Modeling social heterogeneity, neighborhoods and local influences on urban real estate prices

Nederlandse Geografische Studies / Netherlands Geographical Studies

ISSN 978-90-6809-428-2

Netherlands Geographical Studies 385

Modeling social heterogeneity, neighborhoods and local influences on urban real estate prices

Spatial dynamic analyses in the Belo Horizonte metropolitan area, Brazil

Bernardo Alves Furtado

Utrecht 2009

Koninklijk Nederlands Aardrijkskundig Genootschap
Faculteit Geowetenschappen Universiteit Utrecht

This dissertation was partly accomplished with financial support from Urban and Regional research centre Utrecht (URU), and CAPES (Brazilian government research institute) which funded an internship period at Research Institute of Knowledge Systems (RIKS). It is also part of an international partnership agreed upon by the Centre for Regional Development and Planning (CEDEPLAR) at Federal University of Minas Gerais (UFMG), and URU at Utrecht University. The author also benefited from a work leave in 2008 from Centro Universitário UNA.

ISBN 978-90-6809-428-2

Graphic design: GeoMedia (Faculty of Geosciences, Utrecht University) (7517)

Printed in the Netherlands by A-D Druk b.v. – Zeist

Contents

Figures

Tables

Acknowledgements

This journey has been a long, meandering learning process that is fortunately far from over, but an important stage is concluding. Personal motivation and curiosity have remained high throughout the period, without which little would have been achieved. However, if not for the invaluable support of a number of people and institutions, this process would have been nothing but ephemeral dreams.

Chronologically, I have to thank Roberto Luis Monte-Mór for early support in reading the first research project, and the unbiased selection committee of CEDEPLAR/UFMG who was not afraid to welcome an architect to their Regional Economics center. In the early stages of my debut in Economics in 2005, my dear classmates were indispensable. I would like to thank Guilherme Issamu Hirata, Victor Maia Delgado, Guilherme Jonas Costa Silva, Ana Tereza Lanna Figueiredo, Rodrigo Lara Pinto Coelho, Elydia Mariana Gonçalves da Silva, Bruno Dutra Badia, and all the others. The institutional environment at CEDEPLAR has always been supportive and I also appreciated the help of Professors Rodrigo Simões Ferreira, Andre Golgher, Edson Domingues, Sueli Moro, Ana Maria Hermeto Camilo de Oliveira, Marco Crocco, Roberto Luis de Melo Monte-Mór, Mauro Borges and Frederico Gonzaga Junior, as well as the helpful aid of Maria Cecilia da Silva Neto and the team at the secretariat. I also thank Mirtes at the library.

Prefeitura de Belo Horizonte (the City), through the Department of Real Estate Valuation, and managers José Afonso Bicalho Beltrão, Omar Pinto Rodrigues, and Ervio de Almeida provided the data for chapter 6, to whom I am very grateful. I also would like to strongly acknowledge the help of Professors Zhenlin Yang and Liangjun Su who provided the IVQR Matlab procedure.

The four-month period at RIKS was funded by CAPES – the Brazilian Research Foundation – to which I am very grateful.

The year in the Netherlands would not have been possible without the active, unconditional support of Centro Universitário UNA. Specifically, I would like to express my gratitude to Professor Ricardo Paiva – the Dean of Research – who provided the necessary conditions to accomplish the final leap. I am also grateful to Padre Magela – the Dean of Centro Universitário UNA – and my fellow co-workers.

A long-run thank you should go to earlier supervisors and their help with concepts and groundwork: Beatriz Couto, João Gabriel Teixeira (Japão) and João Francisco de Abreu.

I also benefited from discussions with follow PhD students at URU, especially with regard to the economic geography cluster, and with post-doctoral researcher Marco Capasso, with regard to understanding the IVQR routine. I thank Marianne van Jonge and her team at the secretariat for support with all kinds of questions.

Hedwig van Delden has been a friend and a terrific help, from early contact, to the first meeting, to my acceptance and welcoming at RIKS, and to follow-up work. Actually, the teamwork and availability of members at RIKS touched me deeply. Saim Muhammad and Jasper van Vliet – my roommates – responded pleasantly to my battery of questions at the early stages of understanding the foundations and principles of cellular automata modeling. Alex Hagen-Zanker was key to the concepts of validation and literature on the theme. Jelle Hurkens enabled the cornerstone moment of adapting changes to the model so that I could actually run it. Patrick didactically covered a number of issues in programming. My stay and that of my family was made much more pleasant by the friendship and help of all, but especially Hedwig, Roel Vanhout, and Edith Schmeits.

