Laura H. Atuesta, J. Eduardo Ibarra-Olivo, Nancy Lozano-Gracia and Uwe Deichmann

Access to employment and property values in Mexico

Article (Accepted version) (Refereed)

Original citation:
Atuesta, Laura H. and Ibarra-Olivo, J. Eduardo and Lozano-Gracia, Nancy and Deichmann, Uwe (2018) *Access to employment and property values in Mexico*. Regional Science and Urban Economics, 70. pp. 142-154. ISSN 0166-0462
DOI: 10.1016/j.regsciurbeco.2018.03.005

Access to employment and property values in Mexico

Abstract

Location is one of the main characteristics households consider when buying a property or deciding where to live, since it determines accessibility to transport and hence to jobs and employment. Using a geographically-referenced dataset on new housing developments, this paper estimates how households value accessibility in Mexico City. Results are shown considering road accessibility to formal employment subcenters (private accessibility) and distance to the main public transport stations in the city (public accessibility). Results suggest that accessibility to employment subcenters is valued as an amenity by households but being closer to a Metro station is perceived as a disamenity. Moreover, households located in neighborhoods with a greater proportion of informal workers and with lower education levels give a lower value to private accessibility than households located in neighborhoods with a lower proportion of informal workers or in high-educated neighborhoods. These results are evidence of the existence of spatial segregation in the city where disadvantaged households are segregated, not only because of their economic conditions, but because they are located farther away from employment opportunities. The results in this work stress the importance of thinking about integrated land use and transport policies.

Keywords: accessibility, housing location decisions, spatial segregation, hedonic housing models.

JEL codes: R210, R230, R310, R410

1. Introduction

Household choices shape cities. Households select housing units based on affordability and the amenities they provide. A growing literature in urban economics seeks to better understand these choices using information revealed by market prices to estimate the value households assign to various housing characteristics. Arguably the most important characteristic is location which greatly influences access to amenities and employment opportunities.

Infrastructure investments and public services interact with these choices. By investing in roads and transit services, public policies can make different parts of a city more or less attractive, which will be reflected in property values. A better understanding how households value transport options can help design transport policies that increase private, but also social welfare—for instance by making sustainable transport more attractive. Such improved understanding is especially needed in fast growing cities in emerging and developing economies. However, most of the literature in these countries has focused on dwelling unit characteristics rather than on access to opportunities, specifically income earning opportunities.

Studies in emerging and developing economies offer interesting insights because of the spatial and socioeconomic segregation that these cities experience. Recent literature (Suárez et al. 2016; Negrete and Paquette Vassalli, 2011), suggests high-income households have a higher likelihood to own and use private cars and spend less time commuting to their workplace. Conversely, low-income households are located farther away from employment centers, and have access to more limited opportunities, whether it is because they spend more time accessing formal employment, or because they segregate economically finding informal or lower-paid jobs closer to their places of residence (Negrete and Paquette Vassalli, 2011).

This paper uses a unique, geographically referenced dataset on new housing developments to estimate how households value accessibility in Mexico City. Accessibility is measured by access to employment subcenters, access to the historic center (or Central Business District) and access to the two main transportation infrastructures in the city: The Bus Rapid Transport (BRT) system and the Metro system. These findings provide useful insights to understand how households make transport mode decisions for commuting purposes, as well as how they choose the location of their residences.

Our paper builds on a growing literature on the valuation of dwelling unit characteristics but makes several additional contributions. First, for the identification of the employment subcenters, we use the results of Atuesta and Ibarra-Olivo (2017) that estimate nonparametrically the subcenters following McMillen (2001). These results are more robust than those using specific thresholds of employment and allow comparability through time and among cities.

Second, we consider not only private accessibility to employment subcenters using road distance, but also proximity to public transport infrastructure to determine how much households value location. Since employment subcenters are estimated based only on formal employment, other types of accessibility could be more valuable for households that are not employed in formal jobs. Then, heterogeneity in terms of valuing accessibility is observed when analyzing the socioeconomic level of households or the proportion of residents who are informal workers in each neighborhood. Our results account for these differences analyzing the value of accessibility to different types of households.

Third, thanks to a new dataset of housing developments compiled by a private market intelligence firm we are able to base our analysis on repeated sales of very similar housing units, rather than rely on one or repeated cross-sections of market prices (Shiller, 1991; McMillen, 2003). This also helps reduce the omitted variables bias, common in standard hedonic estimates (Wallace and Meese, 1997).

Our results suggest that accessibility to employment subcenters is considered an amenity for households in Mexico City. Being one percent closer to employment subcenters (and controlling for their employment level) increases housing values by between one and almost three percent. As employment subcenters earn importance in the spatial structure of the city, the Central Business District loses significance. In two of the three specifications analyzed, being closer to the Central Business District is perceived as a disamenity that depreciates housing prices, suggesting that, employment decentralization is playing an important role for households' location choices.

In terms of public transport, our estimations suggest that being closer to a Metro station is considered a disamenity for households, but after a certain distance, proximity to a Metro station starts being appreciated as an amenity. This result is expected, since housing prices can be affected by the informal commerce, congestion and insecurity that concentrate at the gates of the Metro stations. Since mass transport systems can be substituted by other modes of transportation, close proximity to public transport stations seems to cause more disadvantages than benefits in Mexico City. Our results also suggest households located in neighborhoods with a greater proportion of informal workers assign a lower value to private accessibility than households located in neighborhoods with a lower proportion of informal workers.

The remainder of this article is organized as follows. The following section provides a brief review of the literature on the spatial structure of cities and the role of public and private transportation for accessibility improvements in Latin America with an emphasis on Mexico City. Section three describes the data sources and section four discusses the econometric estimation approach. Section five summarizes the results and section six concludes and discusses some policy implications.

2. Literature review

The monocentric city model has for long been the workhorse of urban economics. But the assumption of a single central district where economic activity is concentrated is quite unrealistic for large and diverse cities. Such cities will have several centers that offer agglomeration economies for firms while reducing congestion in the CBD and potentially offering lower commuting costs for workers. In this section, we review the literature that focusses on analyzing the spatial urban structure of the city, and its relationship with public and private accessibility and with the economic and spatial segregation of households.

Conclusions regarding the extent of polycentricity in Mexico City are mixed. While some studies suggest that Mexico City has decentralized its employment (Aguilar and Alvarado 2005; Sanchez Trujillo 2012; Aguilar, 2011; Fernandez Maldonado et al; 2013; Romein et al. 2009), others conclude that the city is still behaving as a monocentric city with most employment opportunities concentrated in the CBD (Suárez 2007; Suárez and Delgado 2007; Suárez and Delgado 2009). Specifically, Suárez and Delgado (2009) posit that, while the city’s residents have decentralized, employment opportunities are still concentrated in the center. In fact, according to the Population and Economic Censuses, there are more people working in the central city than residents living there, evidence that the city still has some characteristics of a monocentric structure (Suárez and Delgado 2009). Since more than 50 percent of employment in Mexico City is considered informal (Azuara and Marinescu 2013), Suárez and Delgado (2009) suggest that employment subcenters identified only considering formal employment could be biased and may not fully reflect the real urban structure of the city.

Notwithstanding the degree of monocentricity of a city, a well-connected city can provide high access to opportunities. However, in Mexico City, access to opportunities faces serious challenges given its size and limited infrastructure, leading to high congestion and severe air pollution. In the last couple of years (2016-2017), more than a million vehicles were taken off the streets in order to reduce pollution in the city (Milenio 2016; Excelsior 2017). According to Romein et al. (2009), 80 percent of the daily trips use collective transportation modes (formal or informal) with the remaining 20% using private cars, including taxis. Different modes of transportation exist including private cars, semi-public buses (private concessions for public service), and public transportation (BRT, Metro, and LRT). Hence, a full characterization of accessibility in Mexico City must include both public (access to public transportation), and private accessibility (access to roads where private and semi-public modes of transportation operate).

Indicators of accessibility summarize the ease by which destinations of interest can be reached from a given location (Yoshida and Deichmann 2009). For example, they can be defined as the time households spend getting to job locations. Ideally one would extract this information from detailed origin-destination surveys, but such surveys are rarely available, and when available, they are only representative at an aggregate level, more often the municipal level.[1] Instead, access measures can be approximated by estimating the distance or time to reach destinations using a detailed geographically referenced data set of origins, destinations and the transport network. Gravity type models summarize average accessibility to a range of potential destinations such as employment subcenters with weights assigned based on the number or proportion of jobs offered. The literature on access measures goes back to Hansen (1959) and Wilson (1967). An example application is Cervero et al. (1999) who use an 'occupational match' approach to find consistency between employed residents' skills and employment in specific sectors by neighborhoods.

[1] The last origin and destination survey of Mexico City was conducted by INEGI in 2007 (INEGI 2007).

How does access to job opportunities and public transit affect households' choices on where to live? Although the literature on housing location choices is large, the increasing availability of detailed spatial data has more recently led to more fine-grained analysis. The emerging literature suggests that the effects are by no means uniform across cities in different parts of the world. Adair et al. (2010) recognize three types of access measures based on time taken, cost of travel and convenience, plus the availability of different transport modes. They find significant heterogeneity in the impact of access on property values across local sub-markets in Belfast. Ahlfed (2011) uses employment access measures to estimate land value gradients in Berlin calculating different gravity-type measures for urban rail stations and main roads. In Latin America, Perdomo (2011) and Munoz-Raskin (2010) estimate the impact of the BRT system on housing prices in Bogota. Agostini and Palmucci (2008) estimate the effect of a new Metro line in Santiago, Chile, on housing prices. And Flores Dewey (2010) estimates the effect of a new BRT line on housing prices in the municipality of Ecatepec in Mexico City.

