031	-0.034	-0.058†	0.002	-0.030*
Workers commuting >1-2 hours	-0.040	0.038	-0.053†	-0.144‡	-0.032	-0.081‡
Workers commuting >2 hours	0.046	0.055	-0.032	-0.182‡	-0.075‡	-0.110‡
N	18,418	15,722	23,617	24,277	45,919	80,089
Adjusted R squared	0.201	0.211	0.222	0.212	0.181	0.175
College degree						
Workers commuting 6'-30'	0.072	0.150*	0.078	0.064	0.108†	0.022
Workers commuting >30'-1 hour	0.060	0.052	0.054	-0.020	0.025	0.026

Continued on next page

Table 5 – continued from previous page

Workers commuting >1-2 hours	-0.086	0.058	-0.067	-0.131†	-0.011	-0.053*
Workers commuting >2 hours	-0.162	0.173	0.034	0.007	-0.049	-0.162‡
N	3,949	3,750	5,310	6,252	14,928	28,843
Adjusted R squared	0.245	0.236	0.229	0.257	0.228	0.204

Source: Authors' calculations
Notes: Controls: age, age squared, colour or race, household head, with children up to 15 years old, married, sector of activity, occupation, existence of a formal contract. For the regressions with all individuals, the education attainment of the individual was included as an additional control. Reference category: workers commuting for up to 5 minutes. Significance levels: * p<0.10, †p<0.05, ‡p<0.01. Only male individuals aged 25 to 64 years old living within a distance of 30km from the centre are considered in the analysis. Sampling weights are taken into account with Stata command pweight. Complete tables can be requested from the authors.

Table 6: Logit model for the probability of being unemployed, regressions with all individuals and by education groups

	Macapá - AP	Aracaju - SE	Vale do Rio Cuiabá - MT	Maceió - AL	Florianó-polis - SC	João Pessoa - PB	Grande São Luís - MA
All individuals							
Inverse of distance	1.175	0.464	0.119	0.432	2.925†	1.410	1.713
Inverse of distance squared	0.944	6.556	31.848	1.390	0.455*	0.943	0.970
N	6,034	8,459	8,525	10,020	16,009	11,378	11,291
Pseudo R squared	0.055	0.064	0.038	0.053	0.051	0.067	0.061
Up to incomplete primary school							
Inverse of distance	0.322	0.192	0.003*	0.310	0.112	1.502	1.176
Inverse of distance squared	2.335	36.753	37,017.1*	1.407	4.366	0.912	0.992
N	1,917	3,223	2,952	4,448	4,320	5,083	3,186
Pseudo R squared	0.026	0.048	0.019	0.024	0.053	0.037	0.030
Complete primary school to incomplete tertiary school							
Inverse of distance	2.355	0.621	2.430	0.496	11.255	1.703	2.065*
Inverse of distance squared	0.584	2.704	0.058	1.339	0.138	0.889	0.960*
N	3,300	4,343	4,385	4,464	8,357	4,991	6,963
Pseudo R squared	0.064	0.047	0.038	0.056	0.062	0.068	0.061
Complete tertiary school							
Inverse of distance	0.531	98.561	0.620	1.195	2.783	0.172	1.181
Inverse of distance squared	1.648	0.061	37.412	0.749	0.484	1.852	0.992
N	817	893	1,188	1,108	3,332	1,304	1,142
Pseudo R squared	0.131	0.245	0.104	0.152	0.063	0.187	0.163

	Natal - RN	Grande Vitória - ES	Manaus - AM	Belém - PA	Goiânia - GO	Curitiba - PR	Forta-leza - CE
All individuals							
Inverse of distance	2.487	0.846	0.468	0.697	1.354	1.204	1.151
Inverse of distance squared	0.276	1.252	1.441	1.528	0.782	0.962	1.039
N	13,086	26,231	12,933	16,838	18,004	33,821	27,974
Pseudo R squared	0.053	0.035	0.035	0.038	0.030	0.026	0.046
Up to incomplete primary school							
Inverse of distance	2.789	0.124†	0.176*	0.257*	0.204	1.108	2.385
Inverse of distance squared	0.388	9.381*	2.620	3.555†	4.159	1.161	0.397
N	5,027	8,268	3,916	5,409	6,768	10,976	9,953
Pseudo R squared	0.038	0.026	0.016	0.021	0.030	0.022	0.026
Complete primary school to incomplete tertiary school							
Inverse of distance	1.332	1.989	0.734	1.010	4.190	3.190†	0.646
Inverse of distance squared	0.428	0.473	1.091	1.089	0.286*	0.706*	2.237
N	6,495	13,860	7,583	9,494	8,772	17,410	15,259
Pseudo R squared	0.050	0.031	0.033	0.038	0.032	0.026	0.051

