UNIVERSIDAD EAFIT
Abierta al mundo

Escuela de Economía y Finanzas

Documentos de trabajo
Economía y Finanzas

Centro de Investigación
Económicas y Financieras

No. 13-25
2013

The Importance of Community Attributes in Household Residential Location Decisions

Ospina, Mónica; Bohórquez, Santiago;
Serna, Andrea; Castañeda, Laura

cief
Centro de Investigaciones
Económicas y Financieras
Universidad EAFIT

The Importance of Community Attributes in Household Residential Location Decisions

Ospina, Monica[*]

Bohórquez, Santiago

Serna, Andrea

Castañeda, Laura

Abstract:

This study identifies how community attributes affect household residential location decisions in Medellin, Colombia. The empirical model applies the revealed preference principle: each household is assumed to have made an optimal location decision given a set of alternatives. Using household data, we estimate a conditional logit choice model for residential communities by controlling for both individual and neighborhood characteristics, including environmental attributes. The set of alternatives for each household are defined using the applicable neighborhood's socioeconomic and geographic characteristics. The results provide an estimate of household preferences for the many characteristics of the potential choices in the choice set. In the case of Medellin, we found positive and significant preferences for public provided goods such as public schools and security but relatively low preferences for recreational and cultural spaces; households prefer that the latter be provided by the private sector.

JEL Classification: R21; R23; Q50

Key Words: Housing Demand; Neighborhood Characteristics; Environmental Economics

[*] Universidad EAFIT, mospina6@eafit.edu.co

1.   Introduction

In the early 1960s, many economists began to study the microeconomic and spatial aspects of housing markets in urban areas, the behavior of housing producers, the provision of public services by local governments and the economics of residential location (Ingram, 1977). One of the most notable theoretical advances in residential location was produced by Tiebout (1956), who suggested that under certain conditions consumers might reveal their true preferences for locally provided public goods. This has motivated several authors to study the importance of community attributes in the residential location decision (Friedman 1981, Nechyba and Strauss 1997, Ozturk and Irwin 2001 and Ferguson et al 2007).

According to the revealed preference model, a household deciding the optimal location for its house takes into account all possible alternatives and chooses the option that generates the maximum level of utility. In evaluating the alternatives, the utility of the household and the family and dwelling characteristics will be affected by local public goods, environmental services and neighborhood characteristics.

Understanding the causes that lead to a population concentration in certain communities is not only academically interesting, but it also has profound implications for social well being and public policy design. Knowing how a community's attributes have contributed to its population concentration and what amenities are more appreciated by households can guide both public provision and conservation of residential goods and services.

This paper identifies preferences for community amenities or attributes using the residential location of households in Medellin in 2009. To do so, a conditional logit model is estimated, controlling for attributes of individuals and neighborhood characteristics, including environmental elements. The data employed are from the *Encuesta de Calidad de Vida* (ECV, survey of life quality) for 2009 and from the Área Metropolitana environmental authority for the Valle de Aburrá, where Medellín is located.

Medellin is the second ranked city, in terms of GDP, in Colombia and is the capital of the department of Antioquia. Although 72% of its territory is rural, 90% of the population lives in urban areas. The urban portion of the city is divided into 16 areas, called *comunas*, which are divided into 250 neighborhoods. These neighborhoods are the areas of focus in this study.

Medellin is a city that exhibits socioeconomic divergence on a geographic level; in the northern part of the city, there is a concentration of low-income households, young people, criminal bands and drug sales. However, the southern area of the city contains households with higher income and a greater proportion of the population that is over 60 years old; it also boasts the majority of the employment opportunities and a greater sense of security (Duque, 2011).

This paper is organized as follows: Section 2 presents an overview of the related literature. Section 3 explains the methodology used, while Section 4 describes the data collected. Finally, Section 5 presents the results and main conclusions of the paper.

2.   Related literature

Tiebout (1956) first issued the hypothesis that households regard the residential location decision as choosing a particular package of local public goods and services over other such packages. Tiebout's model suggests that households reveal their preference for public goods and services when they *"vote with their feet"* through choosing where to live. Additionally, it establishes that households in a given community derive the same marginal benefits from local public goods and services.