All these studies find that accessibility is capitalized into housing prices, either measured as proximity to employment subcenters or proximity to public transport stations. However, households at different income levels experienced this appreciation differently, with middle-income households benefiting more than low income households from such appreciation in the city of Bogota, Colombia; instead, in Belfast, low-income households seem to experience larger benefits from similar effects. Again, for Bogota, Avedano Arosemena (2012) finds that proximity to employment centers raises housing prices, while proximity to major transport routes has no or negative effects. This suggests that the road network infrastructure does not deliver significant connectivity benefits in the city. Martinez and Viegas (2009) examine the effect of accessibility to the Metro system, the rail system and the main roads in Lisbon, Portugal, on housing prices, finding strong and significant effects of public transit accessibility on housing prices, with impacts ranging between 5 percent and 10 percent. Road access, in contrast, appears to be a disamenity with a negative effect on housing prices.

The literature has also studied the perception of public transportation for different households in Mexico City, suggesting a substitution effect between public and private transportation. Crotte et al. (2009) found that the Metro system is perceived as a normal good for low-income households, but as an inferior good for high-income households. On the other hand, Guerra (2014) suggested that white-collar workers in Mexico rarely use the public transport system because not all the employment sub-centers are well connected, so they must choose between private transportation and informal public transportation modes.

Other studies have focused on the relationship between public transportation and urban expansion. Negrete and Paquette Vassalli (2011) suggest that only the central city of Mexico City is well connected in terms of public transportation. With the new trends of suburbanization where the low-income households are offered dwellings inside enormous gated communities (in some cases with more than ten thousand properties), accessing formal transportation is difficult, and these households may end up paying more for informal transportation, and spending more time commuting than when they were living in more central locations.

This segregation is also studied by Suárez et al. (2016). Contrary to what was concluded in previous studies (Suárez and Delgado, 2007), the authors present evidence suggesting that the poor travel less in Mexico City than the middle- and high-income households. In many cases, low-income households choose an informal job close to their residence to avoid long and expensive displacements. But remaining close to their residences limits the opportunities they have, and in many cases, leaves them with lower-paying jobs without social security or employment benefits. Then, segregation is exacerbated, since not only the poor live further away, but they also are excluded from formal employment opportunities located in the central city.

Building on previous literature, this article explores how private and public accessibility is capitalized into property values, estimating heterogeneous effects in terms of socioeconomic conditions of households and the level of informality in neighborhoods where these households are located. These results provide insights about the effect of the spatial urban structure in property values and in the economic and spatial segregation within the city.

Description of the data

We use three main sources of data. First, we use a commercial database of housing prices and dwelling characteristics developed by SOFTEC (a housing market consulting firm) which covers the period from 2002 and 2013. Second, we calculate the distance of each household to the closest public transit stations of each type using maps with detailed information on road networks and public transport infrastructure investments. Finally, we use Mexico's Housing and Population Censuses at the lowest level of geographical disaggregation available, called AGEB,[2] to calculate the average level of education and the proportion of informal workers in each neighborhood. In the following paragraphs, we provide additional details on both the housing price data and the censuses including a brief discussion of the specific variables extracted for our analysis.

2.1.Housing data

We use housing information gathered by SOFTEC[3] on all new construction in Mexico City between 2002 and 2013. Information on sales prices and characteristics of projects are provided by private developers to SOFTEC at the beginning of the construction of the project, and every three months until the last property in the construction project is sold. Projects include all types of residential construction, including gated-communities of individual dwellings or buildings, or just a single building. We use housing information gathered for the last trimester each year from 2002 to 2013 (information regarding sales from October, November, and December of the same year).

The data are aggregated at the project or *fraccionamiento* (housing divisions or housing projects) level and include the date when the construction project enters into the SOFTEC inventory, the date in which all units in the same project were sold, and the date when the construction was completed (all the housing units built). Besides providing several

[2] The AGEB (or geographic and statistical area) is defined as a territorial unit inside a municipality. There are urban and rural AGEBs depending on their housing density. An urban AGEB is a geographic area of a group of blocks delimited by streets, avenues, sidewalks or other construction easily identified, in which its land is used mainly for occupational purposes, industries, provision of services or commercial purposes. All urban AGEBs must be located inside urban localities.

[3] SOFTEC is a private company created in 1988 with the objective of collecting data to monitor the housing supply in Mexico. According to conversations with their sales representative, their database represents about 85% of the total housing supply in Mexico, gathering approximately 350,000 observations (projects) per year at the national level.

characteristics of the project as well as the housing units included in it, the dataset also provides information on the median price of properties in the project sold during every trimester.[4] By combining the information of different projects and different years, we are able to construct an unbalanced panel dataset in which the unit of analysis is the project and the number of times it appears in the dataset depends on the number of years it takes for the last unit to be sold.[5]

The number of properties in each project varies, as does the number of observations per year since not all the projects are in the market for the same amount of time. By summing up the average number of properties sold in each development per month, we are able to calculate the monthly average housing supply for each year. Table 1 shows the number of projects available on average each year, the estimated average housing supply, and the mean housing price for each year.

The median price experienced a growth of 2.8 percent during the 11-year period suggesting that the reduction of the housing supply (from 2002 to 2013) was reflected in scarcity of housing units and consequently, an increase in housing prices. Every year in which the average housing supply had negative growth, the median price experienced growth (except for two years in which the rise in the median price was less than one percent, and 2009, a year in which both, the housing supply and the median price, experience a negative growth of two percent).

Table 1: Number of observations (projects), estimated housing supply and median price (2002- 2013)

Year	Average number of projects on sale each year	Average number of properties sold per month per project	Estimated housing supply (total number of properties sold per year)	Median price (MXN$)
2002	821	5.43	4,458.0	882,084
2003	919	3.77	3,464.6	1,140,472
2004	1,067	4.21	4,492.1	1,114,346
2005	1,233	4.05	4,993.7	1,114,212
2006	1,515	3.06	4,635.9	1,131,260
2007	1,542	2.52	3,885.8	1,154,502
2008	1,303	2.85	3,713.6	1,184,655
2009	1,235	2.94	3,630.9	1,160,689
2010	1,141	3.07	3,502.9	1,154,515
2011	1,014	3.46	3,508.4	1,140,002

[4] SOFTEC gathered information regarding the representative property of each project, including its sale price. Therefore, for each project, information regarding the median price is obtained.

[5] This value varies between one and eleven years.

2012	997	3.68	3,669.0	1,131,161
2013	967	3.50	3,384.5	1,158,109

2000 constant prices in MXN$. Information of the last price reported given the representative property in each project.

The behavior of the housing market can be explained partially by the housing policies implemented in the country during the last 20 years. At the beginning of the 2000s, the housing policy in Mexico was focused on increasing the homeownership rate and providing housing opportunities for the low-income population. Following this objective, all formal workers in Mexico began contributing five percent of their salaries to Housing Funds, to be used later for potential mortgages. The mortgage opportunities in the country increased (Vizuet, 2010), and this improvement in credit accessibility led to expectations of increased demand. As a result, developers accelerated housing construction, increasing housing supply. The available credit combined with greater housing supply translated in a stagnation of housing prices. However, housing demand for the income segments to which construction was targeted, did not increase as fast as expected, elevating the housing inventory in the main cities, and the vacancy rate in the country (Monkkonen, 2015). In the second half of the 2000s, prices dropped, inventories stocked, and the frequency of sales decreased. For instance, the housing sales in Mexico City grew from 253 thousand properties sold in 2001 to 665 thousand in 2005, and decreased in 2006 to 544 thousand; while the number of mortgages approved increased by 165% from 2001 to 2007 (Vizuet, 2010), and the national vacancy rate per 100 thousand properties grew 51% from 2005 to 2010 (INEGI).

When analyzing the behavior of specific housing projects, we also observe changes in their characteristics through time. While from 2003 to 2006 more than 300 projects were inaugurated each year, this figure decreased to 190 in 2010. Almost half of the projects (48 percent) in the dataset experienced a negative growth in the median price of their housing units, with an average depreciation of 7.56%. Not all developers maintained the same characteristics of their representative house every year. From 3,321 projects with price information for more than one year, only 1,878 (56 percent) maintain the same housing characteristics (including area size, number of bathrooms, bedrooms, and lot size). For instance, 1,107 projects changed the size of their properties, being smaller than the original size in 57 percent of the cases. Housing and project characteristics are shown in Table 2.

Table 2: Characteristics of the properties and developments. Means and standard deviations in parentheses.

Variables	Average values (standard deviation)
Median housing price of properties sold in the last trimester of the year	1,144,346 (1,497,433)
Average number of years with information (how long in years the development takes to be sold out)	3.109 (2.076)
Average number of properties sold in a month (in the same development)	3.447 (12.00)
Total number of properties sold per development	84.62 (460.2)
Area of the property (mts2)	125.8 (93.46)
Number of bathrooms	2.133 (1.655)
Number of bedrooms	2.368 (0.586)
Number of stories	5.864 (5.124)
Property has parking place	1.619 (0.932)
Property has bedrooms for employees	0.276 (0.447)
Socioeconomic classification: 1. Social 2. Affordable 3. Medium 4. Residential 5. Residential plus	3.771 (0.910)
Development has a pool	0.125 (0.330)
Development has security surveillance	0.777 (0.416)

Source: SOFTEC.