Continued on next page

Table 6 – continued from previous page

Complete tertiary school							
Inverse of distance	1662.660*	3.270	0.542	1.987	0.448	0.117†	1.921
Inverse of distance squared	0.000	0.399	1.265	0.770	1.824	1.817†	0.278
N	1,564	4,103	1,434	1,935	2,464	5,435	2,762
Pseudo R squared	0.134	0.082	0.084	0.060	0.055	0.041	0.096

	Recife - PE	Salvador - BA	Porto Alegre - RS	Belo Horizonte - MG	Rio de Janeiro - RJ	São Paulo - SP
All individuals						
Inverse of distance	0.430†	0.411‡	2.409*	1.007	0.736	1.253
Inverse of distance squared	1.629	1.604†	0.694	0.848	1.176	0.654
N	37,419	28,533	37,203	49,194	84,152	164,255
Pseudo R squared	0.057	0.053	0.027	0.031	0.042	0.032
Up to incomplete primary school						
Inverse of distance	0.536	0.550	13.921*	0.569	0.275*	6.053‡
Inverse of distance squared	1.459	1.201	0.132‡	1.180	4.335	0.183‡
N	13,098	8,711	10,940	18,033	22,383	49,540
Pseudo R squared	0.027	0.023	0.027	0.022	0.018	0.016
Complete primary school to incomplete tertiary school						
Inverse of distance	0.342†	0.393†	2.030	0.865	0.877	0.644
Inverse of distance squared	1.732	1.722†	0.800	1.476	0.965	0.894
N	20,214	15,990	21,323	24,813	46,498	85,260
Pseudo R squared	0.053	0.050	0.024	0.031	0.041	0.029
Complete tertiary school						
Inverse of distance	1.012	0.143	0.803	1.430	54.645*	0.669
Inverse of distance squared	0.979	2.427	1.170	0.388	0.000	1.614
N	4,107	3,832	4,940	6,348	15,271	29,455
Pseudo R squared	0.090	0.080	0.028	0.032	0.055	0.028

Source: Authors' calculations
Notes: Controls: age, age squared, colour or race, household head, with children up to 15 years old, married. For the regressions with all individuals, the education attainment of the individual was included as an additional control. Coefficients are presented as odds-ratios. Significance levels: * $p<0.10$, †$p<0.05$, ‡$p<0.01$. Only male individuals aged 25 to 64 years old living within a distance of 30km from the centre are considered in the analysis. Sampling weights are taken into account with Stata command pweight. Complete tables can be requested from the authors.

A few aspects can be highlighted in relation to these results. On the one hand, unemployment levels may vary throughout the city in an irregular way, with no specific pattern in either monocentric or multicentric cities. In a sense, this conclusion in the Brazilian case matches part of the literature, which finds no regular pattern for the spatial distribution of the unemployment rate.

However, the conclusion goes against recent theoretical predictions that distance to jobs can affect the probability that individuals belonging to low-skilled minorities find a position. If these theoretical predictions are valid, it might be that there are methodological issues driving this unexpected result. First, distance is not measured in relation to an individual, but relates only to her neighbourhood. In addition, we do not take into account the location of job offers and existing jobs. Our database locates individuals by their place of residence. Therefore, there may be difficulties in correctly identifying the centres in the city and in calculating the relative location of each potential worker. Moreover, when distance is measured as the commuting time for workers in the neighbourhood, this may not be the same as the commuting time a potential worker would spend if he or she were in work.

Table 7: Logit model for the probability of being unemployed, regressions with all individuals and by education groups

