Spatial heterogeneity of housing prices in formal and informal settlements: A GWR hedonic model for segmented markets in Cali

Harvy Vivas Pacheco

Ángela María Franco Calderón

Documento de Trabajo
Alianza EFI - Colombia Científica
Junio 2021

Número de serie: WP4-2021-002

Spatial heterogeneity of housing prices in formal and informal settlements: A GWR hedonic model for segmented markets in Cali

Harvy Vivas Pacheco [(a)] **and Ángela María Franco Calderón** [(b)]

Abstract

This paper studies the factors associated with differential prices in segmented housing markets. Despite the breadth of the literature addressing this issue, a gap remains in the investigation of differential prices in areas where formal and informal settlements coexist. Based on primary data from more than 550 households, we calibrated a Geographically Weighted Regressions (GWR) hedonic model in the District 18 of Cali, Colombia, as an approach for housing pricing in areas where heterogeneous occupation patterns prevail. The findings reveal that environmental quality, violence, accessibility to transportation, business and services, and household incomes are the variables that have the greatest impact on price structures in both the formal and informal sub-markets. In this regard, the model of price behavior developed in this research improves predictions and provides accurate information to policymakers about the factors associated with housing prices.

Keywords

Spatial heterogeneity, informal settlements, housing prices, Geographically Weighted Regressions (GWR), hedonic prices.

JEL classification: G28, R21, R28, R31, R38, R58

Acknowledgments

This paper was supported by Lincoln Institute of Land Policy for the application of the survey on informality in Comuna 18 (Research Papers on "Land Policies and Urban Development in Latin America" program) and the "Inclusión productiva y social: programas y políticas para la promoción de una economía formal, código 60185, Alianza EFI (Inclusive Formal Economy), Contingent Recovery Contract No. FP44842-220-2018.". This is a working version of a book chapter on informality in urban settlements in Cali that we are preparing from the urban project, Univalle, Cali. We also express our gratitude to Vicerrectoría de Investigaciones (Universidad del Valle) for their great support to this specific project of the Alliance.We received valuable feedback from Ciro Biderman and Diane Davis, who were designated by the Lincoln Institute as external arbitrators of the preliminary version of this research.

(a) Universidad del Valle, Cali, Colombia. Ciudad Universitaria Meléndez, Department of Economics. harvy.vivas@correounivalle.edu.co (author for correspondence).

(b) Universidad del Valle, Cali, Colombia. Ciudad Universitaria Meléndez, School of Architecture angela.franco@correounivalle.edu.co

I. Introduction

In Latin America, despite institutional efforts and substantial investments, recent figures reveal that the number of urban poor living in informal settlements increased from 104.8m to 113.4m between 1990 and 2012 (UN-Habitat 2012a), indicating that 24 percent of the total population (UN-Habitat 2012b) lives in these types of dwellings. Whereas demographic trends show an ongoing demand for urban homes (Bouillon 2012), the provision of affordable housing continues to be a major challenge in the region. The factors that contribute to the persistence of this situation include insufficient household income, costs and restricted access to mortgages (Bouillon 2012; Dasgupta and Lall 2009; Marx, Stoker and Suri 2013; Woetzel et al. 2014), the weakness of social housing programs (Libertun de Duren 2017), the reinforcement of residential segregation patterns, including those derived from urban planning actions (Fernandes 2008), and ineffective public policies for the integration of poor families into urban socioeconomic dynamics (Smolka 2003; Abramo 2003, 2013).

One of the most important macro-structural factors that influences the perpetuation of informal housing in Latin America is the irregular functioning of land markets and its effect on the spatial heterogeneity of housing prices (Smolka and Mullany 2013; Smolka and Biderman 2009). Despite the relevance of this issue, the difficulties inherent in the study of informal settlements mean that the availability of data on land and housing pricing remains limited. This information is fundamental for market analysis regarding public policies.

Brueckner and Lall (2015) note that despite increasing interest in this issue, there is a lack of empirical work. For instance, studies by Jiménez (1984), Friedman et al. (1988), Lanjouw and Levy (2002), and Kapoor and Le Blanc (2008) have focused on the impact of insecure tenure on housing prices and rents. Manganelli et al. (2014) analyzed segmented markets with hedonic geographic weighting models. More recently, Nakamura (2017) adopted spatial hedonic models to analyze the tenure security premium in informal housing markets in Pune (India), and Talukdar (2018) estimated the presence of housing rent premiums in both formal and informal markets in Nairobi.

Despite the substantial contributions of these studies, a gap remains in the analysis of case studies that juxtapose formal and informal settlements. Hence, this work applies a Geographically Weighted Regression (GWR) model in District 18 in Cali and proposes an approach to housing pricing in areas where heterogeneous occupation patterns prevail based on the identification of the factors that have the greatest impact on differential prices at the local level.

Within this framework, the four premises underpinning this research are as follows. 1) Formal and informal settlements constitute land and housing sub-markets with different entry costs and tenure regimes (Coccato 1996; Payne 2004). 2) Sub-markets work interdependently according to the quality of residential spaces and property rights, overlapping demand and supply curves (Tinsley 1997). 3) These sub-markets form dual structures as the direct result of the inability to absorb the formal sector (Heikkila and Lin 2014). 4) Formal and informal markets are endogenously determined in accordance with municipal regulations (Buckley and Kalarickal 2006); Goytia and Lanfranchi 2009; Brueckner and Selod 2009).

For organizational purposes, this paper is divided into four sections. The first synthesizes the literature related to informal markets, focusing on housing pricing approaches developed for urban sub-markets. The second section introduces the case of Cali and presents the background of residential segregation and housing prices in the city. The third section describes the statistical sources and methods used to model price differentials in the study area. Finally, the fourth section discusses and analyzes the results of the proposed geographical weight model in the empirical strategy, highlighting the importance of the findings for formulating socially inclusive policies.

II. Literature Review

It is difficult to provide comprehensive coverage of the problems associated with the proliferation of informal settlements in Latin America. This is particularly true when attempting to understand the effects of economic and social policies on urban land markets within realities as diverse as those present in each country in the region. However, the literature and global discussions present recurrent topics related to public policies concerning land markets and urban informality. These studies range from broad issues such as regulatory frameworks and their differences between countries (Monkkonen and Ronconi 2013; Henderson 2009) and the impact of housing programs on scenarios of high residential segregation and social exclusion (Ruprah 2011; Greenstein, Sabatini and Smolka and

Larangeira 2007) to specific analyses of the functioning of real estate markets (Smolka and Mullahy 2013), streamlining credit and access to financing (Bouillon 2012), and the existence of dual land and housing markets in metropolitan areas (Brueckner and Selod 2009; Kapoor and Le Blanc 2004; Abramo 2012) as well as endogenous and mutually determined formal and informal areas (Heikkila and Lin 2014).

Regarding the macro-structural factors of urban informality, Smolka and Mullahy (2007) suggest that in Latin America, some factors associated with residential segregation and persistent informality operate through the structure of differential prices. Therefore, the limited purchasing power of vulnerable groups clashes with high prices for developed land and real estate speculation, leading to informal solutions in the urban margins (Abramo 2003, Bouillon 2012). According to Duhau (2013) and Veríssimo (2012), this situation establishes a dual market that combines restrictions on the supply of and access to affordable housing with precarious dynamics in the provision of infrastructure, public goods and urban services. According to this line of reasoning, informal settlements introduce imperfections in market performance that need to be corrected in different ways, as widely discussed in the literature on legalization and regularization (Galiani and Schargrodsky 2010; Field 2005, Smolka and Larangeira 2007, Fernandes 2011).

Opposing the duality theory, Greenstein, Sabatini and Smolka (2007) argue that informality is the result of normal formal land and housing market performance. These authors suggest that restricting vulnerable families' choices limits their access to credit, excludes them from the labor market, and reinforces patterns of socio-spatial segregation. However, authors such as Clichevsky (2000) and Gilbert and Ward (1985) assert that informal areas are expensive for the city as a whole because of their strong environmental and social impact, affecting long-term sustainability.