2.2. Transport infrastructure and accessibility measures

To measure accessibility, we distinguish between "private accessibility", which considers road network distances from each residential point to all employment subcenters, and "public accessibility" which measures distances to public transit options. In this section, we describe the methodology used for identifying subcenters in Mexico City and the accessibility measures used in the empirical specification.

2.2.1. Identification of subcenters

Employment sub centers used to construct private accessibility measures are those identified in Atuesta and Ibarra-Olivo (2017). The authors follow McMillen's (2001) two-stage nonparametric procedure, introducing some adjustments to fit the Mexican case. Below we briefly describe their procedure. All the parameters are reported in Table 3.

1. Using the logarithmic transformation of employment density per square kilometer, the following kernel-weighted local polynomial regression of employment density on distance to the Central Business District (CBD) is estimated:

$$y_i = w_p(x_i) + \varepsilon_i \qquad (1),$$

where y_i is the log employment density in location i, x_i is the Euclidean distance from location i to the CBD, $w_p(x_i)$ is the smoothing function given by a polynomial expansion of order p, and ε_i is a prediction error.

   1.1. To estimate $w_p(x_0) = E[Y|X = x_0]$ with no assumption about the functional form of $w_p(\cdot)$, a local polynomial of order p fits a regression of Y on the polynomial terms $(x_j - x_0) + (x_j - x_0)^2 + \cdots + (x_j - x_0)^p$ for every smoothing point x_0. The terms are weighted by a smoothing function given by a triangle kernel K_h.

   1.2. Potential subcenters are considered to be locations with positive residuals at the 1 percent significance level: $y_i - \hat{y}_i/\hat{\sigma}_i > 2.576$, where $\hat{y}_i$ is the fitted value of employment density y at site i, and $\hat{\sigma}_i$ is the estimated standard error for the prediction. This criterion suggests that these locations significantly contribute to the overall employment density function even when distance to the CBD is considered. Tracts located within a 5 km radius of the CBD are excluded from the list, as they are likely part of it.

   1.3. The remaining census tracts with residuals that exceed 1.5 standard deviations are grouped together if they fall within 3 kilometers from each other. The resulting census tract groups represent reasonably larger potential subcenters as they contain many nearby tracts.

2. The second step is an estimation of a semi parametric regression with the employment density of each census tract on the left-hand side, and the distance to the CBD as well as to possible employment subcenters as independent variables:

$$Ln(ED)_i = g(x_i) + \sum_{j=1}^{S} \left(\delta_{1j} D_{ij}^{-1} + \delta_{2j} D_{ij}\right) + u_i \qquad (2),$$

where $g(x_i)$ is the Fourier transformation of distance of census tract i to the CBD; $-D_{ij}$ and D_{ij}^{-1} are the distance and the inverse distance of census tract i to each of the possible subcenters j of the subsample, respectively. Distance variables enter negatively into the regression to ensure that a coefficient is positive when proximity to a specific subcenter increases densities of census tracts.

2.1. The reduced equation is estimated again until all subcenter distance (and inverse distance) coefficients are positive and significant at or below the 20 percent level. Subcenters with p-values lower than 20 percent and positive estimated coefficients for either the distance or the inverse distance are included in the final list of subcenters for each of the cities.

2.2. The variable distance to the CBD enters the equation non-parametrically using a Fourier expansion. First, the distance variable is transformed to lie between 0 and 2π; then, the Fourier expansion is estimated as follows:

$$g(DCBD_i) = \lambda_0 + \lambda_1 z_i + \lambda_2 z_i^2 + \sum_{q=1}^{0} \left(\gamma_q \cos(qz_i) + \gamma_q \sin(qz_i)\right) \qquad (3),$$

where z_i is the transformed variable of distance to CBD and q is the order of the Fourier expansion which is selected using the AIC criteria (the order that minimizes the AIC of the equation in step 2). The order of the Fourier expansion chosen for the specification as well as the number of final sub centers for each year are shown in Table 3.

Table 3: Parameters used for the identification of employment sub centers and final list of sub centers according to the estimations for each year.

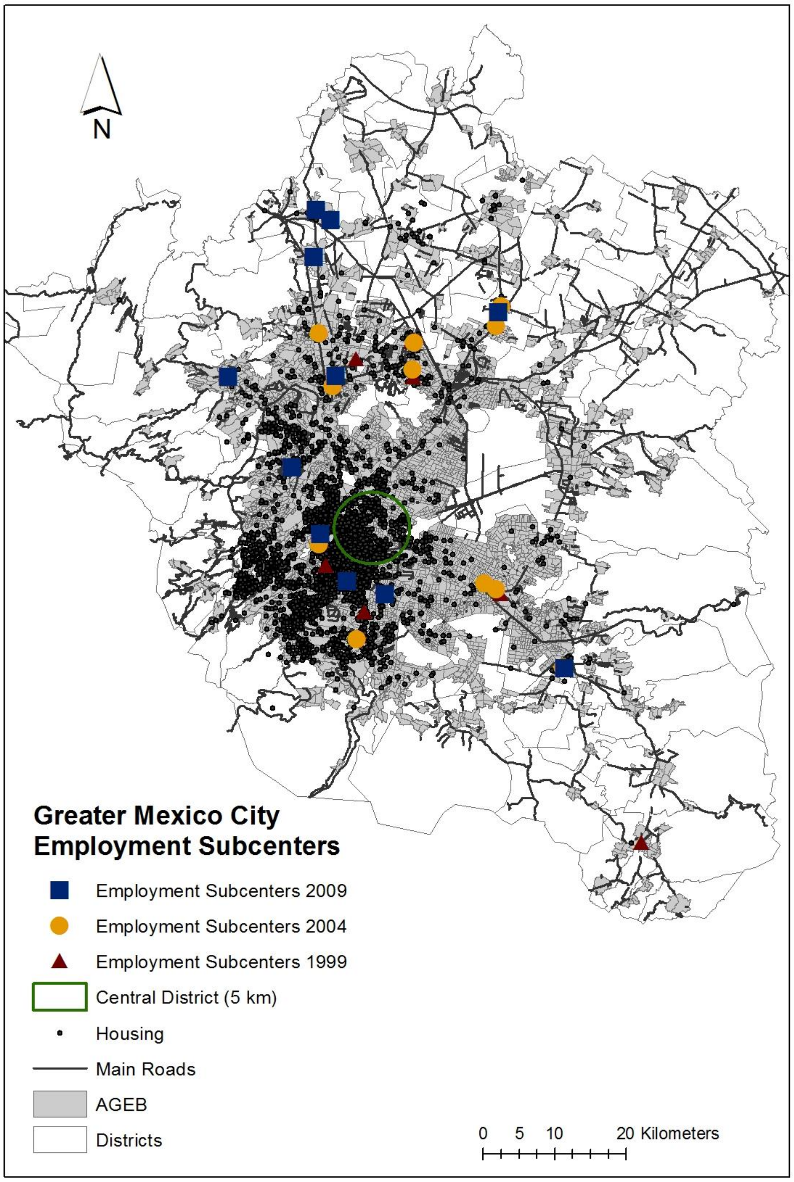
Parameters		
Kernel bandwidth[6]	10	
CBD radius	5 km	
Nearby distance	3 km	
Significant positive residuals	1999	2,100
	2004	2,352
	2009	2,483

6 Several bandwidths were specified. Since no significant changes in the smoothing function were found, a rule-of thumb kernel bandwidth is used, which minimizes the conditional weighted mean integrated squared error.

Residuals above 1.5 Standard Deviation	1999 2004 2009	204 204 205
Resulting potential sub centers	1999 2004 2009	59 64 57
Order of Fourier expansion	1999 2004 2009	6 6 6
Final Sub centers	1999 2004 2009	10 13 11

Figure 1 shows the identified subcenters and the location of the projects included in the dataset. Most of the subcenters are relatively stable across the three years analyzed. Subcenters are located mostly in the Western and Southern parts of the city with some also located in the North. The East side of the city is less developed in terms of employment and is also the area where most of the lower-income families live. Ten subcenters were identified in 1999, 13 in 2004 and 11 in 2009. Most of the subcenters are consistent through the years identifying important economic sites such as Santa Fe (the financial district on the Western side), the Olympic Village on the South side, Pantitlán and the wholesale market in the Southeast, and some important employment centers in the North such as Tlanepantla, Cuautitlán and Ecatepec.

Figure 1: Housing communities and employment subcenters in 1999, 2004 and 2009 for the metropolitan area of Mexico City



Source: SOFTEC provided information of closed communities which were geocoded by the research team. INEGI provided information of the main roads and location of districts and AGEBs. Location of subcenters estimated by Atuesta and Ibarra-Olivo. (2017).