	Macapá - AP	Aracaju - SE	Vale do Rio Cuiabá - MT	Maceió - AL	Florianó-polis - SC	João Pessoa - PB	Grande São Luís - MA
All individuals							
% workers commuting 6'-30'	0.048†	0.664	-0.091	0.293	0.327	0.212	0.507
% workers commuting >30'-1 hour	0.105	0.288	2,219	0.175	0.049†	0.475	0.168†
% workers commuting >1 hour	0.516	0.929	2.128*	1,205	0.428	1,267	0.723
N	6,034	8,459	8,525	10,020	16,009	11,782	11,566
Adjusted R squared	0.056	0.063	0.040	0.053	0.054	0.064	0.061
Up to incomplete primary school							
% workers commuting 6'-30'	0.011*	3,231	-0.161	1,304	0.004†	0.151	1,532
% workers commuting >30'-1 hour	0.005*	0.995	5,118	1,424	0.021*	0.159	0.546
% workers commuting >1 hour	1,398	1,639	2,514	1,715	0.006‡	0.898	1,071
N	1,917	3,223	2,952	4,448	4,320	5,389	3,291
Adjusted R squared	0.026	0.046	0.019	0.025	0.059	0.035	0.030
Complete primary school to high school graduates without college degree							
% workers commuting 6'-30'	0.101	0.299	0.040	0.047†	2,798	0.799	0.291
% workers commuting >30'-1 hour	0.336	0.160	3,518	0.018†	0.078	3,474	0.081†
% workers commuting >1 hour	0.349	0.884	2,588	0.631	1,368	3,524	0.570†
N	3,300	4,343	4,385	4,464	8,357	5,079	7,122
Adjusted R squared	0.064	0.047	0.039	0.058	0.067	0.067	0.063
College degree							
% workers commuting 6'-30'	0	0.012	-0.301	0.337	0.210	0.000	2,668
% workers commuting >30'-1 hour	0.056	0.005	0.031	0.010	0.061	0.002	0.574
% workers commuting >1 hour	0.158	0.003†	0.240	26.935*	4,475	0.002	0.999
N	817	893	1,188	1,108	3,332	1,314	1,153
Adjusted R squared	0.139	0.260	0.102	0.177	0.064	0.162	0.164

	Natal - RN	Grande Vitória - ES	Manaus - AM	Belém - PA	Goiânia - GO	Curitiba - PR	Forta-leza - CE
All individuals							
% workers commuting 6'-30'	0.196*	3,722	0.322	0.869	0.274	1,508	8.937*
% workers commuting >30'-1 hour	0.111†	5,284	5,128	0.357	1,030	1,296	5,723
% workers commuting >1 hour	0.790	1.715*	0.847	2,357	0.756	1,576	4,325
N	13,086	26,231	12,933	16,838	17,626	33,594	28,821
Adjusted R squared	0.054	0.035	0.037	0.039	0.032	0.026	0.046
Up to incomplete primary school							
% workers commuting 6'-30'	0.126	1,701	0.493	36,757	0.210	0.816	2,237
% workers commuting >30'-1 hour	0.257	2,175	5,536	5,896	0.698	0.868	0.353
% workers commuting >1 hour	0.705	1,634	1,028	132.388*	1,024	1,929	3,518
N	5,027	8,268	3,916	5,409	6,579	10,862	10,416
Adjusted R squared	0.038	0.025	0.016	0.022	0.033	0.022	0.028
Complete primary school to high school graduates without college degree							
% workers commuting 6'-30'	0.155	13.582†	0.166	0.102	0.496	2,746	113.524‡
% workers commuting >30'-1 hour	0.038†	14.542†	4,042	0.074	1,371	2,788	110.358‡
% workers commuting >1 hour	0.711	2.230*	0.657	0.223	0.741	1,232	19.355†
N	6,495	13,860	7,583	9,494	8,604	17,310	15,623
Adjusted R squared	0.051	0.032	0.036	0.039	0.032	0.025	0.051
College degree							
% workers commuting 6'-30'	0.343	0.047	13,079	0.266	0.234	0.305	0.042
% workers commuting >30'-1 hour	0.493	1,331	901,885	0.139	10,454	0.064	4,959
% workers commuting >1 hour	0.524	0.180	1,656	0.878	0.071†	3,154	0.004
N	1,564	4,103	1,434	1,935	2,443	5,422	2,782
Adjusted R squared	0.128	0.086	0.086	0.059	0.065	0.038	0.102