From a different perspective, Payne (2012) argues that informal settlements configure spontaneous and unregulated sub-markets that are juxtaposed to formal markets. In some cases, these sub-markets are located very close or even overlapped in small territories, while in other cases, they do not necessarily make up adjacent entities (Coccato 1996) because they are spread throughout the city. Thus, there is no single housing market; instead, sub-markets are interrelated with different hedonic price functions (Xiao and Webster 2017, Goodman and

Thibodeau 1998, 2007; Straszheim 1974) and, in many cases, located in specific geographical areas (Basu and Thibodeau 1998).

Accordingly, the differences in the types of land occupation, the legality or illegality of the settlements and the income of people, among other factors, generate segmented housing markets (Da Piedade Morais and de Oliveira 2009; Lim 1987) with differentiated prices within the geographical space. The structure of these prices reveals aspects such as the quality of both the dwellings and their social environment, differences in the sociodemographic composition of the population (Munro 1986; Allen et al. 1995), accessibility to the area, availability of urban services and possibilities of access to mortgage credit. These variables, among other factors, affect the disposition to pay for the net residential space of each family and, as such, become factors that influence the emergence of sub-markets (Xiao and Webster 2017; Ball and Kirwan 1977).

The existence of heterogeneous housing markets therefore implies that the assumption of stationary prices implicit in the attributes of each dwelling is erroneous because it omits geographical differences, the diversity of conditions intrinsic to the supply and demand of land, and the sociodemographic characteristics of the residents. According to Yu et al. (2007), Fotheringham et al. (2002), and Goodman and Thibodeau (1998), the existence of local sub-markets reveals the segmentation and functional imbalances of the housing markets in a given space. Thus, the same attributes generate different prices in each location, and the economic agents have different valuations and dispositions to pay. In this regard, the urban housing market can be segmented according to the structural characteristics of the buildings, the amenities of the neighborhoods, the income levels of the occupants, the types of jobs, social class and ethnic or racial identity (Goodman and Thibodeau 1998). This market may also be composed of formal and informal housing with different ownership regimes and provision of services. Thus, non-stationary housing markets should be studied with different empirical strategies.

In academic research, there is a tradition of empirical estimates of potential segments that evaluate the significance of price differences between sub-markets. In some cases, these segments are imposed exogenously, and in other cases, they come from classification algorithms. The techniques are diverse: hedonic regressions for each sub-market and differences tests and linear hierarchical models (Goodman and Thibodeau 1998), factorial analysis of principal components (Can 1990; Bourassa et al. 1999), spatial syntax models (Xiao and Webster 2017), and geographically weighted hedonic regressions (GWR) (Fotheringham and Park 2018; Li et al. 2017; Yao and Fotheringham 2016; Yu et al. 2007; Fotheringham et al. 2002).

Some authors, such as Wen et al. (2017), compare the results of applying spatial expansion methods against both the traditional hedonic model and the GWR and conclude that the latter constitutes a powerful technique for capturing spatial variation patterns of prices. As Borst (2007) asserts, this method is an effective means to detect market partitions and the variability of the coefficients of each of the hedonic price function regressors without predefining the areas or sub-markets.

In this framework and considering that in many Latin American cities such as Cali formal and informal housing coexist in the same districts or even in the same neighborhoods, this research elaborates on the structure of the housing markets in areas where good-quality residential neighborhoods with high prices converge with depressed areas with low prices. As an attempt to build on previous studies while addressing the non-stationary effects of certain variables on the hedonic price structure, this study models the behavior of housing values in areas with formal and informal sub-markets. This allows us to refine predictions and information on the spatial distribution of prices.

III. Residential segregation, land and housing prices in Cali

Cali is one of Colombia's three main economic and financial centers, along with Medellin and Bogota. In demographic terms, one of the most important characteristics of Cali is that in absolute numbers, it houses the country's highest concentration of Afro-descendants (Urrea et al. 2007). This ethnic-racial condition has been a determining factor in the persistence of residential segregation patterns (Barbary et al. 1999; Barbary 2004; Urrea et al. 2007; Vivas 2013). These patterns are evident in the formation of residential clusters of ethnic minorities and poor immigrants located on the margins of the city, as occurs in District 18. In these sectors, social groups are concentrated in conditions of vulnerability with restricted mobility, poor access to urban services, and little possibility of entering the formal labor markets.

Exploratory analyses of the city as a whole, using a human capital density metric of the working population[1] in each of the census tracts, confirm that Cali has clearly differentiated clusters. The indicator obtained is a proxy for the quality of potential social interactions or the environment in which individuals live: values far from zero denote a higher density of a well-educated working population in relation to the population with a low education. As values approach 100, the ratio between the skilled and unskilled population indicates better potential conditions for micro-local interaction. The structure of this metric by census tract takes the following form[2]:

$$[H/L]^{*}_{j} = \sum_{j=1}^{N} \left\{ \frac{[H/L]_j - [H/L]_{\min}}{[H/L]_{\max} - [H/L]_{\min}} \right\} \quad (1)$$

The results of the measurement for the entire city show that the population with the highest levels of human capital is concentrated along the city's central longitudinal axis. Conversely, the vulnerable population with low levels of human capital lives at the eastern and western borders (see Figure 1).

[1] Obtained from the National Administrative Department of Statistics (DANE, 2005), statistical database Redatam, through the (H/L) ratio between individuals who have completed a high school education or above including vocational and technical careers (H) and those with incomplete secondary, primary or no education (L). This metric is re-scaled with an adjustment factor that considers maximum and minimum values observed inside the city, taking values in the [0-100] range that indicate the relative weight of the qualified population and its spatial distribution.

[2] This indicator fulfills some of the desirable properties of the segregation indices and could be considered a good approximation for the study of spatial cluster formation inside the city. It was suggested in previous research on residential segregation in Cali (Vivas, 2007, 2012 and 2013).