Subcenters in Mexico City have been identified before by using employment thresholds, but none of these studies have used a parametric or semiparametric approach, but instead impose a priori assumptions on thresholds of employment concentration. For instance, Aguilar and Lozano (2014) identified subcenters as areas with an employment concentration above the city mean plus one standard deviation; Sanchez and Trujillo (2012) defines subcenters as those locations that exceed a given threshold of employment, following the approach used by Giuliano and Small (1991); finally, Casado Izquierdo (2012) uses the methodology proposed by Coombes et al. (1986) which consists of five steps to final subcenter identification. Most of the subcenters detected by these previous studies are also identified in this work. The main difference is that, by identifying the CBD also as a subcenter, the threshold methodology fails to detect peaks in employment density which are achieved by a smoothing function. The nonparametric approach is a flexible way of subcenter identification that allows to smoothen out the employment density function, detect local peaks, and avoid the use of thresholds.

2.2.2. Private and public accessibility measures

Private accessibility measures are estimated from each housing project to each subcenter. The private accessibility index uses an exponential distance decay function as follows (Song, 1996):

$$A_i = \sum_j (E_j) \exp -d_{ij} \qquad (4),$$

where A_i is the accessibility index for each project i; E_j is the number of employees working in subcenter j; and d_{ij} is the road distance variable from project i to employment subcenter j.[7] The distances between each pair of housing project i to employment subcenter j are calculated using the urban road maps obtained from INEGI and hence reflect the road travelled by individuals rather than the straight-line (Euclidean) distance. Employment in each subcenter was calculated using the Economic Censuses of 1999, 2004 and 2009, disaggregated to the AGEB (census tract) level.

For the estimation of public accessibility, we considered the road distance from each housing project to the main public transport investments—the Metro system and the BRT system (called *Metrobús* in the central city and *Mexibús* in the suburban area). The Metro system is much older than the BRT. The first line of the Metro opened in 1969 and the last line was finished in 2012. It has in total twelve lines covering 226 kilometers. The whole system has 195 transit stations with 24 transfer stations. The BRT system (Metrobus) has

[7] Summary of statistics of accessibility measures are available in the appendix.

six lines covering 125 kilometers. The Metrobus is newer than the Metro with its first line opened to the public in 2005 and the last one in January 2016. Mexico City also has other public transportation systems working at the local level or connecting to the Metro or the Metrobus:

- The Mexibus in the state of Mexico;
- The electric Trolebus (zero emissions) with 15 lines;
- The RTP buses connecting different Metro stations to other areas of the city;
- The LRT in the south of the city with 16 stations and two terminals; and
- Microbuses or small buses (capacity of 24 passengers) that use the main roads of the city and stop every time a passenger requires a stop.

Table 4 shows the information for each line opened in the period between 2002 and 2013, as well as the dates in which they were inaugurated and the dates in which the construction was announced (the announcement date is important to consider anticipation effects in the analysis).

Table 4: Announcement and inauguration dates of different transportation infrastructures during the 2002-2013 period.

	Distance	**Number of stations**	**Inauguration date**	**Announcement date**
Metro line 12	24.5km	20	October 2012	August 2007
Metrobus line 1	30km	47	First part: June 2005 Second part: March 2008	First part: 2003 Second part: 2007
Metrobus line 2	20km	36	December 2009	January 2007
Metrobus line 3	17km	33	February 2011	November 2009
Metrobus line 4	28km	35	April 2012	November 2010
Metrobus line 5	10km	18	November 2013	February 2013
Mexibus	52km	95		October 2010

Some transportation investments are not included in Table 5 because they were announced, inaugurated and constructed in a period different than 2002-2013.

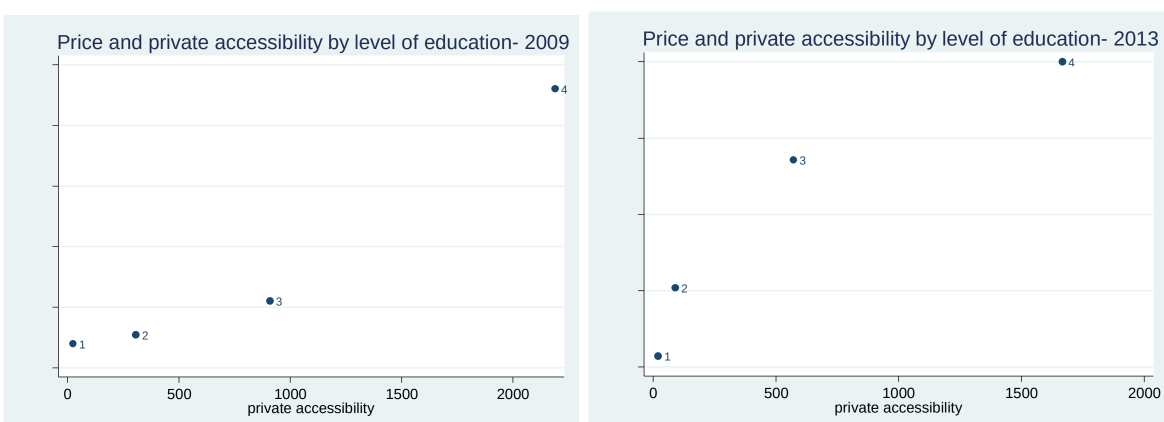
2.3.Housing and Population Censuses at the AGEB (census tract) level

Finally, to capture heterogeneous effects between households belonging to different socioeconomic levels or between neighborhoods with different proportion of residents employed in the informal sector, we use information from the Housing and Population Censuses of 2000, 2005 and 2010, disaggregated at the AGEB level.

We analyzed two neighborhood characteristics: the average years of education, and the proportion of individuals affiliated with the *Seguro Popular*. The first variable is used as proxy for measuring the socioeconomic status of each neighborhood. The second one, as a proxy of the proportion of individuals employed in the informal sector. According to ILO (2003) guidelines, informal workers are those who are not subject to national labor legislation, income taxation, social protection, or employment benefits. In 2002, the Mexican federal government implemented the *Seguro Popular* to provide health insurance benefits to those workers employed in the informal sector. The insurance is free for the majority of the workers and it is mostly used by low-income households not employed in the formal sector (Azuara and Marinescu, 2013).[8] The proportion of workers affiliated with *Seguro Popular* is an estimate of the proportion of informal workers in each neighborhood.

Figures 2 and 3 below show the relationship between median housing prices and both private and public accessibility measures (respectively), by each education level quartile. Figures 4 and 5 show similar relationships but categorizing the neighborhoods by their level of informality. In each figure, the panel on the left shows the relationship for 2009 and the panel on the right for 2013.

Figure 2: Relationship between housing prices and private accessibility by quartiles of education level

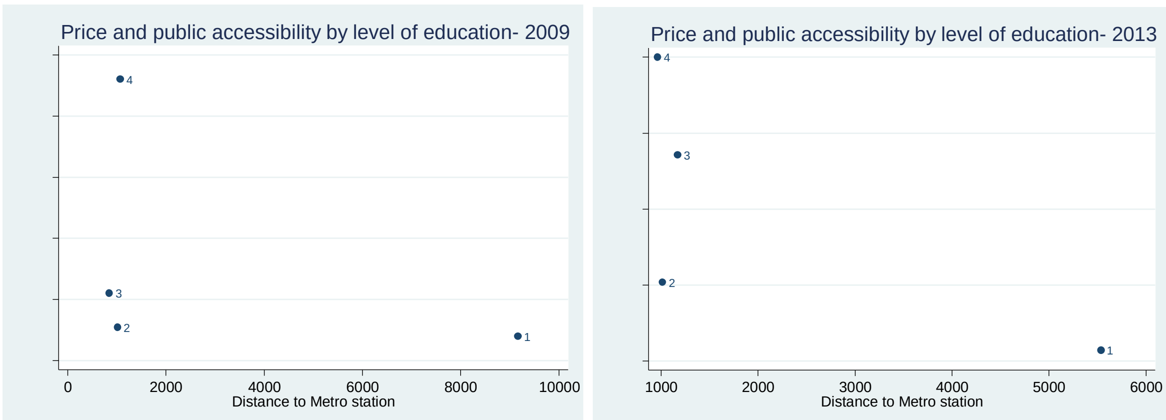


Note: *Private accessibility measured by accessible employment opportunities in employment subcenters, discounted by the distance between each property and each subcenter.

8 Another definition of informality could be the proportion of workers that are not affiliated to any social security program. However, in this definition, self-employed professionals would be included. This group of population has very different socioeconomic conditions than those individuals employed in the informal sector. Then, we decide to use as proxy affiliation to Seguro Popular.

Figure 2 shows a positive relationship between private accessibility, housing prices and education. Neighborhoods with the highest levels of education (neighborhoods located in the fourth quartile) are those that are better located in terms of proximity to employment subcenters and have the greatest housing prices (measured at the median level), both in 2009 and 2013.