Continued on next page

Table 7 – continued from previous page

	Recife - PE	Salvador - BA	Porto Alegre - RS	Belo Horizonte - MG	Rio de Janeiro - RJ	São Paulo - SP
All individuals						
% workers commuting 6'-30'	0.545	0.503	0.360	0.487	1,264	0.679
% workers commuting >30'-1 hour	0.982	0.264	0.433	0.901	0.909	0.771
% workers commuting >1 hour	1,370	1,047	0.376†	0.733	1,158	1,020
N	37,669	30,873	43,722	50,753	88,531	164,684
Adjusted R squared	0.057	0.051	0.024	0.031	0.042	0.032
Up to incomplete primary school						
% workers commuting 6'-30'	0.273	1,333	1,986	0.278	0.935	0.628
% workers commuting >30'-1 hour	0.609	0.358	1,593	1,120	0.696	0.779
% workers commuting >1 hour	0.699	2,183	1,376	0.738	1,321	0.981
N	13,207	9,678	13,696	18,924	24,195	49,331
Adjusted R squared	0.027	0.023	0.022	0.024	0.019	0.016
Complete primary school to high school graduates without college degree						
% workers commuting 6'-30'	0.670	0.404	0.067†	0.358	1,048	0.902
% workers commuting >30'-1 hour	1,243	0.249	0.111†	0.396	0.948	0.873
% workers commuting >1 hour	1,854	0.837	0.121‡	0.526	1,049	1,094
N	20,350	17,292	24,572	25,418	48,893	85,450
Adjusted R squared	0.053	0.048	0.023	0.031	0.042	0.029
College degree						
% workers commuting 6'-30'	6,864	0.015	17,634	12,783	0.689	0.337
% workers commuting >30'-1 hour	1,456	1,056	18,005	55,136	0.345	0.512
% workers commuting >1 hour	7,570	0.148	2,771	2,117	0.523	0.915
N	4,112	3,903	5,454	6,411	15,443	29,903
Adjusted R squared	0.091	0.080	0.024	0.030	0.058	0.029

Source: Authors' calculations
Notes: Controls: age, age squared, colour or race, household head, with children up to 15 years old, married. For the regressions with all individuals, the education attainment of the individual was included as an additional control. Coefficients are presented as odds-ratios. Significance levels: * p<0.10, †p<0.05, ‡p<0.01. Only male individuals aged 25 to 64 years old living within a distance of 30km from the centre are considered in the analysis. Sampling weights are taken into account with Stata command pweight. Complete tables can be requested from the authors.

As robustness checks, some additional results are provided in Table A.4, in the appendix[5]. We run the models for all individuals without dividing the database between metropolitan areas. Then, we also include in these models the control for the metropolitan area of residence. As it can be seen, the regression for the logarithm of the hourly wage against distance measures indicate that the farther away from the city centre the lower the wage received on average. A higher education attainment is associated to higher wages, which are also present in some of the largest metropolitan areas. However, when Model 2 is considered, the partial correlation of commuting distance and wages does not have the expected sign (longer commuting should be associated with lower wages). This is due to the fact that longer time periods are more common in larger urban areas, which are associated to more populated metropolitan areas. This is an indication that there are iterative effects of commuting distance and city size which should be controlled for (what is done in the estimations presented in Table 5).

For the models related to the probability of unemployment, the results indicate that a lower chance of unemployment is associated with longer commuting times in the weighting area (once again, an unexpected result). This is most likely caused by the fact that iterations between metropolitan areas and commuting times are not taken into account. In addition, the probability of unemployment is lower for more educated individuals. These considerations make our previous estimations preferable in relation to this additional exercise.

[5]We thank the contribution of an anonymous referee who suggested that we estimated these alternative models to enrich our analysis.

5 Final remarks

There is significant spatial mismatch in the labour market in Brazilian metropolitan areas. The influence of spatial location and distance to jobs on labour market outcomes is stronger for larger urban areas, and wages are more strongly related to distance to jobs and to distance to the centre than unemployment rates are. In addition, the difference in the commuting time for poor and rich workers is larger in labour markets with 500,000 workers or more.

The literature on spatial mismatch suggests that this phenomenon is predominantly urban and that it is more relevant for low-skilled minorities in larger urban areas for whom congestion costs are relatively more important. In addition, these minorities may face more limitations in their social interactions, with a significant impact on their ability to find a better match in the job market.

In this paper, we have attempted to investigate whether this negative relationship between spatial mismatch and labour market outcomes is valid in Brazil after controlling for individual characteristics. Our conclusions indicate that there is no clear relation between two different measures of accessibility to jobs and the probability of being unemployed. However, for wages there is a clear correlation, which is stronger in larger metropolitan areas.

These results indicate that in the Brazilian case, the spatial mismatch is more relevant to determine individual wages (in accordance to the relationship mentioned by Gobillon et al. 2007). On the other hand, the probability of unemployment may not be affected as much by it. This can be a result of the empirical strategy adopted here, in which commuting time spent by workers is used to calculate the potential commuting time an unemployed person would have spent in case she was employed. It may also be an indication that the spatial mismatch has a stronger effect than alternative measures of unbalance in the labour market, such as unemployment duration, as it was found in the literature (see for instance Rogers 1997). Finally, the adequate estimation strategy should allow for iterative effects between accessibility measures and metropolitan areas. This means that each metropolitan area has a particular dynamic in the labour market.