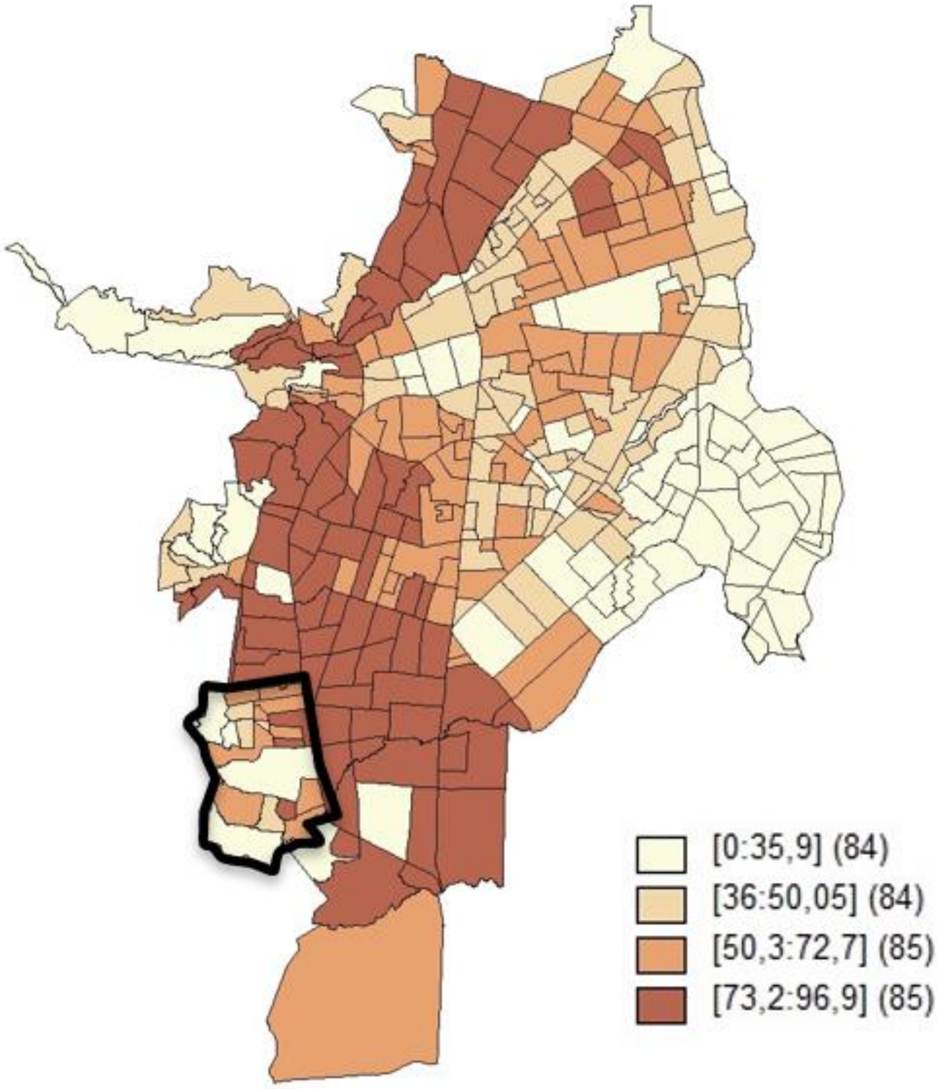


Figure 1. Distribution of human capital in districts by neighborhood metric

Source: Authors' estimates in GeoDa. The area within the polygon corresponds to District 18 along with each of its neighborhoods[3]. The values in brackets correspond to the number of neighborhoods in each category of the quantiles.

Based on these results, the Moran global correlation coefficient was calculated, excluding *missing* values at 0.673. The result of Moran's univariate comparison allows us to reject the null hypothesis of random population distribution according to the human capital density

[3] As shown in Figure 1, in District 18, the majority of neighborhoods have low values for the human capital index with the exception of the most consolidated formal neighborhoods located in the flat area, for which the metric is above 70. This indicates that a high proportion of the working population has an educational level above high school.

metric. This result can be visualized through the Moran scatter plot that provides graphical evidence of the positive spatial dependence of re-scaled [H/L]j (see Figure 2). The horizontal axis corresponds to the metric used, and the vertical axis shows the values corresponding to the first-order spatial lags.

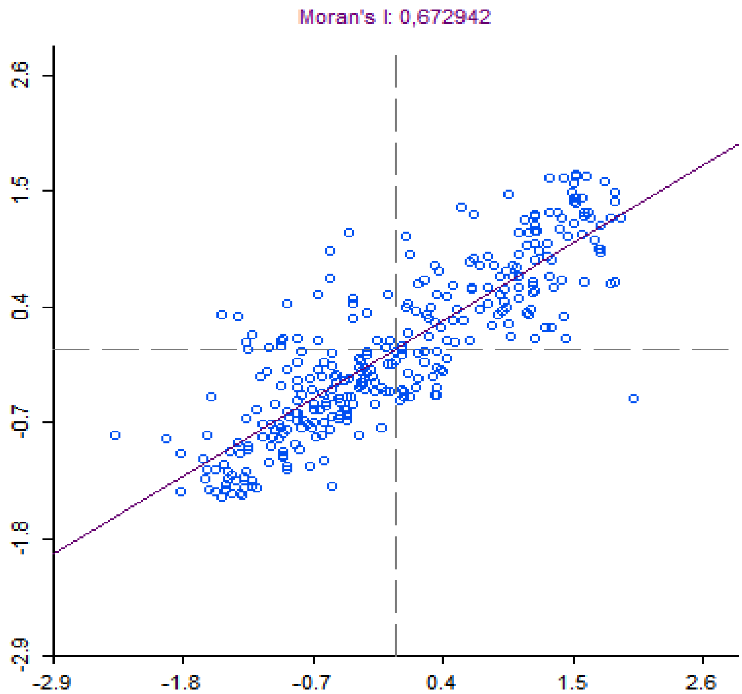
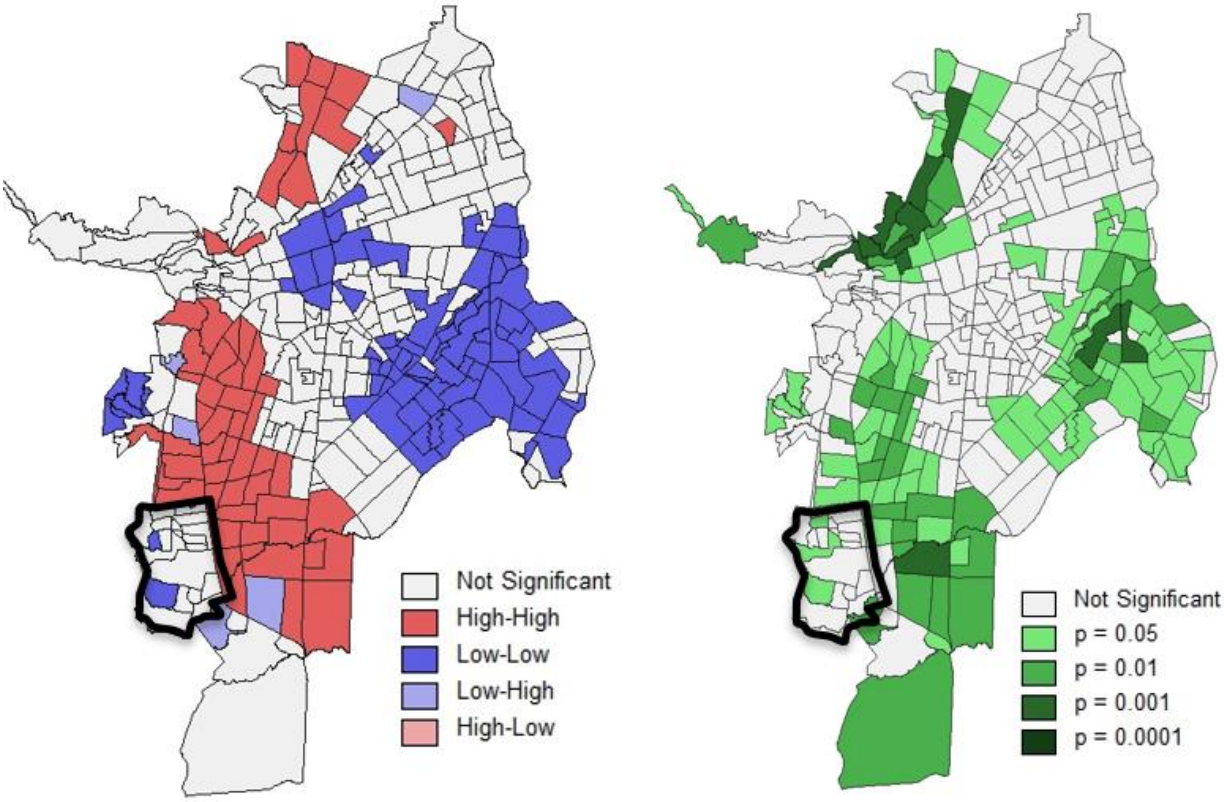


Figure 2 Univariate Moran Scatter Plot
Source: Authors' estimates in GeoDa.

These results are corroborated using information about land prices from real estate market valuations in the urban sector (2010, 2012). Similar to previous findings based on human capital, Moran's univariate index calculated for the land prices reaches a relatively high value of 0.598. To verify the presence of areas with significantly high associations, we adopt Local Indicator of Spatial Association (LISA) indicators. According to the results (see Figures 3 and 4), the central longitudinal axis of the city forms a high-high [H-H] cluster in which spatial units of high prices per square meter predominate surrounded by neighborhoods in the same condition. In the eastern part of the city where the Afro-descendant population is predominant and in the hillside neighborhoods in the west where a blend of Afro-descendant, poor Mestizo and Indigenous populations prevail, the presence of neighborhoods with low prices that have a low-low [L-L]) first-order spatial association is detected. Hence, the analysis reveals that the peripheral neighborhoods in District 18 (at the top of the hillside) have the lowest spatially correlated prices.



Figures 3 and 4: LISA Comparisons for Cali Land Prices
Source: Authors' estimates in GeoDa. Left: cluster formation. Right: significant p-values.

Finally, Moran's bivariate index between the composition of human capital and land prices, estimated for all the polygons that compose the city, reached a value of 0.553, providing additional evidence of the presence of highly segregated residential clusters on the local scale. Based on these findings on the urban scale, the following sections present the methods and results obtained for the calculation of differential housing prices in the case of District 18.