Figure 3 Relationship between housing prices and public accessibility* by quartiles of education level



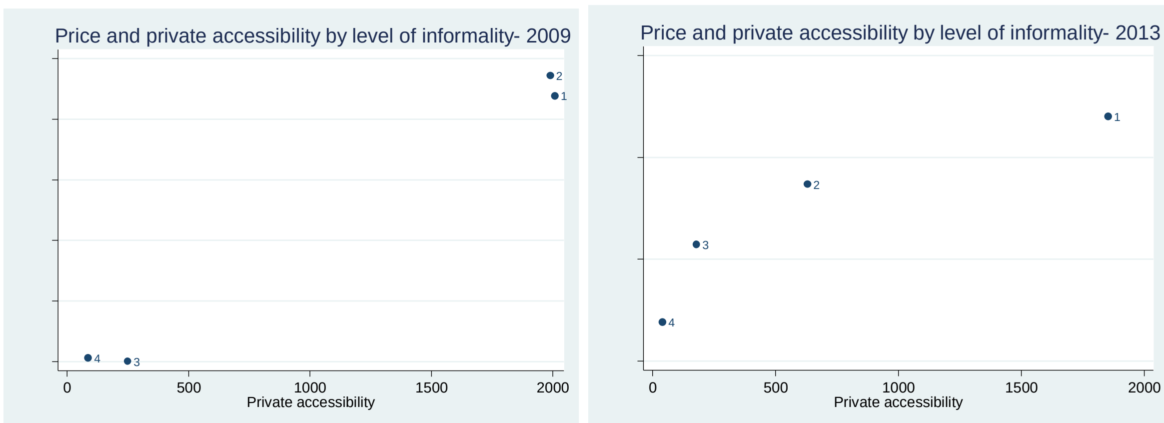
Note: *Public accessibility is measured by distance to the Metro station

Figure 3 shows the relationship between housing prices and public accessibility (measured by distance to the Metro station) by each education quartile. High-education neighborhoods are still showing the highest housing prices, measured at the median, and are also the closest to the Metro stations. Neighborhoods in the second and third education quartiles are also closer to the Metro station. The main difference is observed with neighborhoods with low levels of education, which are significantly farther away from a Metro station than other types of neighborhoods. These graphical results show the existence, not only of socioeconomic differences between neighborhoods, but also of a spatial segregation within the city.

A possible reason why low-educated households value proximity to employment subcenters less is that they are mostly employed by the informal sector. To analyze this explanation further, we conduct a similar analysis in Figures 4 and 5 (for private and public accessibility, respectively) but categorizing the neighborhoods by informality quartiles (using as a proxy the proportion of individuals in the neighborhoods affiliated with the

Seguro Popular). Neighborhoods located in the lowest quartile are neighborhoods with the lowest level of informality, while neighborhoods in the highest quartiles are neighborhoods with the greatest proportion of informal workers.

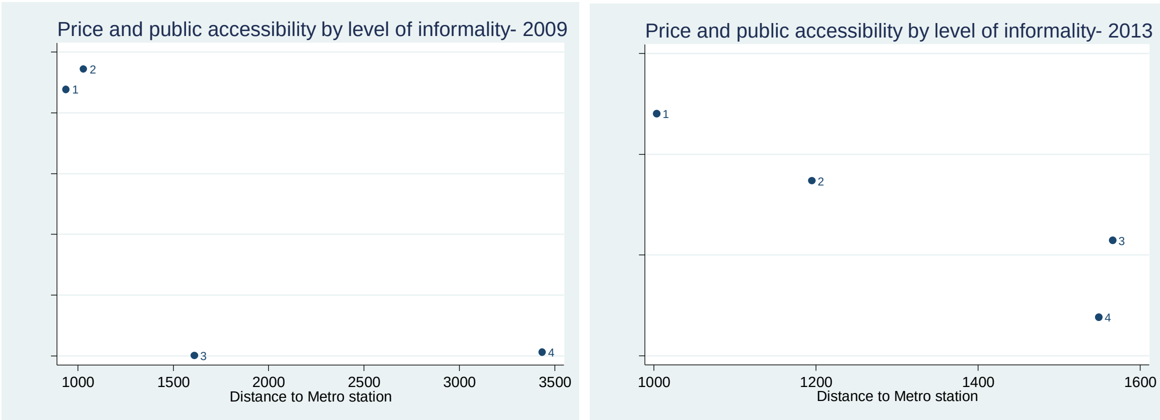
Figure 4: Relationship between housing prices and private accessibility* by quartiles of informality**



Note: * Private accessibility measured by accessible employment opportunities in employment subcenters, discounted by the distance between each property and each subcenter. **Informality measured by the proportion of informal workers in the living in the neighborhood

Figure 4 shows an inverse relationship between the proportion of formal workers in the neighborhood and accessibility to employment subcenters. Households located in neighborhoods with lower informality (first quartile) appreciate more than households located in neighborhoods with greater levels of informality (fourth quartile). On the other hand, private accessibility is lower in neighborhoods with greater levels of informality. Since the employment subcenters were estimated using information of formal employment, these results are not surprising: informal workers do not capture proximity to employment subcenters as an amenity in their housing location choices.

Figure 5: Relationship between housing prices and public accessibility* by quartiles of informality**



Notes: *Public Accessibility measured by distance to the Metro station. **Informality measured by the proportion of informal workers in the living in the neighborhood, neighborhoods in quarter four has the greatest proportion of informal workers.

Figure 5 suggests that households located in neighborhoods with the greatest levels of informality are further away from Metro stations and have lower housing prices than households located in neighborhoods with lower levels of informality. The spatial segregation we identified when we split the sample by education status is also observed when we categorize observations by level of informality. Households located in low-education neighborhoods, as well as households located in high-informality neighborhoods tend to be far from employment subcenters and Metro station.

From these descriptive statistics, we can anticipate two conclusions from the analysis: First, in terms of housing prices, households located in high-education neighborhoods and in neighborhoods with low levels of informality face greater housing prices than households located in low-education or high-informality neighborhoods. These results are evidence of the existence of economic segregation in the city. Second, high-education households and households located in neighborhoods with low levels of informality are better located: both in terms of access to employment subcenters, and proximity to public transport infrastructure. Hence, on top of the economic segregation, spatial segregation and a reduction of opportunities for lower income households are observed in Mexico City.

3. Methodology

We estimate housing price models to measure the effect of both private and public accessibility on housing prices. We split the sample in two subsamples in order to capture different dynamics and compare more homogeneous projects: first, we consider all projects that were built and sold in one year and thus appear in the SOFTEC inventory only once; second, we consider only those projects that took more than one year to get sold. Finally, we pooled both subsamples and create an unbalanced panel in which some projects appear only once in the dataset, while others appear more than once, depending on how long it took to sell all their housing units.[9]

Projects sold in one year and those selling over multiple years have different characteristics, although not all these differences are statistically significant. The most important difference is that projects that are sold within a year are smaller, i.e. have fewer properties than projects that take more time to sell. Projects that sold out in one-year have on average 116 properties, while multiple-year projects have on average 252 properties. Also, projects that are sold within a year have properties with greater median prices, greater number of bedrooms and greater number of bathrooms, although the difference in area size is not significant. Projects that are sold out within a year are on average closer to the Metro and BRT stations than projects that take longer to be fully sold.

To test accessibility impacts on prices for those projects that are sold within a year, we estimate a pooled OLS because we are pooling projects from different years and we control for the housing and community characteristics as well as for year-fixed effects. The equation estimated is the following:

$$\ln P_{ij} = \beta_0 + \sum_{t}^{T} year_t + \sum_{t}^{T} \delta_t A_{it} I(year_t) + \beta_1 DCBD_i + \beta_2 D_i + \beta_3 D_i^2 + \sum_{k}^{K} \beta_k X_{ki} + \sum_{p}^{P} \beta_p Z_{pj} + \varepsilon_i \quad (5),$$

where $\ln P_{ij}$ is the natural log of price of the representative property of the project i located in AGEB j; $\sum_t^T year_t$ are year-fixed effects; D_i is distance to the closest public transportation station; $DCBD_i$ is distance from each development to the CBD; X_{ki} are

[9] In a working paper of this article, we included a repeated-sales estimation considering only the first and the last price reported for each project. Repeated-sales estimators assume that all characteristics remain unchanged through time. Then, characteristics are not included in the regression. However, as noticed in the descriptive section, almost half of projects with price information for more than one year, changed the characteristics of their representative house. Therefore, we decide to control for characteristics every time the project enters the dataset.

housing and development characteristics; and Z_j are characteristics of AGEB j in which the project i is located. The accessibility measure, A_i enters the equation as an interaction term with the year dummies to estimate a different coefficient for each year. We also include the square of distance to the closest public station because being closer to the line could be seen as a disamenity, although the positive effects may be captured by properties that are slightly farther away from the line, yet still close enough to the public station.[10]

The second specification uses an unbalanced panel data with random effects including only projects that took more than one year to get sold. Projects that were sold in only one year are excluded from this estimation because we want to capture the time effects in the valuation of accessibility. Time effects are given because the location and the number of subcenters change every five years (using Census data from 1999, 2004 and 2009), and because new public transportation lines were inaugurated, changing also the distance of some projects to the closest public transport station. This specification is estimated through the following equation:

$$lnP_{ijt} = \alpha_t + \sum_t^T \delta_t A_{it} I(year_t) + \beta_1 DCBD_i + \beta_{2t} D_{it} + \beta_{3t} D_{it}^2 + \sum_k^K \beta_{kt} X_{kit} + \sum_p^P \beta_p Z_{pj} + \varepsilon_{it} \qquad (6),$$

where lnP_{ijt} are the natural logarithm of prices of properties sold in year t belonging to project i located in AGEB j; α_t are time fixed effects; A_{it} are the log of private accessibility indexes entering into the equation as an interaction term with the year dummies; $DCBD_i$ is the distance from each development to the CBD; D_i is the distance from each project i to the closest public transit station (proxy for public accessibility); X_{it} are the characteristics of properties in each project i sold in year t (including the number of properties sold since the first year the community was on sale, and the average number of properties sold monthly in the last trimester of each year); and Z_j are characteristics of AGEB j in which the project i is located.