In any case, city size and skill level seem to be relevant aspects for the chances an individual has to perform well in the labour market. Intra-urban policies should aim to reduce inequalities in terms of accessibility. Since education attainment is strongly related to income, poorer neighbourhoods, which are also less served by public policies – and in the peripheries are usually far away from jobs – should be the main focus of transportation policies in the short run and education programs for middle and long run results.

This is intended to be an exploratory work. In this sense, we have explored correlations between labour market outcomes and measures of accessibility to jobs for Brazilian metropolitan areas. Our results depend on strong identification hypotheses to avoid bias related to simultaneous location decisions of workers and firms within the city (Ihlanfeldt 2006). If these conditions do not hold, our results may not represent a causal relationship, but will be meaningful in the sense of providing a better understanding of the conditional distribution of wages and the unemployment rate in the biggest metropolitan areas of Brazil.

The broader analysis of urban labour markets in Brazil provides an indication that there are relevant differences in the way workers and firms interact in space, and urban scale seems to be important to this relationship. Future work should investigate these issues more thoroughly. In this sense, different proximity dimensions could be included in the analysis, in order to investigate the factors that generate the spatial mismatch. However, this approach would require a more comprehensive database of the characteristics of Brazilian labour markets and the local interaction between individuals, which are not available yet.

Acknowledgement

Eduardo Amaral Haddad acknowledges financial support from CNPq.

References

Andersson F, Haltiwanger JC, Kutzbach MJ, Pollakowsky HO, Weinberg DH (2014) Job displacement and the duration of joblessness: The role of spatial mismatch. NBER working paper series, WP 20066. CrossRef.

Åslund O, Östh J, Zenou Y (2010) How important is access to jobs? Old question – improved answer. *Journal of Economic Geography* 10: 389–422. CrossRef.

Brueckner JK, Thisse JF, Zenou Y (2002) Local labor markets, job matching, and urban location. *International Economic Review* 43[1]: 155–171. CrossRef.

Bruzelius N (1979) *The Value of Travel Time*. Croom Helm, London

Capello R (2007) *Regional Economics*. Routledge Advanced Texts in Economics and Finance, Abingdon

Chauvin JP, Glaeser E, Ma Y, Tobio K (2016) What is the difference about urbanization in rich and poor countries? Cities in Brazil, China, India and the United States. *Journal of Urban Economics* 98: 17–49. CrossRef.

Da Mata D, Deichmann U, Henderson JV, Lall SV, Wang HG (2007) Determinants of city growth in Brazil. *Journal of Urban Economics* 62: 252–272. CrossRef.

Di Paolo A, Matas A, Raymond JL (2016) Job accessibility and job-education mismatch in the metropolitan area of Barcelona. *Papers in Regional Science* 96[S1]: S91–S112. CrossRef.

Fujita M, Thisse JF (2012) *Economics of Agglomeration: Cities, industrial location and regional growth* (2nd ed.). Cambridge University Press, Cambridge

Gibb K, Osland L, Pryce G (2014) Describing inequalities in access to employment and the associated geography of wellbeing. *Urban Studies* 51[3]: 596–613. CrossRef.

Gobillon L, Selod H (2013) Spatial mismatch, poverty, and vulnerable populations. In: Fischer MM, Nijkamp P (eds), *Handbook of Regional Science*. Springer, 93-107. CrossRef.

Gobillon L, Selod H, Zenou Y (2007) The mechanisms of spatial mismatch. *Urban Studies* 44[12]: 2401–2427. CrossRef.

Haddad EA, Barufi AMB (2017) From rivers to roads: Spatial mismatch and inequality of opportunity in urban labor markets of a megacity. *Habitat International* 68: 3–14. CrossRef.

Holzer HJ (1991) The spatial mismatch hypothesis: What has the evidence shown? *Urban Studies* 28[1]: 105–122. CrossRef.

Houston D (2005a) Employability, skills mismatch and spatial mismatch in metropolitan labour markets. *Urban Studies* 42[2]: 221–243. CrossRef.

Houston D (2005b) Methods to test the spatial mismatch hypothesis. *Economic Geography* 81[4]: 407–434. CrossRef.

Hu L (2014) Changing job access of the poor: Effects of spatial and socioeconomic transformations in Chicago. *Urban Studies* 51[4]: 675–692

Hu L, Giuliano G (2014) Poverty concentration, job access, and employment outcomes. *Journal of Urban Affairs* 39[1]: 1–16. CrossRef.

Ihlanfeldt KR (2006) A primer on spatial mismatch within urban labor markets. In: Arnott RJ, McMillen DP (eds), *A Companion to Urban Economics*. 404-417, Oxford. CrossRef.