IV. Data and Methods

To study District 18, we used statistical information from three sources: the national census on population and housing; microdata from the Beneficiary Selection System III (SISBEN), which identifies potential households for social assistance programs and targeted expenses (DNP, 2008); and a survey on the informal land and housing market conducted in District 18. Socio-demographic variables were taken primarily from the population census and SISBEN III, which include data on housing units such as property type, materials, access to public services, and the frequency of water supply. This information also includes microdata such as household size, economic distribution of the population, health coverage, birth rate, occupation, and household appliances. The information on land and real estate prices was obtained through the survey that used housing as a unit of analysis, enabling the study of price differentials in the District.

Thus, using the census and SISBEN III distribution as a reference, a representative sample of conglomerates with proportional allocation was defined, considering that the study area is formed of heterogeneous populations distributed in the flat area and the hillside. Of the 659 surveys projected, 571 were conducted, equivalent to 86.6%[4]. After the database was filtered, 558 valid surveys were obtained, which was 84.7% of the number initially projected.

The survey included information on household identification, socioeconomic environment, household unit variables, land and housing occupancy, current occupation, means of transport, and other questions directed toward the heads of the household. In addition to questions on identification and socio-demographic conditions, the survey included a query about house prices, rental value, and willingness to pay when the property is not owned. Owners were asked for the price at which they would sell the property if they received an offer. Finally, tenants were asked about their willingness to pay for the property in the event that they planned to buy.

[4] Difficulties accessing some of the informal settlement addresses from the initial sample as well as logistic and security problems prevented the completion of the initially projected number of surveys. Nevertheless, both the number of surveys conducted and their distribution guarantee the instrument's statistical representativeness.

After processing the data, it was possible to conduct a differential price analysis through micro-econometric exercises that enabled the identification of the variables with the greatest impact on transverse price behavior on the different blocks that make up the neighborhoods included in the sample. For this reason, the methodological strategy explicitly considers the spatial dimension in price variation through GWR methods and uses the houses' different location points in the census blocks as units of analysis. This approach considers information on housing prices, the houses' predominant characteristics, and the socioeconomic and socio-demographic environment, among other attributes that are detailed in the results.

The micro-fundamentals of this type of study are often supported by the hedonic pricing approach, which considers a subject's willingness to pay for a given good based on its characteristics or attributes and according to a set of restrictions, particularly the household level of income or expense structure. As is common in these analyses, the preference structure is represented through a utility function with a good, H (housing), with k characteristics, a compound good, C, which includes all other goods and services, and a budget constraint, Y:

$$U = f(H_k, C) | Y \quad (2)$$

The microeconomic solution of the analytical model enables us to infer that the good's price and attributes provide information about the willingness of agents to pay. In this regard, the analyses of Rosen (1974), Lancaster (1966), and Can (1992) on the real estate market and the works of Yu et al. (2007) and LeSage and Pace (2009), Anselin, L., and Lozano-Gracia, N. (2009) on housing markets with empirical applications constitute conceptual and methodological references.

In our case, while the data used in the econometric model estimations come from the sample in District 18, we do not adopt a pure hedonic pricing approach. This is because the variables stemming from the survey and from secondary sources approximate real estate market values but not transactions actually carried out. Indeed, due to the conditions under which transactions occur in informal markets, there is no statistical database for actual home transactions in a given period as required by the pure hedonic pricing approach.

The primary information from both the formal (flat area) and the informal (hillside area) sectors provides data on the value of actual rent for leased units and rent estimated by owners who inhabit the property. These data also include information on the value estimated by tenants in the event they are able to buy the property along with the value assigned by owners if they are willing to sell[5].

The differential pricing model considers a vector of typical attributes of the houses $\vec{v}$ = $\{v_1, v_2, v_3,\ldots, v_v\}$, including the formal and informal nature of the area, location, physical attributes and accessibility, among other relevant variables from the sample. A vector of the closest variables from the socioeconomic and socio-demographic environment was also constructed, $\vec{s}$ = $\{s_1, s_2, s_3,\ldots, s_v\}$, with the block as the unit of analysis. These data come from the pairing process between the units contained in the sample and the geo-referenced databases along with data from Cali Municipal Planning, SISBEN III, and the census. It should be noted that the surveys contain corresponding keys or link codes that provide accurate, geo-referenced locations of houses, even in the informal sectors.

The $\vec{z}$ vector contains all other relevant variables, such as those that are actually institutional, and some dichotomous variables that attempt to capture whether the household surveyed was a beneficiary of entitlement programs or other housing improvement initiatives promoted by the local government. In the same vector, mobility and transportation variables are considered along with indicators associated with infrastructure projects that affect the study area. Thus, the price vector can be expressed as $\vec{p} = f(\vec{v}, \vec{s}, \vec{z}; \vec{u})$, where u corresponds to the stochastic error term. The GWR local regression model considers the spatial variability of the parameters associated with price predictors. The characteristic feature of these models lies in the assumption of structural permanence of the regression model coefficients.

The literature on hedonic prices and their variants has broadly argued that the existence of spatial self-correlation in prices, along with the presence of substantive heterogeneity, violates the assumptions of conventional models estimated by ordinary least squares (OLS). In addition, the non-random distribution of variables in the space, the presence of clusters, and

[5] Pearson's correlation between the purchase and sale provisions of real estate with revealed rental values reaches 0.55 with the complete sample.

local variability introduce biases and inconsistencies in the ordinary estimates that need to be corrected by appropriate methods. Anselin (1988) and Can (1992), among others, suggest strategies that attempt to improve the quality of estimates and move toward new diagnostic elements that ensure greater accuracy. In reality, the presence of non-stationary processes in space is reflected in the differential impacts of the same variables when considering the specific location or presence of segmented land and housing markets. Thus, the attribute valuation structure has significantly different effects according to spatial references such as centroids or geographic coordinates.

Therefore, considering the sample distribution, the primary information from the survey, and social, economic, and demographic variables pulled from the SISBEN, we estimated the GWR models based on the line of research suggested by Fotheringham et al. (2002). These models are suitable under the condition of random non-spherical or heteroscedastic disturbances. The advantage of this methodological strategy is that it allows us to obtain vectors of parameters that vary in the geo-referenced space. A generic specification of these models is shown below:

$$\vec{p} = \alpha(x_i,y_i) + \beta(x_i,y_i) * \vec{v} + \theta(x_i,y_i) * \vec{s} + \phi(x_i,y_i) * \vec{z} + \vec{u} \qquad (3)$$

Furthermore, $\forall$i= 1, 2, 3,... units in an (x_i,y_i) geographic coordinate system; therefore, the parameter vectors are estimated according to

$$\hat{\beta}_{(xi,yi)} = [Z'W_{(xi,yi)}Exo]^{-1} Z'W_{(xi,yi)}End \text{ ,} \qquad (4)$$

where Exo is the vector of exogenous variables, End is endogenous and $W(x_i,y_i)$ is a diagonal matrix of weights defined as a function of distance that assigns greater weight to closer observations. One possible behavior of the $\vec{u}$ vector considers that its distribution takes the following form so that the variance depends on location:

$$\vec{u} \approx N(0, \sigma^2_{(x,y)}) \qquad (5)$$

As Fotheringham et al. (2002) note, heteroscedasticity in these models may be addressed by introducing a weighting system that uses the inverse variance of each observation.

V. Results and Discussion

Based on the estimates and what they represent, we can detect the magnitudes of the predicted price differentials according to the specific values of the exogenous variable vectors. Once the best model was validated, in accordance with established validation protocols, the estimates made it possible to conduct simulation and impact exercises on prices based on the variables of interest in the micro-local environment, which are presented below.

A) Spatial Exploratory Analysis

In District 18, the exploration of the LISA-univariate comparison for house prices shows that there is a differentiated structure with neighborhood effects. In this structure, two areas can be distinguished: some with high-price levels and high-high [H-H] spatial correlation and others with low prices correlated with neighbors with low-price levels in contiguous areas (L-L) (see Figure 5).