Finally, the last specification is also a panel in which all projects (those sold in only one year and those that are sold in more than one year) are included. The specification is the

[10] Since not all transportation investments were available every year (new BRT lines were inaugurated and a new Metro line was inaugurated in 2012), public accessibility is not constant through time. See Table 3 for more information.

same as in equation (3) but with all observations in the dataset included. Both panels are estimated using random effects because some variables are different for each project but constant through time (for example, some of the housing and project characteristics, or distance to the CBD).

Heterogeneous effects

Finally, to capture heterogenous effects of accessibility for households with different socioeconomic characteristics, we estimate the same models of equations (2) and (3), but interact the accessibility measure (either private or public) with a categorical variable indicating the education or informality quartile for the area in which the project is located. Both equations (2) and (3) are transformed as follows:

$$\ln P_{ij} = \beta_0 + \sum_t^T year_t + \sum_t^T \delta_t A_{it} I(year_t) + \beta_1 DCBD_i + \sum_1^4 \gamma_j Z_j D_i + \beta_2 D_i^2 + \sum_k^K \beta_k X_{ki} + \sum_p^P \beta_p Z_{pj} + \varepsilon_i \quad (7),$$

and,

$$lnP_{ijt} = \alpha_t + \sum_t^T \delta_t A_{it} I(year_t) + \beta_1 DCBD_i + \sum_1^4 \gamma_j Z_j D_i + \beta_{3t} D_{it}^2 + \sum_k^K \beta_{kt} X_{kit} + \sum_p^P \beta_p Z_{pj} + \varepsilon_{it} \ (8),$$

where all variables are similar to the equations previously specified, but a new interaction between education/informality of neighborhood j (Z_j), and accessibility (private or public-D_i) is included ($Z_j D_i$). This interaction term would provide information regarding the valuation of accessibility that households have depending on the quartile of education/informality that they belong.

4. Results

The results of the three models described above are shown in Table 5, while the results of the heterogeneous effects are shown in Tables 6 and 7 for education and informality quartiles, respectively. Column A in Table 5 shows the first specification in which only one-year-sold projects are included; column B shows the results of the unbalanced panel data with random effects in which only multiple-year sold projects are included; and finally, column C shows the unbalanced panel data results with all projects, in which some of them enter only once in the database (if they were sold in one year) or multiple times, depending on how long they took to sell all their properties. All three specifications cluster the standard errors at the neighborhood level to capture unobservable neighborhood characteristics that influence housing prices.

Table 5: Single-year developments and panel data specifications. Dependent variable: log of housing prices (2000 constant prices).

	Column A	**Column B**	**Column C**
VARIABLES	**One-year observations**	**Repeated observations**	**All observations**
Private accessibility 2002	0.0144*** (0.00442)	-0.00138 (0.00431)	-0.00384 (0.00328)
Private accessibility 2003	0.0241*** (0.00879)	0.00913*** (0.00337)	0.00777** (0.00337)
Private accessibility 2004	-0.00748 (0.0152)	0.00867*** (0.00315)	0.00709** (0.00310)
Private accessibility 2005	0.0248 (0.0321)	0.0124*** (0.00302)	0.0108*** (0.00305)
Private accessibility 2006	0.0224*** (0.00717)	0.0120*** (0.00290)	0.0116*** (0.00293)
Private accessibility 2007	0.0253*** (0.00617)	0.0143*** (0.00289)	0.0141*** (0.00285)
Private accessibility 2008	0.0102 (0.00891)	0.0124*** (0.00263)	0.0120*** (0.00261)
Private accessibility 2009	0.0169** (0.00680)	0.0148*** (0.00304)	0.0151*** (0.00297)
Private accessibility 2010	0.0132** (0.00532)	0.0160*** (0.00311)	0.0164*** (0.00303)
Private accessibility 2011	0.0295*** (0.00582)	0.0188*** (0.00332)	0.0197*** (0.00328)
Private accessibility 2012	0.0247*** (0.00678)	0.0206*** (0.00349)	0.0218*** (0.00345)
Private accessibility 2013	0.0288*** (0.00559)	0.0221*** (0.00378)	0.0242*** (0.00366)
Distance to CBD (in logs)	-0.00367 (0.0303)	0.114*** (0.0343)	0.0649* (0.0337)
Distance to the closest Metro station (in logs)	0.0269 (0.0648)	0.225*** (0.0711)	0.125** (0.0581)
Distance to the closest Metro station squared	-0.00133 (0.00496)	-0.0167*** (0.00510)	-0.00947** (0.00424)
Distance to the closest BRT station (in logs)	0.0335 (0.0494)	-0.0172 (0.0263)	0.000931 (0.0266)
Distance to the closest BRT station squared	-0.00311 (0.00368)	-0.00153 (0.00199)	-0.00205 (0.00194)
Year-fixed effects	yes	yes	Yes
Housing and project	yes	yes	yes

characteristics			
Observations	2,762	8,329	11,091
R-squared	0.817		
Number of projects		2,916	5,678

Clustered standard errors at the neighborhood level in parentheses. Panel data estimations (column B and C) with random effects. Type of property: townhouse, dropped to avoid collinearity. OLS estimation in column A has year fixed effects.
*** p<0.01, ** p<0.05, * p<0.1

At first glance, the results of Table 5 suggest that households value accessibility to employment subcenters, and their preferences are capitalized in housing prices. The valuation of private accessibility is greater in projects that are sold within a year than in projects that spend more than one year on the market. On average, a one percent increase in private accessibility is reflected in an appreciation of housing prices of 2.2 percent for one-year-sold projects, and 1.4 per cent for multiple-year-sold projects. This difference could be the reflection of non-linearities in the valuation of accessibility; since one-year sold projects are on average closer to employment subcenters, private accessibility is more valuable. In most of the years of analysis, private accessibility was significant in explaining changes in housing prices for one-year sold projects, except for 2004, 2005 and 2008.

Distance to the Central Business District, is not statistically significant for explaining housing prices in the single-year communities. However, in the panel specifications, the effect of being closer to the CBD becomes negative (positive coefficient) suggesting that being closer to the center is perceived as a disamenity by homeowners: being one percent farther away from the center appreciates housing prices between 6 and 11 percent.

Public accessibility measured by proximity to a public station has a negative effect on housing prices. Proximity to public transportation is measured by the distance to the closest public transport station as well as by its squared term, following the hypothesis that being too close to the station or to the line could be also seen as a disamenity because of the noise, congestion and crime that the public transportation could generate. For multiple-year sold projects and for the sample containing all projects in the market, being one percent closer to a Metro station is reflected in a depreciation of housing prices of 22 percent and 12 percent, respectively.

While these coefficients are positive (suggesting a negative relationship between proximity to the Metro station and appreciation of housing prices), their quadratic terms are negative and significant for columns B and C models. These negative signs on the quadratic forms suggest that the slope of the coefficient changes as the estimation involves properties that are located farther away from the station. However, the quadratic term is smaller in size than the singular term suggesting that the negative perception of proximity to a public station is greater than its benefits.[11]

These results suggest that Mexico City is perceived by households as a polycentric urban area, in which proximity to local employment subcenters is more important than proximity to the historic downtown. Households value accessibility to the subcenters (either by private transportation or buses), but being located close to a public transit station, either BRT or Metro, seems not to be as important. This result could have several explanations: the negative coefficient of public accessibility could be reversed if we control for crime and other negative externalities that public stations bring (however, this information is not available). Alternatively, it can also be explained because the transportation system does not efficiently connect households to their sites of employment, or because households have a negative perception of the public transportation structures available in the city.

Because the value that is assigned to different types of accessibility may change by type of household, with households that work in the formal sector valuing accessibility to job centers more than households that work in the informal sector—and hence do not necessarily seek job opportunities in the formal employment centers—we estimate heterogeneous effects of accessibility on housing prices. Tables 5 and 6 show the valuation of private and public accessibility for different types of households, categorized by the proportion of informal workers living in the neighborhood where the project is located, and by the average years of education of the neighborhood where the project is located. Three models are estimated in each case, similar to those in Table 4, and are presented in columns A, B and C of Tables 6 and 7, respectively. Two different regressions are estimated for

[11] Since private accessibility is measured by an exponential distance decay function, it cannot be compared to the public accessibility measures. In the Online Appendix we estimate private accessibility as the average distance to employment subcenters. Furthermore, we categorize subcenters in stable and new subcenters to measures heterogenous effects. Estimations are conducted for log-log and log-nonlog specifications. Results are robust for all specifications and are available upon request.

each case. The first includes as the independent variable the interaction between informality/education quartiles and private accessibility; the second includes the interaction between informality/education quartiles and public accessibility, measured by distance to the closest Metro station.

The results corroborate findings from the previous literature regarding the valuation of location for different types of households. Employment subcenters used for calculating private accessibility considered only formal employment. Hence, one would expect that housing prices of projects located in neighborhoods with lower level of informality are more sensible to proximity to employment subcenters, than prices in projects located in neighborhoods with greater level of informality. The results confirm this hypothesis: for projects that are sold within a year, while one percentage point increase in private accessibility is reflected in an appreciation of housing prices of 2.2 percent in the lowest quartile of informality, this figure is only 1.1 percent in the highest percentile of informality, and not significant in the third quartile (with a negative sign). In the case of multiple-year sold projects and considering all projects in the sample, the appreciation of housing prices for being closer to employment subcenters doubles when comparing projects located in neighborhoods with high proportion of informality and projects located in neighborhoods with low proportion of informality.