Johnson RC (2006) Landing a job in urban space: The extent and effects of spatial mismatch. *Regional Science and Urban Economics* 36: 331–372. CrossRef.

Kain JF (1968) Housing segregation, negro employment, and metropolitan decentralization. *Quarterly Journal of Economics* 82: 32–59. CrossRef.

Kain JF (1992) The spatial mismatch hypothesis: Three decades later. *Housing Policy Debate* 3[2]: 371–459. CrossRef.

Kain JF (2004) A pioneer's perspective on the spatial mismatch literature. *Urban Studies* 41[1]: 7–32. CrossRef.

Krugman P (1995) *Development, Geography, and Economic Theory.* The MIT Press, Cambridge, Mass

Lucas Jr RE, Rossi-Hansberg E (2002) On the internal structure of cities. *Econometrica* 70[4]: 1445–1476. CrossRef.

McCann P (2013) *Modern Urban and Regional Economics* (2nd ed.). Oxford University Press, Oxford

Morrison PS (2005) Unemployment and urban labour markets. *Urban Studes* 42[1]: 2261–2288. CrossRef.

Ottaviano G (2004) Chapter 58 – Agglomeration and economic geography. In: Henderson JV, Thisse JF (eds), *Handbook of Regional and Urban Economics, Volume 4*. Elsevier, Amsterdam, 2563–2608

Partridge M, Rickman DS, Ali K, Rose-Olfert M (2009) Agglomeration spillovers and wage and housing cost gradients across the urban hierarchy. *Journal of International Economics* 78: 126–140. CrossRef.

Roback J (1982) Wages, rents and the quality of life. *Journal of Political Economy* 90[6]: 1257–1278. CrossRef.

Rogers CL (1997) Job search and unemployment duration: Implications for the spatial mismatch hypothesis. *Journal of Urban Economics* 42: 109–132. CrossRef.

Topa G, Zenou Y (2015) Neighborhood and network effects. In: Duranton G, Henderson JV, Strange WC (eds), *Handbook of Regional and Urban Economics, Volume 5A*. Elsevier, Amsterdam, 561–624. CrossRef.

Tyndall J (2015) Waiting for the R train: Public transportation and employment. *Urban Studies* 54[2]: 520–537. CrossRef.

Vieira R, Haddad EA (2015) An accessibility index for the metropolitan area of São Paulo. In: Kourtit K, Nijkamp P, Stough RR (eds), *The Rise of the City: Spatial Dynamics in the Urban Century.* Edward Elgar Publishing, Cheltenham, UK, 242–258. CrossRef.

Zenou Y (1999) Unemployment in cities. In: Huriot JM, Thisse JF (eds), *Economics of Cities.* Cambridge University Press, Cambridge, 343–389

Zenou Y (2000) Urban unemployment, agglomeration and transportation policies. *Journal of Public Economics* 77: 97–133. CrossRef.

Zenou Y (2006) Efficiency wages and unemployment in cities: The case of high-relocation costs. *Regional Science and Urban Economics* 36: 49–71. CrossRef.

Zenou Y (2009) *Urban Labor Economics* (1st ed.). Cambridge University Press, New York. CrossRef.

Zenou Y (2013) Spatial versus social mismatch. *Journal of Urban Economics* 74: 113–132. CrossRef.

A Appendix

Table A.1: Average hourly wage (in Brazilian real) in each weighting area by the distance to the main business centre, 2010.

	Distance from centre (in kilometer)							
	<2.5	2.5 to <5	5 to <10	10 to <20	20 to <30	30 to <40	40 to <50	50 or more
Macapá - AP	12.96	10.98	10.00	7.61		8.47		7.32
Aracaju - SE	11.80	14.75	9.08	9.07	5.29			
Vale do Rio Cuiabá - MT	13.95	24.74	11.97	8.83	8.03			6.57
Maceió - AL	19.11	6.60	10.14	6.53	5.28	4.90		
Florianópolis - SC	24.60	20.15	15.59	11.07	9.30	8.31	7.71	7.64
João Pessoa - PB	7.58	10.99	12.01	7.04	3.72	3.91	4.65	4.48
Grande São Luís - MA	8.79	14.08	13.28	9.52	5.33	4.58		
Natal - RN	5.60	6.14	11.04	13.95	4.62	5.64	5.24	5.24
Grande Vitória - ES	13.59	15.30	12.84	10.00	8.10	7.79	10.41	10.41
Manaus - AM	12.44	14.44	15.31	9.42			4.58	6.68
Belém - PA	20.14	13.85	11.42	8.44	6.19	5.32	5.73	5.73
Goiânia - GO	28.69	18.71	13.19	9.21	6.69	7.12	7.21	9.00
Curitiba - PR	28.58	31.99	14.51	9.45	9.10	6.62	7.38	6.68
Fortaleza - CE	6.13	7.67	12.46	10.37	6.73	4.44	3.84	4.19
Salvador - BA	12.63	23.79	9.69	8.68	10.41	7.32	8.89	7.62
Recife - PE	21.29	12.68	11.67	8.25	6.46	4.66	6.67	6.38
Porto Alegre - RS	26.99	31.46	17.78	10.09	9.15	8.65	9.05	8.41
Belo Horizonte - MG	34.98	19.37	14.90	9.85	7.50	8.90	6.88	7.79
Rio de Janeiro - RJ	7.74	15.58	17.51	14.84	11.77	8.91	8.08	8.25
São Paulo - SP	23.47	21.54	21.51	16.88	10.69	11.14	10.09	9.96