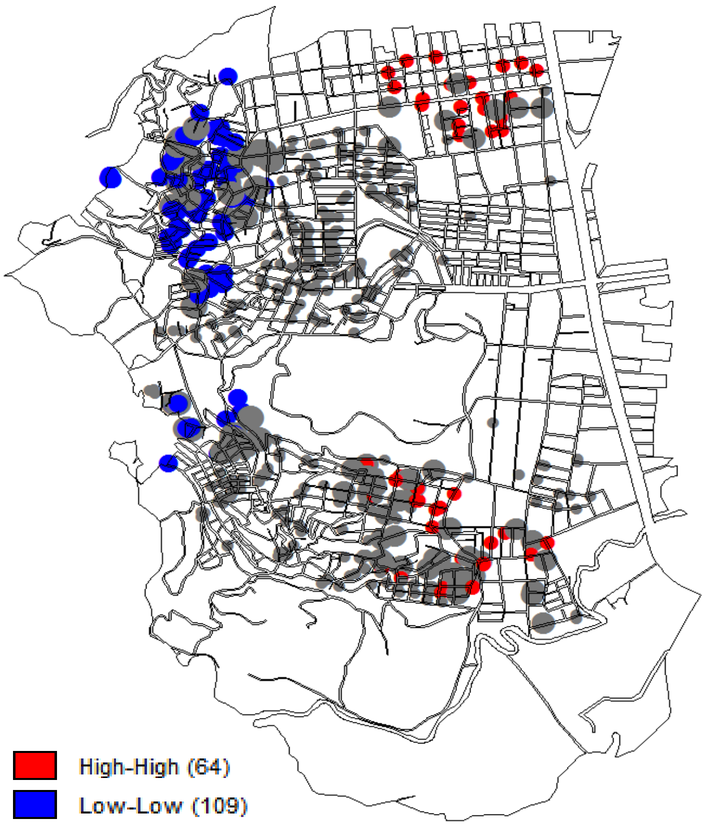


Figure 5. LISA-Univariate Housing Prices in District 18
Source: Authors' estimates from the District 18 survey.
* The values in parentheses correspond to the number of polygons that make up the study area.

These results suggest the presence of spatial heterogeneity in housing prices in the study area, which is likely connected to a combination of factors such as a lack of safety, violence, drug addiction, accessibility differentials, property scale, and the spatial distribution of household income. For these reasons, it is possible that the overall effect estimated in the aggregate models, although relevant, does not reflect the local behavior of each of these factors and their differential impact on house prices, making it necessary to introduce a strategy that includes the variability of the spatial effects of the regressors considered in the pricing models.

B) Geographic Weighting Model (GWR)

The GWR model that is specified below in (x_i, y_i) coordinates uses the logarithm of house prices as a dependent variable and includes the following regressors: *A*, the built surface; *viol*, the variable for the presence of crime and drug addiction in the neighborhood; *time*, the metric for access to business and service locations; g, the per capita expenses of the spatial micro-unit; and u, the stochastic error term.[6]

$$Ln(p_i) = \alpha_{(x_i,y_i)} + \beta_{1(x_i,y_i)} viol + \beta_{2(x_i,y_i)} time + \theta_{(x_i,y_i)} \ln(g) + \delta_{(x_i,y_i)} \ln(A) + u \,. \qquad (6)$$

The results obtained for each of the spatial units in the relevant variables are expressed in weighted averages with the appropriate statistics of strength and correlation of the regressors with the dependent variable. In addition, the estimated models show the local coefficient of determination (R^2), which measures the goodness-of-fit between the systematic part of the model and the dependent variable.

The coefficients obtained show strong correlation and significance in the majority of blocks, and the local R^2 reaches relatively high values throughout the district. The overall goodness-of-fit in the entire study area was 78%, with the best coefficients in the neighborhood blocks in the flat area. The median price elasticity of income was 0.044, the price elasticity of property size was 0.057, and the medians of the local perception of violence and accessibility variables were 0.10 and 0.08, respectively (see Table 1, which also shows estimated values according to quartiles).

---

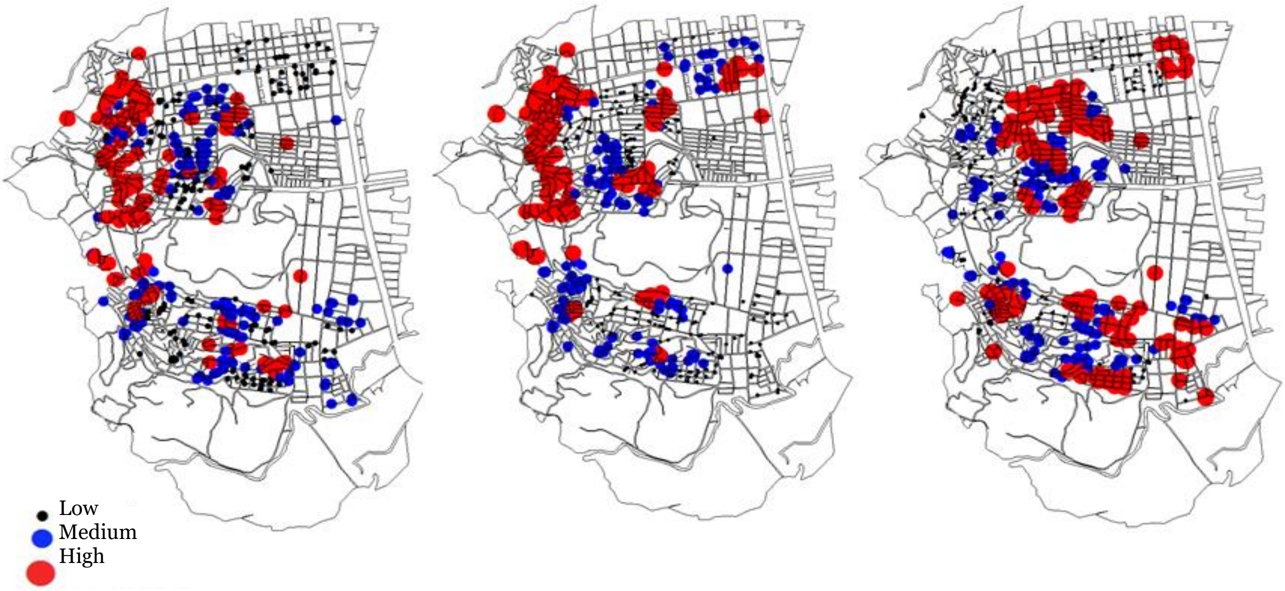
[6] Local estimates of the parameters for the GWR linear model were calculated using the maximum likelihood method in SpaceStat (version 4.0; BioMedware, 2013) and GWR4 (version 4.09; Nakaya et al., 2013) using a Gaussian kernel. Spatial references use the centroid of each block where the sample was located.

Variable	Q1	Median	Q3	Diff-Criterion*
Intercept	3.889444	3.912615	3.934763	1.397571
Violence (viol)	-0.130719	-0.108189	-0.088146	-0.122104
Time (time)	-0.028209	-0.008485	0.013602	0.263360
Ln(A)	0.030029	0.057538	0.102008	-24.675806
Ln(g)	0.032216	0.043924	0.064465	-6.388776

Table 1. Semiparametric geographically weighted regression
Source: Authors' estimates in GWR4 software (Nakaya, et al. (2013)). *Geographical variability tests of local coefficients.

The highest coefficients of the *viol* variable (which measure the local impact of this variable on prices) were detected predominantly in the informal sectors located in the hillside area, whereas the lowest values were identified in the neighborhoods in the flat area. Figure 6 shows the distribution of the low, medium, and high values of the estimated violence coefficient (β1), Figure 7 presents the accessibility metrics and coefficient distribution (β2), and Figure 8 presents the variability of the impact of expenses (income proxy, Θ).