However, when analyzing which households value more being closer to a Metro station, we find that neither households located in high-informality neighborhoods nor households located in low-informality neighborhoods value being closer to a Metro station. In all informality quartiles, being closer to a Metro station is considered a disamenity. In the case of projects that are sold within a year, proximity to a Metro station does not have an effect on housing prices. For multiple-year sold projects, proximity to a Metro station is always a disamenity for all quartiles of neighborhood informality. On absolute terms, the effect is greater for projects located in neighborhoods with a greater proportion of informal workers. The results are consistent when all projects are included in the estimation.

Table 6: Heterogeneous effects categorizing the sample by the proportion of informal workers in each neighborhood where the project is located. The table only shows the coefficients of the interaction between informality quartile and private/public accessibility.

	Column A	**Column B**	**Column C**
	One-year	**Repeated**	**Panel all obs**

	observations	observations	
Regression including the interaction between private accessibility (to employment subcenters) and informality quartiles			
Private accessibility in 1st quarter of informality	0.0226*** (0.00570)	0.0180*** (0.00304)	0.0172*** (0.00316)
Private accessibility in 2nd quarter of informality	0.0252*** (0.00406)	0.0169*** (0.00354)	0.0141*** (0.00341)
Private accessibility in 3rd quarter of informality	-0.0135 (0.00827)	0.0124*** (0.00351)	0.00783** (0.00380)
Private accessibility in 4th quarter of informality	0.0118*** (0.00429)	0.00865*** (0.00251)	0.00947*** (0.00261)
Regression including the interaction between public accessibility (to a Metro station) and informality quartiles			
Public accessibility in 1st quarter of informality	0.0119 (0.0671)	0.238*** (0.0684)	0.106* (0.0568)
Public accessibility in 2nd quarter of informality	0.0178 (0.0639)	0.250*** (0.0679)	0.126** (0.0564)
Public accessibility in 3rd quarter of informality	0.106 (0.0671)	0.268*** (0.0676)	0.150*** (0.0559)
Public accessibility in 4th quarter of informality	0.0336 (0.0638)	0.272*** (0.0680)	0.137** (0.0563)

Clustered standard errors at the neighborhood level in parentheses. Panel data estimations (column B and C) with random effects. OLS estimation in column 1 has year fixed effects. Different regressions for private and public accessibility. All regressions include the same control variables as in Table 4.
*** p<0.01, ** p<0.05, * p<0.1

When the heterogeneous effects are analyzed by levels of education of neighborhoods where the projects are located, results are mixed (results shown in Table 7). According to the literature, private accessibility would be more valuable for households with higher levels of education, while public accessibility would be more valuable for low-education households; then, the Metro system could be considered an inferior good. Our results suggest that projects that are located in low-education neighborhoods value private accessibility less than projects that are located in high-education neighborhoods. For example, for projects located in low-education neighborhoods, one percent improvement in private accessibility increases housing values between 1.2 and 1.8 depending on the specification analyzed; this range increases to 1.6-1.9 percent for projects located in high-educated neighborhoods.

The differences between education quartiles are very small when analyzing the effect of proximity to the Metro station on housing prices. When only one-year sold projects are considered, proximity to a Metro station does not have any impact on housing prices. In the other two estimations (multiple-year sold projects and all projects in the sample), the negative effect of proximity to the Metro station varies between 11 and 23 percent for low-education neighborhoods, and between 10 and 23 percent for high-education neighborhoods.

Table 7: Heterogeneous effects categorizing the sample by the level of education in each neighborhood where the project is located. The table only shows the coefficients of the interaction between education quartiles and private/public accessibility.

	(3)	(1)	(5)
	One-year observations	**Repeated observations**	**Panel all obs**
Regression including the interaction between private accessibility (to employment subcenters) and education quartiles			
Private accessibility in 1st quarter of education	0.0184*** (0.00393)	0.0140*** (0.00331)	0.0124*** (0.00284)
Private accessibility in 2nd quarter of education	0.0183*** (0.00687)	0.0142*** (0.00322)	0.00331 (0.00339)
Private accessibility in 3rd quarter of education	0.0209*** (0.00512)	0.00911*** (0.00308)	0.0134*** (0.00252)
Private accessibility in 4th quarter of education	0.0197*** (0.00707)	0.0163*** (0.00278)	0.0173*** (0.00314)
Regression including the interaction between public accessibility (to a Metro station) and education quartiles			
Public accessibility in 1st quarter of education	0.0415 (0.0667)	0.235*** (0.0730)	0.110** (0.0558)
Public accessibility in 2nd quarter of education	0.00282 (0.0675)	0.233*** (0.0703)	0.148** (0.0586)
Public accessibility in 3rd quarter of education	0.0295 (0.0681)	0.254*** (0.0722)	0.121** (0.0555)
Public accessibility in 4th quarter of education	0.0384 (0.0681)	0.230*** (0.0732)	0.103* (0.0563)

Clustered standard errors at the neighborhood level in parentheses. Panel data estimations (column B and C) with random effects. OLS estimation in column 1 has year fixed effects. Different regressions for private and public accessibility. All regressions include the same control variables as in Table 4.
*** p<0.01, ** p<0.05, * p<0.1

Conclusions

In this paper, we estimate the effect of accessibility on housing prices in Mexico City. We consider private accessibility as proximity to employment subcenters, and public accessibility as proximity to public transit stations. Furthermore, we estimate heterogeneous effects for households located in neighborhoods with different education levels and neighborhoods with different proportions of informal workers. We estimate this effect using different specifications: an OLS estimation of all projects that were sold within a year, a panel specification including only projects that took more than one year being sold, and a panel specification considering all projects in the sample.

Residents in Mexico City value proximity to employment subcenters as an amenity and such value is capitalized on housing prices. For all specifications, being one percent closer to employment subcenters (and controlling for their employment level) increases housing values between one and almost three percentage points. As employment subcenters earn importance in the spatial structure of the city, the Central Business District loses significance. In two of the three specifications analyzed, being closer to the Central Business District is perceived as a disamenity, negatively affecting housing prices. Efforts to revive the center of the city would be required to change the current negative perception captured in the negative effects on house prices.

On the other hand, proximity to public transit stations has a different effect on housing prices. While being closer to a BRT station does not have an effect on housing prices, being closer to a Metro station is considered a disamenity. The coefficient of the quadratic term is of opposite sign, suggesting that housing prices increase with distance to the stations at a decreasing rate.

This negative perception of public transport is problematic for a city the size of Mexico City, where congestion and pollution are exacerbated by the common use of private vehicles -over 20 percent of the population commute by private vehicles compared to about 12 percent in a city like Tokyo (Institute for Transportation and Development Policy, 2012). These results are relevant for policy makers and urban planners when planning future public and private infrastructure projects. The fact that private accessibility is more important than public accessibility (measured by distance to a public transport station) poses a challenge for cities that experience fast and unplanned growth. The widespread use of private vehicles (including individual and collective transport), despite the existence of Metro and BRT systems, is not sustainable, even if new roads and highways are built. As polycentricity of cities increases policymakers should think about policies that allow people to locate closer to their jobs and find ways to make public transit options more attractive to all residents. Coordinated land use and transport policies are an important step in this direction. Policies that put emphasis on the revival of areas in and around public transport stations could could make transit use and nearby housing more attractive.

Further, our analysis considering heterogeneity of household types suggests clear differences in preferences across neighborhoods. Given the lack of information on household income, we use education and job informality as proxies for income levels in the areas analyzed. Descriptive statistics show that areas with high-educated households or neighborhoods with low proportion of informal workers face higher housing prices (evaluated at the mean), suggesting socioeconomic segregation within the city. Moreover, these households not only live in high-value dwellings but are also better located in terms of accessibility to employment subcenters and proximity to public transport. These results are also evidence of the existence of spatial segregation in the city. Disadvantaged households then, are segregated; not only because of their economic conditions, but because they are located farther away from economic centers with less access to employment opportunities. Although estimations suggest homogeneous effects in terms of the negative perception of public transport, households living in areas with lower average education levels and higher informality value being closer to formal employment subcenters less than households in areas with higher average education and lower job informality levels.

This analysis raises a series of questions regarding segregation and accessibility to jobs that points at an important area for future research. Further analyses could complement these findings with more detailed information on informal employment. Since subcenters in this paper are estimated using only formal employment, we cannot draw conclusions regarding the effect of accessibility to informal employment centers on housing prices. Further, jointly considering the constraints in transport and housing accessibility would be essential to better understand the constraints and tradeoffs households in the lower income brackets face as they choose where to live and where to work.

References

Adair, A., McGreal, S., Smyth, A., Cooper, J., and Ryley, T. (2000). House prices and accessibility: The testing of relationships within the Belfast urban area. *Housing studies*, *15*(5), 699-716, doi: https://doi.org/10.1080/02673030050134565

Agostini, C.A., and Palmucci, G.A. (2008). The Anticipated Capitalisation Effect of a New Metro Line on Housing Prices. *Fiscal studies*, *29*(2), 233-256, doi: 10.1111/j.1475-5890.2008.00074.x.

Ahlfeldt, G. (2011). If Alonso was right: modeling accessibility and explaining the residential land gradient. *Journal of Regional Science*, *51*(2), 318-338, doi: 10.1111/j.1467-9787.2010.00694.x.