In the model, houses are considered a spatially fixed good. When a household purchases a dwelling, it obtains access to, not only the house but also a series of community goods and services. Therefore, the value of the house depends both upon the structural attributes of the house and the amenities and services available in its location. One of the limitations of Tiebout's theory is that it assumes that households have the ability to choose between a great number of feasible alternatives; in reality, this is improbable because of cost and location limitations (Ozturk et al, 2001).

As a complement to the theoretical model developed by Tiebout, McFadden (1974) formulated an econometric model that allowed modeling decisions with qualitative options. In this model, the selection objects, sets of available alternatives, individual selection and behavior patterns of the population characterized the decisions of individuals. McFadden (1978) applied this model to the choice of residential location and established that individuals weighted the attributes of each alternative in making a decision.

Friedman (1981), based on Tiebout's theory and McFadden's econometric model, used a conditional logit model to test the effects of local public services and community attributes on residential choice. Friedman used data from 29,000 houses located in the San Francisco bay area in 1965. The results of the study suggested that local public services play a minor role in residential choice; the major determinant is the quantity of housing services that the household can obtain in the community. In addition, parks and recreation services were found to positively affect the location decision, while longer time and distance to the workplace and the felony crime rate negatively affected the decision.

Nechyba and Strauss (1997) used a discrete choice approach to estimate the impact of local fiscal and other variables on individual community choices. Their dataset encompassed 90% of the homeowners in six school districts in Camden County, New Jersey. Using a random utility model and a mixed polytomous choice model, they found that public goods have a significant impact on community choice. Likewise, investment in education, community entry prices, degree of commercial activity and distance to the metropolitan area also appeared relevant to the location decision. A high violent crime rate and higher marginal housing prices decreased the probability of choosing a particular community.

Ozturk and Irwin (2001) used the housing and migration data from 823 households in Franklin County, Ohio in 1995 to estimate the probability that a household decides to stay in or move away from a school district. Using a spatial probit model, they concluded that the quality of the schools, criminality rates and presence of children were important determinants in the households' decisions.

In searching for alternative explanations, many articles have included environmental amenities of the community as explanatory variables. Earnhart (2001) estimated the aesthetic benefits generated by the presence and the quality of environmental services associated with residential location in Fairfield, Connecticut. The study employed two applications of discrete–choice hedonic analysis of revealed data and choice-based conjoint analysis of stated data. The study found that environmental services increase the utility of individuals.

Ortúzar and Rodríguez (2002) used a stated preference model to estimate willingness to pay for a reduction of atmospheric pollution in Santiago de Chile. To that end, a survey of 107

families was conducted to assess 10 residential options with varying attributes. A multinomial logit was used to estimate that individuals were willing to expend 1.14% of their family income to improve air quality.

Other articles have approached the subject from a wider perspective, taking into account amenity and economic variables as well as control variables, including agglomeration, demographics, human capital and geographic and social capital. One such article, written by Ferguson et al. (2007), examined population growth in Canada between 1991-2001, accounting for the urban and rural communities separately, as well as five age groups. Using weighted least squares and spatial-error model; the study concluded that variable groupings have different effects on rural and urban communities. For example, in urban communities, amenities and economic factors appear to be about equally important; however, in rural communities, economic factors dominate.

3.   Methodology

To determine the effects of the environmental and other amenities in the different neighborhoods on the spatial location of families, we used a random utility model, such as the one proposed by McFadden (1978). When families face a decision on the optimal place to live, they examine all possibilities and choose the option that maximizes their utility. This model only considers demand because it assumes that the residential location of one family does not affect the housing market or behavior of the local government. This model was later used by Friedman (1981), whose principal hypothesis was that when choosing a residential location, families reveal their preferences toward amenities and characteristics of the house.

Following the work conducted by Friedman (1981), we estimate a conditional logit model where the units of decision are the neighborhoods that households choose to live. The model assumes that families make a choice from several similar alternatives; therefore, we apply a cluster aggrupation of the neighborhoods and eliminate those that are not part of the same cluster as possible destinations. All of the alternatives are assigned the same probability of being chosen.