In determining whether the spatial variations of the effects of price regressors were significant, we found from the *diff-criterion* (AICs, BIC, and CV) comparisons that the intercept and the access to urban facilities variables do not show geographical variability in their effects. In contrast, the *viol*, property size, and income variables do affect prices in a heterogeneous manner, a result consistent with the hypotheses initially raised. Estimates show that the quality of the local environment, as measured by an indicator of the presence of drug addiction and violence in the immediate vicinity of the property location, along with access to transportation, businesses, and services, is a relevant factor in explaining price variability in District 18. Global and local estimates effectively verified the importance of these variables along with income using per capita household expenses as a proxy.



Figures 6, 7, and 8: Estimated Coefficient Distributions (β1, β2, AND Θ) in District 18
Source: Authors' estimates from the District 18 survey.

Some interpolation exercises with the GWR model show that in the hillside areas where informal settlements predominate, housing prices between US $172 and US $194 per square meter (see Figure 9) are predicted, whereas data from the survey reveal that monthly household income ranges from US $107 to US $209. This indicates that households would have to allocate nearly all of their income to buy a house from government programs because the average monthly mortgage payment calculated for a group of social housing projects in the study area ranged between US $100 and US $193.

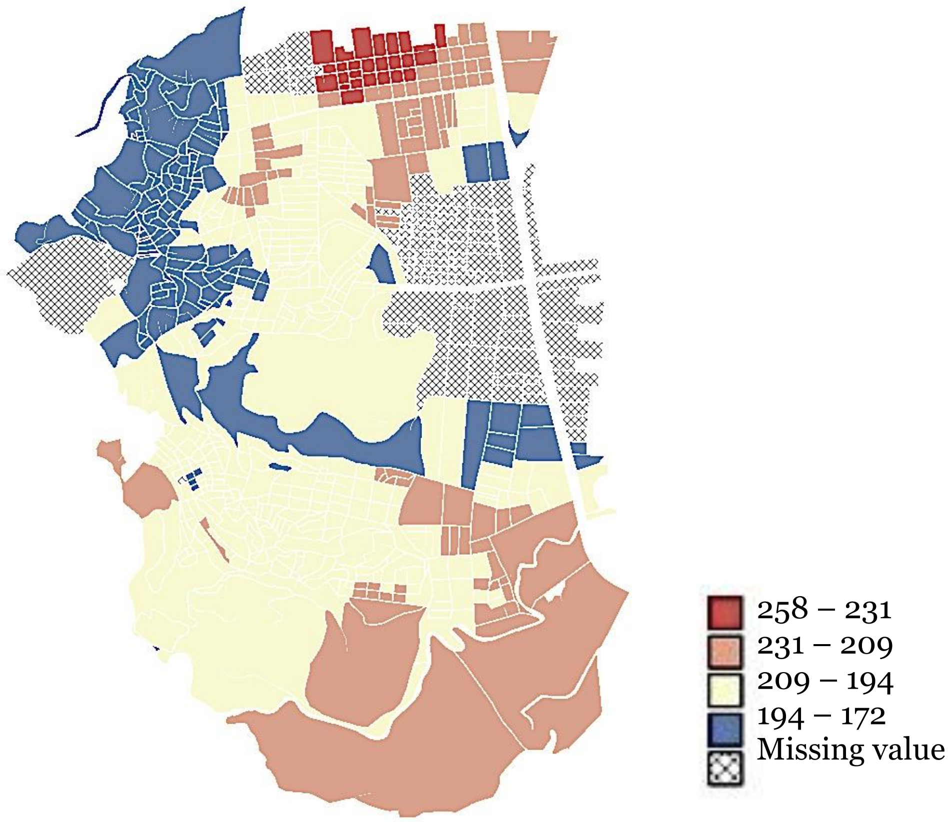


Figure 9: Housing Price Predictions per m² (US$)[7]

Source: Authors' estimates from the adjusted GWR model and using a Gaussian kernel.

---

7 Estimated exchange rate: US$1=COP$2,946 https://themoneyconverter.com/ES/USD/COP.aspx Retrieved 8 May 2017.

This result is consistent with the thesis that insufficient income adds to a combination of restrictions that families face when attempting to acquire social housing, forming the so-called 'affordability gap' (Stone 1993; Kutty, 2005; Bouillon, 2012) that leads to worsening and persistent poverty conditions. In relation to the price differentials predicted by the model, our findings corroborate that there is a significant gap between the flat area where formal housing predominates and the hillside areas where informal housing prevails. The differences range from 34.3% of the value per square meter between the most expensive and cheapest areas to differentials of 11% in the poorest formal areas and the hillside areas that are in better condition than the other neighborhoods in this conglomerate.

Furthermore, our results reveal that in what we call *transition zones,* located on the limits between the flat and hillside areas and where formal and informal settlements are mixed, small differences exist in housing prices despite the property regime (yellow areas in Figure 9). According to our model, the social conditions of violence coupled with the perception of insecurity in these areas constitute the characteristic features that explain the small differences in housing prices between formal and informal housing. Local regression estimates thus confirm this fact. This result therefore demonstrates that in places where the line between formal and informal settlements is thin, people's perception of the surrounding environment has an important weight in property prices. In this type of area, where the identification of sub-markets is difficult due to the coexistence of formal and informal housing, GWR hedonic models are the ideal tool.

VI. Final Remarks

The results obtained with GWR models show that the house pricing structure is significantly heterogeneous and that the incidence of each explanatory variable studied has important variations in the district's geographic space. These variations in the coefficients can be satisfactorily measured by the strategy adopted in this study, yielding data consistent with those obtained in the exploratory phase developed in section 4.1. During this first phase, the comprehensive estimates showed that the variables associated with a house's immediate environment, such as the presence or absence of drug addiction and violence along with access to businesses and services, play a principal role in explaining price variability within the neighborhoods that make up District 18.

In this context, the presence of spatial clusters in the distribution of prices between flat and hillside areas provided new evidence of the imperfections of the housing market when accessibility and the quality of social environment (crime and drug addiction) are considered. Technically, as LeSage and Pace (2009) note in some of their applied exercises of spatial econometrics and as Yu et al. (2007) observe in their work on housing prices in Milwaukee, the "non-stationarity" of housing market prices reveals the presence of segmented markets or sub-markets. These segmented markets require new, alternative methodological strategies for modeling rather than the conventional strategies.

Accordingly, the strategy implemented in District 18 represents an alternative for modeling the spatial heterogeneity of housing prices where formal and informal settlements converge as a tool to shape social housing policies with solid data on housing market behavior.

References

Abramo, Pedro (2003). La teoría económica de la favela: cuatro notas sobre la localización residencial de los pobres y el mercado inmobiliario informal. *Ciudad y territorios: Estudios territoriales*, 35, 136-137.

Abramo, Pedro (2012). La ciudad com-fusa: mercado y producción de la estructura urbana en las grandes metrópolis latinoamericanas. *EURE*, 38 (114), 35-69.

Abramo, Pedro (2013). Mercado informal y producción del hábitat: la nueva puerta de acceso a los asentamientos populares en América Latina. In Bolívar, T. & Erazo, J. (Eds.), *Los lugares del hábitat y la inclusión*, (pp.29-58). Quito: FLACSO.

Alcaldía de Cali (2015). *Cali en Cifras*. Cali: Alcaldía Municipal. https://planeacion.cali.gov.co/caliencifras/Documentos%20pdf/Caliencifras2014.pdf Accessed 10 June 2018.

Allen, Marcus; Springer, Thomas; Waller, Neil (1995). Implicit pricing across residential rental submarkets. *The Journal of Real Estate Finance and Economics*, 11(2), 137–151.

Anselin, Luc (1988). Spatial Econometrics: Methods and Models. Boston: Kluwer Academic.

Anselin, Luc, & Lozano-Gracia, N. (2009). Spatial hedonic models. In *Palgrave Handbook of Econometrics* (pp. 1213-1250). Palgrave Macmillan, London.

Ball, M., & Kirwan, R. (1977). Accessibility and supply constraints in the urban housing market. *Urban Studies*, 14(1), 11–32.

Barbary, Olivier; Ramírez, Héctor Fabio & Urrea, Fernando (1999). Población afrocolombiana y no afrocolombiana en Cali: segregación, diferenciales sociodemográficos y de condiciones de vida. In Cubides, F. & Domínguez, C. (Eds.), *Desplazados, migraciones internas y reestructuraciones territoriales*, (pp. 301-336). Bogotá: Centro de Estudios Sociales (CES) y Universidad Nacional de Colombia.