Source: IBGE

Table A.2: Average individual hourly wage (in Brazilian real) by commuting time from home to work, 2010.

	Up to 5 min.	6 min to $\frac{1}{2}$ hour	>$\frac{1}{2}$ to 1 hour	>1 to 2 hours	>2 hours
Macapá - AP	10.96	11.04	8.32	7.39	12.36
Aracaju - SE	11.72	12.16	9.27	7.85	16.58
Vale do Rio Cuiabá - MT	18.88	14.47	11.81	7.75	15.31
Maceió - AL	7.96	10.01	9.21	6.87	11.14
Florianópolis - SC	13.52	14.19	13.48	11.10	15.77
João Pessoa - PB	9.56	10.97	7.99	6.56	6.93
Grande São Luís - MA	11.16	12.48	10.49	7.76	10.94
Natal - RN	10.81	11.35	7.66	6.61	11.45
Grande Vitória - ES	14.40	13.23	11.05	8.31	10.10
Manaus - AM	10.01	13.60	10.38	7.82	8.81
Belém - PA	12.18	11.49	10.70	7.99	12.06
Goiânia - GO	17.91	13.44	9.99	7.42	14.08
Curitiba - PR	14.54	15.32	12.55	8.54	9.45
Fortaleza - CE	10.04	10.45	8.80	6.33	8.43
Salvador - BA	9.45	10.85	11.23	11.08	13.33
Recife - PE	10.58	10.17	10.50	8.36	8.29
Porto Alegre - RS	12.27	13.50	11.55	9.82	9.97
Belo Horizonte - MG	13.18	13.02	11.55	9.25	9.05
Rio de Janeiro - RJ	14.22	12.87	13.29	12.77	10.46
São Paulo - SP	17.53	16.79	15.83	13.15	11.95

Source: IBGE

Table A.3: Average unemployment rate in each weighting area by the distance to the main business centre, 2010

	Distance from centre (in kilometer)							
	<2.5	2.5 to <5	5 to <10	10 to <20	20 to <30	30 to <40	40 to <50	50 or more
Macapá - AP	8.4%	6.7%	6.6%	10.3%		12.4%		4.7%
Aracaju - SE	10.0%	6.6%	6.9%	9.5%	7.3%			
Vale do Rio Cuiabá - MT	4.1%	3.2%	4.4%	4.3%	7.7%			8.9%
Maceió - AL	5.4%	8.6%	7.5%	8.3%	13.0%	11.8%		
Florianópolis - SC	2.5%	3.8%	2.7%	1.8%	3.4%	3.9%	1.6%	1.9%
João Pessoa - PB	9.1%	5.3%	5.5%	6.8%	8.0%	8.0%	13.9%	8.8%
Grande São Luís - MA	7.7%	10.1%	7.2%	7.7%	6.0%	4.1%		
Natal - RN	7.3%	8.0%	7.4%	5.9%	8.7%	5.4%	8.6%	
Grande Vitória - ES	5.9%	4.5%	4.4%	4.9%	5.7%	4.9%	7.3%	
Manaus - AM	6.0%	8.2%	6.4%	7.7%			4.2%	6.9%
Belém - PA	6.8%	7.4%	5.3%	7.3%	7.6%	5.8%	8.1%	
Goiânia - GO	3.5%	3.2%	3.0%	3.6%	3.5%	4.4%	3.8%	5.8%
Curitiba - PR	3.4%	2.8%	3.0%	2.9%	3.4%	1.5%	2.2%	3.3%
Fortaleza - CE	5.2%	6.8%	5.4%	5.2%	6.7%	6.9%	5.1%	6.3%
Salvador - BA	8.6%	5.7%	8.2%	9.5%	9.3%	12.2%	12.9%	14.6%
Recife - PE	6.7%	7.9%	9.2%	9.7%	11.2%	11.6%	9.6%	11.9%
Porto Alegre - RS	3.7%	4.4%	4.0%	3.7%	3.9%	3.6%	3.1%	2.7%
Belo Horizonte - MG	2.7%	3.8%	4.7%	4.3%	4.1%	4.2%	4.7%	3.7%
Rio de Janeiro - RJ	6.0%	5.1%	5.1%	5.3%	6.0%	6.3%	6.7%	7.8%
São Paulo - SP	5.5%	5.2%	5.1%	5.9%	6.1%	5.5%	5.0%	4.1%