The cluster identification is performed using the k-means method and using socioeconomic variables from the 2005 census. These variables included the percentage of households

without sons, percentage of single-head households, population density, ratio of children that attend public schools over the total number of kids that attend school and the distance from the center of the neighborhood to the Medellin River (one side was assigned positive distances and the other negative because the Medellin River divides the city in two, and mobility from one side of the river to the other is not common). The spatial distribution of these variables is shown in Figure 1.

**Figure 1. Map representation of cluster variables**
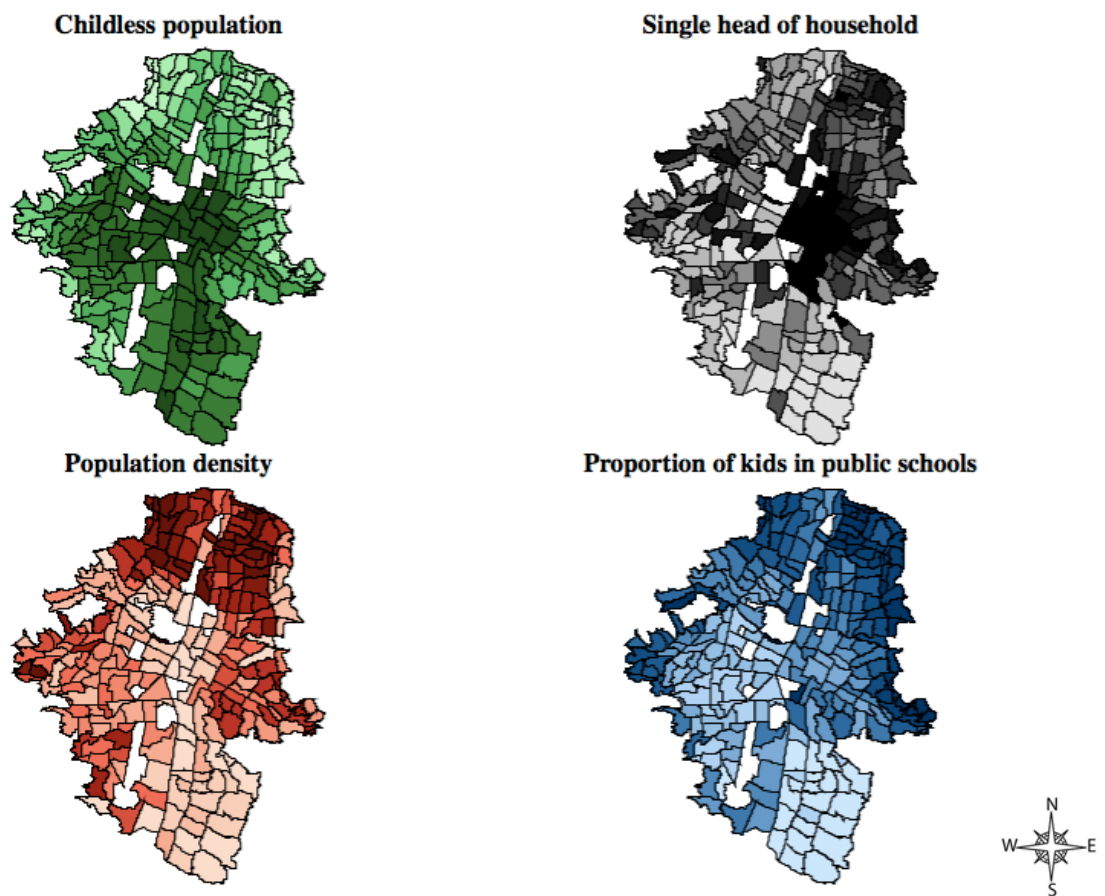
Source: based on 2005 Census data, 2013.

Note: A greater intensity of color is an indicator of a greater proportion of population that fulfills the characterization.