Barbary, Olivier (2004). El componente socio-racial de la segregación residencial en Cali. In Barbary, O. & Urrea, F. (Eds.), *Gente negra en Colombia. Dinámicas sociopolíticas en Cali y el Pacífico*. (pp. 157-194). Medellín: Cidse, IRD, Colciencias, Editorial Lealon.

Basu, Sabyasachi & Thibodeau, Thomas (1998). Analysis of spatial autocorrelation in housing prices. The Journal of Real Estate Finance and Economics 17, 61–85.

Biomedware (2013). SpaceStat 4 [Computer software]. Washington, D.C.: Biomedware.

Borst, Richard A. (2007). *Discovering and applying location influence patterns in the mass valuation of domestic real property* (Doctoral dissertation, University of Ulster). http://www.ipti.org/wp-content/uploads/2014/05/RABThesisFinal-8.5x11.0.pdf Accessed 1 April 2018.

Bouillon, César (Ed.) (2012). *Room for Development: Housing Markets in Latin America and the Caribbean*. New York: IDB, Palgrave, Macmillan.

Bourassa, Steven C,; Hoesli, Martin; Foort, Hamelink; MacGregor, Bryan D. (1999). Defining housing submarkets. *Journal of Housing Economics*, 8(2), 160–183.

Brueckner, Jan. K., & Lall, Somik V. (2015). Cities in developing countries: fueled by rural–urban migration, lacking in tenure security, and short of affordable housing. In *Handbook of regional and urban economics* (Vol. 5, pp. 1399-1455). Elsevier.

Brueckner, Jan. K. & Selod, Harris (2009). A theory of urban squatting and land-tenure formalization in developing countries. *American Economic Journal: Economic Policy*, 1(1), 28-51.

Buckley Robert M. & Kalarickal Jerry (Eds.) (2006) *Thirty Years of World Bank Shelter*

Lending: what have we learned? World Bank, Washington DC. http://documents.worldbank.org/curated/en/958091468159898471/Thirty-years-of-World-Bank-shelter-lending-What-have-we-learned Accessed 8 May 2018.
Can, Ayse (1992). Specification and estimation of hedonic housing price models. *Regional science and urban economics*, 22 (3), 453-474.
Da Piedade Morais, Maria, & de Oliveira Cruz, Bruno (2009). Housing demand, tenure choice, and housing policy in Brazil. In *Urban Land Markets* (pp. 253-282). Springer.
Clichevsky, Nora (2000). *Informalidad y segregación urbana en América Latina: una aproximación*. Santiago de Chile: CEPAL. https://repositorio.cepal.org/bitstream/handle/11362/5712/1/S00100859_es.pdf Accessed 14 April 2016.
Coccato, Marcelo (1996). *Alternatives to Home Ownership: Rental and Shared Sub-Markets in Informal Settlements*. Montreal, Canada: McGill University. Thesis. https://www.mcgill.ca/mchg/student/alternatives Accessed 2 January 2018.
Dasgupta, Basab & Lall, Somik V. (2009). Assessing benefits of slum upgrading programs in second-best settings. In Lall S.V., Freire M., Yuen B., Rajack R. & Helluin JJ. (Eds.) *Urban Land Markets* (pp. 225-251). Netherlands: Springer.
Departamento Nacional de Estadística (DANE) (2005). *Censo general 2005*. Bogotá: DANE, 2005.
Departamento Nacional de Planeación (DNP) (2008). Dirección de Desarrollo social. Grupo de Calidad de Vida. *Diseño del índice Sisben en su tercera versión*. Resumen ejecutivo. Bogotá: DNP.
Duhau, Emilio (2013). La ciudad informal: ¿precariedad persistente o hábitat progresivo? In Bolívar, T. & Erazo, J. (Eds.) *Dimensiones del hábitat popular latinoamericano*, (pp. 59-85). Quito: FLACSO.
Fernandes, Edesio (2008). Consideraciones generales sobre las políticas públicas de regularización de asentamientos informales en América Latina. *EURE*, 34 (102), 25-38.
Fernandes, Edesio (2011). *Regularization of informal settlements in Latin America*. Cambdridge, MA: Lincoln Institute of Land Policy. https://www.lincolninst.edu/sites/default/files/pubfiles/regularization-informal-settlements-latin-america-full_0.pdf Accessed 3 March 2016.
Field, Erica (2005). Property rights and investment in urban slums. *Journal of the European Economic Association*, 3(2-3), 279-290.
Fotheringham, Stwart & Park, Bumsub (2018). Localized spatiotemporal effects in the determinants of property prices: A case study of Seoul. *Applied Spatial Analysis and Policy*, 11(3), 581-598.
Fotheringham, Alexander; Brunsdon, Chris & Charlton, Martin (2002). *Geographically weighted regression: the analysis of spatially varying relationships*. Chichester: Wiley.
Friedman, Joseph; Jimenez, Emmanuel & Mayo, Sephen K.(1988). The demand for tenure security in developing countries. *Journal of development Economics*, 29(2), 185-198.
Galiani, Sebastian & Schargrodsky, Ernesto (2010). Property rights for the poor: Effects of land titling. *Journal of Public Economics*, 94(9), 700-729.
Gilbert, Alain & Ward, Peter M. (1985). *Housing, the State and the Poor: Policy and Practice in Three Latin American Cities*. Cambridge (UK): Cambridge University Press.
Goodman, Allen & Thibodeau, Thomas (2007). The spatial proximity of metropolitan area housing submarkets. *Real Estate Economics*, 35(2), 209–232.
Goodman, Allen & Thibodeau, Thomas (1998). Housing market segmentation. *Journal of Housing Economics*, 7(2), 121–143.

Goytia, Cynthia & Lanfranchi, Gabriel (2009) Informal neighborhoods in the Buenos Aires metropolitan region: understanding the effects of land regulation on the welfare of the poor. In Lall SV et al (Ed.) *Urban Land Markets: improving land management for successful urbanization*. Springer, New York, pp 163–190.

Greenstein, Rosalind; Sabatini, Francisco & Smolka, Martim (2007). Segregación espacial urbana: fuerzas, consecuencias y respuestas normativas [Electronic version]. In Smolka, M. & Mullahy, L. (Eds.), *Perspectivas urbanas: temas críticos en políticas de suelo en América Latina* (pp. 329-333. Cambridge, MA: Lincoln Institute of Land Policy). https://www.lincolninst.edu/sites/default/files/pubfiles/perspectivas-urbanas-cd-full.pdf Accessed 3 March 2017.

Heikkila, Eric, & Lin, Michael (2014). An integrated model of formal and informal housing sectors. *The Annals of Regional Science*, 52(1), 121-140.

Henderson, J. Vernon (2009). The effect of residential land market regulations on urban welfare. In Lall S.V., Freire M., Yuen B., Rajack R. & Helluin JJ. (Eds.) *Urban Land Markets* (pp. 25-49). Netherlands: Springer.

Jimenez, Emmanuel (1984). Tenure security and urban squatting. *The review of economics and statistics*, 556-567.

Kapoor, Mudit & Le Blanc, David (2008). Measuring the risk on housing investment in the informal sector: theory and evidence from Pune, India. Policy Research Working Paper No. 3433. (Washington, D.C.: World Bank). https://openknowledge.worldbank.org/handle/10986/14235 Accessed 2 February 2015.

Kutty, Nandinee K. (2005). A new measure of housing affordability: Estimates and analytical results. *Housing Policy Debate*, 16 (1), 113-142.

Lancaster, Kelvin J. (1966). A new approach to consumer theory. *The journal of political economy*, 132-157.

Lanjouw, Jean Olson & Levy, Philip I. (2002). Untitled: A study of formal and informal property rights in urban Ecuador. *The Economic Journal*, 112(482), 986-1019.

LeSage, James & Pace, R. Kelley (2009). *Introduction to Spatial Econometrics*. LLC: Taylor & Francis Group.