Aguilar, A. (2011). Metropolitan transformation and polycentric structure in Mexico City: Identification of urban subcenters 1989-2009. Presented in the IGU Urban Commission Meeting 2011.

Aguilar, A.G., & Alvarado, C. (2005). La reestructuración del espacio urbano de la Ciudad de México. ¿Hacia la metrópoli multinodal?. *Procesos metropolitanos grandes ciudades. Dinámicas recientes en México y otros países*, 265-308.

Aguilar, A.G., & Lozano, J.H. (2014). Metropolitan transformation and polycentric structure in Mexico City: Identification of urban subcenters, in Donoghue, D.P.(ed). *Urban transformations: Centres, peripheries and systems*, Taylor and Francis, New York.

Atuesta L. & Ibarra-Olivo J.E. (2017). Polycentricism in Mexico: an estimation of employment subcenters in three metropolitan areas, *mimeo.*

Avendano Arosemena, J. (2013). Three essays on urban spatial structure in Bogota DC. Doctoral thesis. Universitat Autonoma de Barcelona, url. https://ddd.uab.cat/record/113051

Azuara, O., & Marinescu, I. (2013). Informality and the expansion of social protection programs: Evidence from Mexico. *Journal of health economics*, 32(5), 938-950, doi: https://doi.org/10.1016/j.jhealeco.2013.07.004

Cervero, R., Rood, T., and Appleyard, B. (1999). Tracking accessibility: employment and housing opportunities in the San Francisco Bay Area. *Environment and Planning A*, *31*(7), 1259-1278, doi: https://doi.org/10.1068/a311259.

Crôtte, A., Noland, R.B., and Graham, D.J. (2009). Is the Mexico City Metro an inferior good?. *Transport Policy*, *16*(1), 40-45, doi: https://doi.org/10.1016/j.tranpol.2009.02.009.

Excelsior (2017). "Gutiérrez: Por contingencia dejan de circular millón y medio de autos". May 16th, 2017. http://www.excelsior.com.mx/comunidad/2017/05/16/1163733 (accessed on July 7th, 2017).

Fernández-Maldonado, A.M., Romein, A., Verkoren, O., & Parente Paula Pessoa, R. (2014). Polycentric structures in Latin American metropolitan areas: Identifying employment sub-centres. *Regional Studies*, 48(12), 1954-1971, doi: https://doi.org/10.1080/00343404.2013.786827

Flores Dewey, O. (2010). The Value of a Promise: Estimating the Effects of a Promise to Build Mass Transit Infrastructure on Housing Prices in Ecatepec, Mexico. *The Lincoln Land Institute*, url: http://www.brt.cl/wp-content/uploads/2011/01/FYP-Working-draft-June1.pdf.

Guerra, E. (2014). Mexico City's suburban land use and transit connection: The effects of the Line B Metro expansion. *Transport Policy*, *32*, 105-114, doi: https://doi.org/10.1016/j.tranpol.2013.12.011

Hansen, W.G. (1959). How accessibility shapes land use. *Journal of the American Institute of Planners*, *25*(2), 73-76, doi: https://doi.org/10.1080/01944365908978307

ILO (2003) *Guidelines concerning a statistical definition of informal employment, endorsed by the 17th International Conference of Labour Statisticians*, 2003. http://www.ilo.org/wcmsp5/groups/public/---dgreports/---dcomm/---publ/documents/publication/wcms_222979.pdf

Institute for Transportation and Development Policy (2012). "Transforming Urban Mobility in Mexico. Towards accessible cities less reliant on cars.". March (2012) http://mexico.itdp.org/wp-content/uploads/Transforming-Urban-Mobility-in-Mexico.pdf

INEGI (2007). *Encuesta Origen Destino de los viajes de los Residentes de la Zona Metropolitana del Valle de México.* Mexico: Instituto Nacional de Estadística y Geografía.

[dataset] INEGI (1999) Censo Económico 1999, url. http://www.inegi.org.mx/est/contenidos/proyectos/accesomicrodatos/ce1999/default.aspx

[dataset] INEGI (2004) Censo Económico 2004, url. http://www.inegi.org.mx/est/contenidos/proyectos/accesomicrodatos/ce2004/default.aspx

[dataset] INEGI (2009) Censo Económico 2009, url. http://www.inegi.org.mx/est/contenidos/proyectos/accesomicrodatos/ce2009/default.aspx

[dataset] INEGI (2010) Censo de Población y Vivienda, url. http://www.inegi.org.mx/est/contenidos/proyectos/accesomicrodatos/cpv2010/default.aspx

Martínez, L., and Viegas, J. (2009). Effects of transportation accessibility on residential property values: Hedonic Price Model in the Lisbon, Portugal, metropolitan area. *Transportation Research Record: Journal of the Transportation Research Board*, (2115), 127-137, doi: https://doi.org/10.3141/2115-16

McMillen, D.P. (2001). Nonparametric employment subcenter identification. *Journal of Urban Economics*, *50*(3), 448-473, doi: https://doi.org/10.1006/juec.2001.2228

McMillen, D.P. (2003). The return of centralization to Chicago: using repeat sales to identify changes in house price distance gradients. *Regional Science and Urban Economics*, *33*(3), 287-304, doi: https://doi.org/10.1016/S0166-0462(02)00028-5

Milenio (2016). "No circulan hoy más de un millón de vehículos. March 27th, 2016, url: http://www.milenio.com/politica/no_circulan_vehiculos-hoy-millon-danos_salud_CdMx-contingencia_ambiental_en_CdMx-EPN_0_702529771.html (accessed on July 7th, 2017).

Monkkonen, P. (2015) The Role of Housing Finance in Mexico's Vacancy Crisis. UCLA Luskin School of Public Affairs, url. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2765658

Munoz-Raskin, R.C. (2007). Walking accessibility to bus rapid transit in Latin America: does it affect property values? The case of Bogotá, Colombia. In *Transportation Research Board 86th Annual Meeting* (No. 07-1138), url. https://trid.trb.org/view.aspx?id=801511

Negrete, M.E., & Paquette Vassalli, C. (2011). La interacción entre transporte público y urbanización en la Zona Metropolitana de la Ciudad de México: un modelo expansivo que llega a sus límites. t*erritorios*, (25): 15-33, ISSNe: 2215-7484.

Perdomo, J.A. (2011). A methodological proposal to estimate changes of residential property value: Case study developed in Bogota. *Applied Economics Letters*, *18*(16), 1577-1581, doi: https://doi.org/10.1080/13504851.2011.554360

Romein, A., Verkoren, O., & Fernandez, A.M. (2009). Polycentric metropolitan form: application of a 'Northern'concept in *Latin America. Footprint*, 3(2), 127-146, doi: http://dx.doi.org/10.7480/footprint.3.2.712

Sanchez Trujillo, V. (2012). El impacto de los subcentros de empleo sobre la densidad de población en la zona metropolitana del Valle de México. Master's thesis. Universitat Autónoma de Barcelona.

Shiller, R.J. (1991). Arithmetic repeat sales price estimators. *Journal of Housing Economics*, *1*(1), 110-126, doi: https://doi.org/10.1016/S1051-1377(05)80028-2

[dataset] SOFTEC (2013). Informe tendencial del mercado inmobiliario Ciudad de México (microdata), url. https://www.softec.com.mx/web/index.php/herramientas/reporte-tendencial2

Song, S. (1996). Some tests of alternative accessibility measures: A population density approach. *Land Economics*, *72(4)*, 474-482, doi: 10.2307/3146910

Suárez, M. (2007) Mercados de trabajo y 1ocalización residencial en la ZMCM. PhD thesis, Universidad Nacional Autónoma de Mexico

Suárez-Lastra, M., & Delgado-Campos, J. (2007). Estructura y eficiencia urbanas. Accesibilidad a empleos, localización residencial e ingreso en la ZMCM 1990-2000. *Economía, sociedad y territorio*, 6(23), url. http://www.redalyc.org/articulo.oa

Suárez, M., & Delgado, J. (2009). Is Mexico City polycentric? A trip attraction capacity approach. *Urban Studies*, 46(10), 2187-2211, doi: https://doi.org/10.1177/0042098009339429

Suárez, M., Murata, M., & Delgado Campos, J. (2016). Why do the poor travel less? Urban structure, commuting and economic informality in Mexico City. *Urban studies*, *53*(12), 2548-2566, doi: https://doi.org/10.1177/0042098015596925

Vizuet, G.I. (2010). Política de vivienda y movilidad residencial en la Ciudad de México. *Estudios Demográficos y Urbanos 2(74)*, 277-316, url. http://www.jstor.org/stable/20787554

Wallace, N.E., & Meese, R.A. (1997). The construction of residential housing price indices: a comparison of repeat-sales, hedonic-regression, and hybrid approaches. *The Journal of Real Estate Finance and Economics*, *14*(1-2), 51-73, doi: https://doi.org/10.1023/A:1007715917198

Wilson, A.G. (1967). A statistical theory of spatial distribution models. *Transportation research*, *1*(3), 253-269, doi: https://doi.org/10.1016/0041-1647(67)90035-4

Yoshida, N., & Deichmann, U. (2009). Measurement of accessibility and its applications. *Journal of Infrastructure Development*, *1*(1), 1-16, doi: https://doi.org/10.1177/097493060900100102