With these variables 12 clusters were formed, as shown in Figure 2. It is important to clarify that areas appearing in white were eliminated from the study because of a lack of information. Such areas include 2 neighborhoods (Tenche and Blanquizal, which are mostly industrial
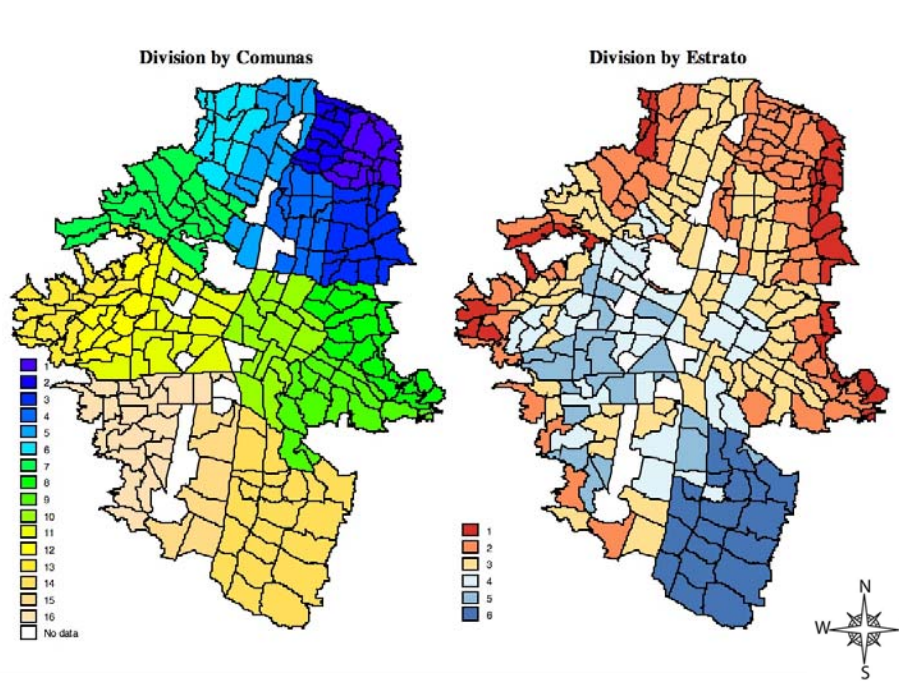
neighborhoods) and 22 special areas, such as universities, military bases and recreational areas that have a wide range of features and affect estimations at a neighborhood level.

Figure 2. City of Medellin clusters according to selected socioeconomic variables
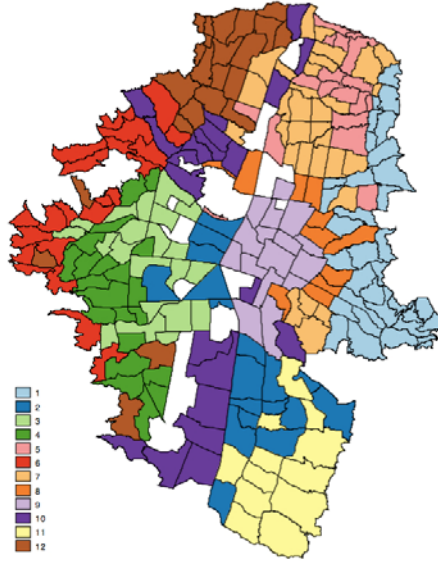
Source: Based on 2005 Census data, 2013

Figure 3. City of Medellin divided by *Comunas* and *Estrato*

Source: Based on data from Area Metropolitana, 2013

Along with these variables, we used the 2009 ECV for Medellin, and chose only households that have moved in the past 5 years to identify residential location decisions that were made given recent community attributes. Each of these households was randomly assigned 4 possible neighborhoods in the same cluster as the neighborhood in which they were located. This was repeated 1000 times to avoid biased results. To test the validity of the results, we estimated three additional models: first, we used the *comunas* of Medellin instead of the cluster division; second, we used the most common social *estrato* of the neighborhoods (the *estrato* is a scale, ranging from 1 to 6, applied to households that depends on house value and size; a higher *estrato* yields higher utility taxes), see Figure 3; and finally, a model in which only 3 possible alternatives were given was employed as a method of testing the consistency of our results.

4. Data

The data from the study were drawn from both the ECV and the data provided by Área Metropolitana. The data were divided into three categories: local public goods, environmental services and neighborhood characteristics. All the data were at the neighborhood level for 2009 (Table 1).

Table 1. Independent variables included in the models

	Variables
	Number of public schools
Local public goods	Total recreational and cultural space (in m^2)
	Private recreational and cultural space (in m^2)
Environmental services	Green zone (in m^2)
	Number of flood episodes
	Number of mass movements
	Socioeconomic *estrato*
Neighborhood characteristics	Homicide rate in 2008
	Size of the neighborhood (in m^2)
	Population density

Note: Variables included in the model.