Li, Shengwen, Ye, Xinyue, Lee, Jay, Gong, Junfang & Qin, Chenglin (2017). Spatiotemporal analysis of housing prices in China: a big data perspective. *Applied Spatial Analysis and Policy*, 10(3), 421-433.

Libertun de Duren, Nora (2017). The Social Housing Burden: Comparing Households at the Periphery and the Centre of Cities in Brazil, Colombia, and Mexico. *International Journal of Housing Policy*, 1-27.

Lim, Gill Chin (1987). Housing policies for the urban poor in developing countries. *Journal of the American Planning Association*, 53 (2), 176–185.

Manganelli, Benedetto; Pontrandolfi, Piergiuseppe; Azzato, Antonello, & Murgante, Beniamino (2014). Using geographically weighted regression for housing market segmentation. *International Journal of Business Intelligence and Data Mining* 13, 9(2), 161-177.

Marx, Benjamin; Stoker, Ttihomas & Suri, Tavneet (2013). The economics of slums in the developing world. *The Journal of Economic Perspectives*, 27 (4), 187-210.

Monkkonen, Paavo & Ronconi, Lucas (2013). Land use regulations, compliance and land markets in Argentina. *Urban Studies*, 50 (10), 1951-1969.

Munro, Moira (1986). *Testing for segmentation in the private housing market in Glasgow*. Centre for Housing Research, Discussion Paper No 8.

Nakamura, Shohei (2017). Tenure security premium in informal housing markets: A spatial hedonic analysis. *World Development*, 89, 184-198.
Nakaya, Tomoki; Charlton, Martin; Brunsdon, Chris; Lewis, Paul; Yao, Jing; Fotheringham, Stewart (2013). GWR4 [Computer software]. Ireland; Glasgow; Arizona: Microsoft Windows.
Payne, Geoffrey (2004). Land tenure and property rights: An introduction. *Habitat International*, 28(2),167–179.
Payne, Geoffrey & Majale, Michael (2012). *The urban housing manual: Making regulatory frameworks work for the poor*. Routledge.
Rosen, Sherwin (1974). Hedonic prices and implicit markets: product differentiation in pure competition. *Journal of Political Economy*, 82 (1), 34-55.
Ruprah, Inder (2011). The Welfare Impacts of Social Housing Programs in Latin America: A Meta-Impact Analysis. Working paper No. 0511 [Electronic version]. (Washington, D.C.: Inter-American Development Bank). http://idbdocs.iadb.org/WSDocs/getdocument.aspx?docnum=36182379 Accessed 2 May 2015.
Smolka Martim (2003). Informality, urban poverty and land market prices. Land Lines 15(1) 4-7. https://www.lincolninst.edu/publications/articles/informality-urban-poverty-land-market-prices Accessed 8 May 2015.
Smolka, M. & Larangeira, A. (2007). *¿Cediendo Terreno? Informalidad y Políticas Urbanas en América Latina*. Cambridge (MA): Lincoln Institute of Land Policy.
Smolka, M. & Mullahy, Laura (2007). *Perspectivas urbanas: temas críticos en políticas de suelo en América Latina*. Cambridge, MA: Lincoln Institute of Land Policy. https://www.lincolninst.edu/sites/default/files/pubfiles/perspectivas-urbanas-cd-full.pdf Accessed 30 June 2016.
Smolka, M. & Biderman, Ciro (2009). Measuring Informality in Housing Settlements: Why Bother? *Land Lines*, 21 (2), 14-19. https://www.lincolninst.edu/sites/default/files/pubfiles/1581_798_Article%203.pdf Accessed 30 June 2016.
Smolka, Martim & Mullahy, Laura (ed.) (2013). *Políticas de suelo urbano: Perspectivas internacionales para América Latina*. Cambdridge (MA): Lincoln Institute of Land Policy. https://www.lincolninst.edu/sites/default/files/pubfiles/politicas-de-suelo-urbano-full.pdf Accessed 8 June 2016.
Straszheim, Mahlon (1974). Hedonic estimation of housing market prices: A further comment. *The Review of Economics and Statistics*, 56(3), 404–406.
Talukdar, Debabrata (2018). Cost of being a slum dweller in Nairobi: Living under dismal conditions but still paying a housing rent premium. *World Development*, 109, 42-56.
Tinsley, Elaine A. (1997). Mechanics of informal land and housing markets: A theoretical exposition (MCP thesis, Massachusetts Institute of Technology). https://dspace.mit.edu/bitstream/handle/1721.1/70293/37776955-MIT.pdf;sequence=2 Accessed 14 July 2017.
UN-Habitat. (2012a). *State of the World's Cities 2012/2013*. (T. Naudin, Ed.). Malta: UN-Habitat. http://mirror.unhabitat.org/pmss/listItemDetails.aspx?publicationID=3387 Accessed 4 February 2016.
UN-Habitat. (2012b). *State of Latin American and Caribbean Cities 2012. Towards a new urban transition*. Naples: UN-Habitat. http://caribbean.cepal.org/content/state-latin-american-and-caribbean-cities-report-2012 Accessed 4 February 2016.
Urrea, Fernando; Viáfara, Carlos; Ramírez, Héctor Fabio & Botero, Waldor (2007). Las desigualdades raciales en Colombia: Un análisis socidemografico de condiciones de vida,

pobreza e ingresos para la ciudad de Cali y el departamento del Valle del Cauca [Electronic version]. In Mosquera, C. & Barcelos, L.C. (Eds), *Afro-reparaciones: memorias de la esclavitud y justicia social reparativa para negros afrocolombianos y raizales* (pp. 691-710), Bogotá: Universidad Nacional de Colombia, Observatorio del Caribe Colombiano.

Veríssimo, Antônio Augusto (2012). Programas de regularización y formación de las plusvalías en las urbanizaciones informales. In Bolívar, T. & Erazo, J. (Eds.), *Dimensiones del hábitat popular latinoamericano*. (pp. 45-68), Quito: FLACSO.

Vivas, Harvy (2007) *Educación, background familiar y calidad de los entornos locales en Colombia*. Tesis doctoral. Barcelona: Universitat Autónoma de Barcelona.

Vivas, Harvy (2012). Segregación residencial y distancia socioeconómica en Cali. In *Memorias del Congreso Nacional de Sociología. Herencia y ruptura en la Sociología Colombiana Contemporánea*. Cali: Universidad del Valle-ICESI.

Vivas, Harvy (2013). Persistencia de la segregación residencial y composición del capital humano por barrios en la ciudad de Cali. *Ensayos sobre política económica*, 31 (70), 121-155.

Wen, Haizhen; Jin, Yilan & Zhang, Ling (2017). Spatial heterogeneity in implicit housing prices: evidence from Hangzhou, China. I*nternational Journal of Strategic Property Management*, 21(1), 15-28.

Woetzel, Jonathan; Ram, Sangeeth; Mischke, Jan; Garemo, Nicklas & Sankhe, Shirish (2014). A blueprint for addressing the global affordable housing challenge. McKinsey Global Institute. Retrieved 10-02-2016, from: http://globalhousingindicators.org/en/content/mckinsey-global-institute-blueprint-addressing-global-affordable-housing-challenge Accessed 10 February 2018.

Xiao, Yang; & Webster, Chris (2017). *Urban morphology and housing market*. Singapore: Springer.

Yu, Dalin; Wei, Yehua Dennis & Wu, Changshan (2007). Modeling spatial dimensions of housing prices in Milwaukee, WI. *Environment and Planning B: Planning and Design*, 34 (6), 1085-1102.

Agradecimientos

Esta serie de documentos de trabajo es financiada por el programa "Inclusión productiva y social: programas y políticas para la promoción de una economía formal", código 60185, que conforma Colombia Científica-Alianza EFI, bajo el Contrato de Recuperación Contingente No.FP44842-220-2018.

Acknowledgments

This working paper series is funded by the Colombia Científica-Alianza EFI Research Program, with code 60185 and contract number FP44842-220-2018, funded by The World Bank through the call Scientific Ecosystems, managed by the Colombian Ministry of Science, Technology and Innovation.