Source: IBGE

Table A.4: Regressions for the whole database

	Model 1 ln(hourly wage) OLS	Model 2 ln(hourly wage) OLS	Model 3 unemp. (P = 1) Logit	Model 4 unemp. (P = 1) Logit
Inverse of distance	1.177‡		1.024	
Inverse of distance squared	0.990‡		0.999	
Commuting time of workers in the weighting area				
% workers commuting 6' to 30'				0.547†
% workers commuting more than 30' to 1 hour				0.543†
% workers commuting more than 1 hour				0.876
Individual commuting time (Reference: up to 5')				
6' to 30'		0.994		
More than 30' to 1 hour		0.979‡		
More than 1 hour to 2 hours		0.930‡		
More than 2 hours		0.932‡		
Metropolitan area (Reference: Belém - PA, 402,170 men 25-64)				
Macapá - AP (85,494 men 25-64)	0.996	1.007	1.078	1.093
Aracaju - SE (159,838 men 25-64)	0.935‡	0.932‡	1.168‡	1.198‡
Vale do Rio Cuiabá - MT (160,638 men 25-64)	1.154‡	1.146‡	0.647‡	0.666‡
Maceió - AL (216,904 men 25-64)	0.842‡	0.842‡	1.323‡	1.334‡
Florianópolis - SC (217,208 men 25-64)	1.186‡	1.175‡	0.439‡	0.447‡
João Pessoa - PB (230,930 men 25-64)	0.811‡	0.826‡	0.995	1.026
Grande São Luís - MA (244,017 men 25-64)	0.997	0.987	1.134†	1.138†
Natal - RN (258,207 men 25-64)	0.878‡	0.872‡	1.150‡	1.171‡
Grande Vitória - ES (353,561 men 25-64)	1.107‡	1.102‡	0.804‡	0.800‡
Manaus - AM (378,496 men 25-64)	1.106‡	1.105‡	1.078	1.073
Goiânia - GO (415,541 men 25-64)	1.150‡	1.139‡	0.547‡	0.550‡
Curitiba - PR (623,103 men 25-64)	1.167‡	1.157‡	0.523‡	0.525‡
Fortaleza - CE (666,504 men 25-64)	0.875‡	0.869‡	0.874‡	0.878‡
Salvador - BA (723,297 men 25-64)	0.989	0.987	1.309‡	1.279‡
Recife - PE (745,952 men 25-64)	0.862‡	0.856‡	1.522‡	1.512‡
Porto Alegre - RS (807,268 men 25-64)	1.081‡	1.068‡	0.627‡	0.629‡
Belo Horizonte - MG (1,115,715 men 25-64)	1.105‡	1.097‡	0.679‡	0.665‡
Rio de Janeiro - RJ (2,402,075 men 25-64)	1.131‡	1.122‡	0.916†	0.865‡
São Paulo - SP (3,953,270 men 25-64)	1.236‡	1.225‡	1.006	0.947
Education attainment (Reference: up to incomplete primary school)				
Complete primary school to incomplete college	1.304‡	1.305‡	0.782‡	0.788‡
Complete college	2.576‡	2.588‡	0.426‡	0.440‡
N	583,184	583,184	621,359	621,359
Pseudo R squared			0.048	0.048
Adjusted R squared	0.406	0.405		

Source: Authors' calculations
Notes: Controls for Models 1 and 2: age, age squared, colour or race, household head, with children up to 15 years old, married, sector of activity, occupation, existence of a formal contract. Controls for Models 3 and 4: age, age squared, colour or race, household head, with children up to 15 years old, married. Coefficients are presented as odds-ratios. Significance levels: * p <0.10, †p <0.05, ‡p <0.01. Only male individuals aged 25 to 64 years old living within a distance of 30km from the centre are considered in the analysis. Sampling weights are taken into account with Stata command pweight. Complete tables are available under request to the authors.