To better understand the model, some clarification of these variables is necessary. Recreational and cultural spaces are places that aim to generate social and urban transformation; those provided by the local government are usually located in marginal areas of the city that have a greater deficit of cultural and recreational spaces. Green zones are calculated in total m^2 by neighborhood and by *estrato* of each neighborhood to obtain a more precise measure of the valuation among households. Finally, the *estrato* is a scale ranging from 1 to 6 that depends on the value and size of a house and applies a socioeconomic rating to the household. In the model, the *estrato* is divided into three qualitative categories: low (1-2), medium (3-4) and high (5-6).

5.   Results and discussion

This section examines the empirical results of the conditional logit estimation. Each of the neighborhood alternatives is described in terms of local public services, environmental services and neighborhood characteristics. The results are shown in detail in Table 2. The local public services show mixed results. However, the public schools have a positive and significant result in the models, which demonstrates the interest of households in public schools when choosing a residential location. Furthermore, the total recreational and cultural spaces have a negative result; however, when only the private spaces are considered, the result becomes positive. This might be explained by the location of the recreational areas: while the spaces provided by the local government are located in marginal areas of the city, the private spaces are located in the most valued areas.

On the other hand, most of the environmental services have unexpected results. The mass movements and floods both have a positive sign; however, this does not imply that households value these natural disasters as positive elements of their residential location. Instead, this result may arise because of the expansion of urban sprawl that has led to the conception of peripheral areas as sites for construction and a disregard of their tendency to suffer natural disasters. In regard to the green zones, they yield a negative and significant result. Nevertheless, when the quantity is separated by *estrato*, these results change, yielding a positive result for households of medium and high socioeconomic levels. This may be explained by different conceptions of a green zone in neighborhoods of different social *estrato*. Whereas medium and higher *estrato* view green zones as recreational spaces that

improve the aesthetics of dwellings, the lower *estrato* considers them places where bad habits, such as drug sales or consumption, tend to develop.

Finally, with regard to neighborhood characteristics, the results show that the predominant *estrato* of the neighborhood, taken as an indicator of the socioeconomic characteristics of the population, has a negative coefficient in the included models. This might be explained by the similarities among the neighborhoods included in the divisions: households may want to live in houses with a lower *estrato* to avoid higher utility taxes, but that have good amenities available to them. The homicide rate has, as established in the literature, a negative sign; this means that households consider security an important attribute of the neighborhood where they choose to live (Table 2).

Table 2. Results of Conditional Logit Estimation

Variables	Cluster	*Comunas*	*Estrato*	Alternatives
Middle *estrato*	-0.3908889***	-0.1406294***		-0.1379203***
	0.036391	0.0326311		0.0393598
High *estrato*	-0.1605304***	-0.0337933		-0.0336736
	0.0331626	0.0308531		0.0364025
Green areas (in m^2)	-0.4573521***	-0.7840645***	-1.308383***	-0.7862267***
	0.0628104	0.0709944	0.2433062	0.0807637
Green areas 2	0.3434983***	0.5823721***		0.5817217***
	0.0654237	0.0747838		0.0859155
Green areas 3	1.980234***	1.959182***		1.995377**
	0.1585869	0.1707383		0.2043365
Total recreation spaces (in m^2)	-3.208348***	-2.578018***	-3.423162***	-2.540376***
	0.3208189	0.3007568	0.3480337	0.3559925
Private recreation spaces (in m^2)	3.103219***	1.766035***	2.074989***	1.628457***
	0.4286478	0.4155785	0.4630081	0.4959688
Homicide rate 2008	-0.1854271***	-0.1860109***	-0.1967371***	-0.1862581***
	0.0088548	0.0091605	0.0105872	0.0111718
Mass movements	0.0019333***	0.0018741**	0.0030471***	0.001785*
	0.0006845	0.0007789	0.0009088	0.0009317
Size of the neighborhood (in m^2)	0.00000149***	0.00000161***	0.00000158***	0.00000164***
	4.52E-08	5.24E-08	6.25E-08	0.000000065
Flood episodes	0.0012235	0.002851***	-0.0025091**	0.0025768**
	0.0010081	0.0010043	0.0011724	0.0012159
Schools	0.033894***	0.0485528***	0.0429973***	0.0480255***
	0.0066823	0.0064001	0.0073163	0.007998

Population density	29.29744***	20.46196***	21.73822***	20.73326***
	1.235535	0.7889331	0.968173	0.9434489

+b/ sd
+* 0.1 ** 0.05 *** 0.01

6. Conclusions

This paper aimed to identify preferences for community amenities or attributes using the residential location of households in Medellin in 2009. To achieve this, a conditional logit model was estimated that controlled for individual attributes and neighborhood characteristics, including environmental elements. Data from the ECV for 2009 and from the Área Metropolitana were also employed.

The results show a positive value for public provided goods, such as public schools and security, but relatively low values for recreational and cultural spaces, with households tending to prefer those that are privately provided. This is relevant in terms of the investments made in Medellin in local public goods during recent years. Understanding how these spaces affect the residential location decision of households can guide the city's public expenditure decisions.

The results also show that environmental services are not the most important determinants of residential location for most of the population. This could reflect a lack of awareness of the importance of environmental services for the well being of the population or the need for a greater availability of quality environmental information. In either case, this is a message for the local government to improve both the environmental education of the population and the quality of the information available.

Future research is required to include additional environmental amenities, such as air quality and public goods, like transportation accessibility. Additionally, future studies must include, upon availability, housing prices to control for household budget constraints.

7.   References

- Área Metropolitana del Valle de Aburra. Atlas Área Metropolitana del Valles de Aburra. Medellín: Área Metropolitana Valle de Aburra, 2010. ISBN: 978-958-8513-40-9.

- Brown, L. A., & Moore, E. G. (1970). The Intra-Urban Migration Process: A Perspective.

- Cameron, A.C & Trivedi, P. (2005). Microeconometrics: Methods and applications. Cambridge University Press, 491-528.

- Cameron, A. C., & Trivedi, P. K. (2009). Microeconometrics using stata. Texas: A stata press publication.

- Duque, J.(2011) Desarrollo Económico y Competitividad. In: BIO 2030 Plan Director Medellín, Valle de Aburrá. Urbam Centro de Estudios Urbanos y Ambientales de la Universidad EAFIT.

- Earnhart, D. (2001). Combining Revealed and Stated Preference Methods to Value Environmental Amenities at Residential Locations. The University of Wisconsin Press, 113-131.

- Fergusson, M et al. (2007). Voting with Their Feet: Jobs versus Amenities. Growth and Change. Vol. 38 No. 1 (March), pp. 77–110

- Flórez Días, J. (2006). The process of residential desicion making: a conceptual model and the atributes assessed.

- Freese, J., & Long, J. S. (2000). Test for the multinomial logit model. University of Winsconsin - Madison.

- Hoyt, W. H., & Rosenthal, S. (1997). Household Location and Tiebout: Do Families Sort According to Preferences for Locational Amenities? Journal of Urban Economics.

- Ingram, K. G. (1977). Introduction to "Residential Location and Urban Housing Markets". In: Residential Location and Urban Housing Markets. National Bureau of Economic Research. URL: http://www.nber.org/chapters/c4305

- Mc. Fadden, D. Conditional logit analysis of qualitative choice behavior. University of California, California.

- Mc. Fadden, D. Modelling the choice of residential location. University of California, Department of Economics, Berkeley.

- Magnac, T. (2005). Logit models of individual choices. Université de Toulouse.

- Moreno Serrano, R., & Vayá Valcarce, E. (2000). Técnicas econométricas para e tratamiento de datos espaciales: La econometría espacial. Universidad de Barcelona, Barcelona.

- Nechyba, T. J., & Strauss, R. P. (1997). Community choice and local public services: A discrete choice approach. National Bureau Of Economic Research, Massachusetts.

- Ortúzar, J. d., & Rodríguez, G. (2002). Valuing reductions in environmental pollution in a residential location context. Pontificia Universidad Católica de Chile, Department of Transport Engineering, Santiago de Chile.

- Ozturk, E., & Irwin, E. G. (2001, Mayo). Explaining Household Location Choises Using a Spatial Probit Model.