At URU, I was lucky to have my enthusiasm shared early on by Ronald van Kempen, and then by my supervisors Frank van Oort and Dick Ettema. They, together, made invaluable suggestions to transform a work in progress into a full product. Their reading and discussion was both careful and thorough. They helped translate general sentences and collections of calculations into cohesive readable material. I cannot thank them enough for their support.

My supervisors in Brazil, Clélio Campolina Diniz and Ricardo Machado Ruiz, provided invaluable help. Campolina ensured that the logic of the research structure was maintained, and facilitated important overarching organization for the research, the necessary main points and the completeness of the thesis. Ricardo first introduced me to self-organizing systems and has tolerated quite pleasantly my constant e-mailing, questioning, haggling and requests for advice. Later on, both of them were highly supportive of the trip to ERSA 2005 in Amsterdam, the internship at Research Institute of Knowledge Systems (RIKS), and the co-tutorship with URU at University of Utrecht. In addition to providing constant conceptual and technical advice, they have become close friends and colleagues.

My parents, Dimas and Pitucha, also endured their quota of reading, commenting and encouragement. My enthusiasm has been shared throughout by my sisters, Raquel and Clarissa.

Finally, I have to express my immense gratitude to my wife Claudia and her continuous support, listening, encouragement, and sharing, as she has made all phases of the work a full family plan. My young son Arthur, too, has always been understanding, patient and a source of joyful distraction and mind-clearing moments.

Introduction

This study aims to bring together the insights of two strands of literature regarding urban development: urban economics – originally considered in a monocentric spatial-economic framework – and self-organizing systems and cellular automata (CA) modeling, which aim to explain multinodal and evolutionary urban development. We do so by examining urban real estate markets that reflect the complexity of urban development. Innovatively, socioeconomic heterogeneity and the notion of neighborhoods – viewed as essential in planning and urban studies – are introduced in our models. Herein, neighborhoods as identity communities and their local influence are hypothesized to be crucial elements for explaining urban economic and morphological development in a research framework applied to the metropolis of Belo Horizonte in Brazil. A better understanding of the processes of urbanization is likely to benefit policy-makers' by providing them with a wider array of hermeneutic tools.

Urban development in epistemic research communities

Cities are complex socioeconomic phenomena that spread physically in space, and exhibit a permanent and continuous evolutionary development trend over time (Jacobs, 1970; Serra, 1987; Soja, 2000).

Urban environments can be analyzed from different conceptual viewpoints. After identifying and defining the city as an object of study in the nineteenth century (Benevolo, 1980), many scientific disciplines have been developing theories, concepts and methodologies to explain and describe cities' intricate processes.

Urban sociologists proposed ecological models early in the early twentieth century (Park, Burgess *et al.*, 1925), aiming for a better understanding of social and spatial segregation within the city. In the 1960s, the study of urban economics was theoretically boosted by Alonso's seminal work on location and land-use (1964). His Cartesian approach served as a reference for a subsequent rich theoretical discussion (Capello and Nijkamp, 2004b) that considers distance to the city center as the paramount factor in explaining the footprint of economic interaction. At about the same time, Jane Jacobs (1970) approached the (macro)economics of the city in a different way, claiming that the diversity and proximity of actors – provided exclusively by city context – enables urban development and growth.

City morphology and city structure have received increasing attention since the 1990s, with a focus placed on implicit economic factors (White and Engelen, 1993), segregation issues (Portugali, 2000), morphology (Batty, M. and Longley, 1994) or transport and land-use functions (Waddell, 2002). In a broader scope, social scientists and urban geographers confirm the socially-divided configuration of cities (Caldeira, 2000). Cities' complexity can be further viewed as highly differentiated in terms of social composition, unequal economic power and segregated spatial characteristics (Lemos, Ferreira *et al.*, 2004), which are more pronounced in urban agglomerations in developing countries. These factors inspire us to include social and spatial heterogeneity as a relevant factor of our urban analysis.

Urban real estate markets and spatial complexity

Urban real estate markets influence and are influenced by cities' characteristics and inhabitants' perceived views, because these markets derive a large part of their value from their location within the city. This dual causality makes real estate markets as complex and multifaceted as cities. The effects of attraction and repulsion among actors are dynamic, where household location is determined, in part, by business locations and vice-versa (Mills and Nijkamp, 1987), or as some suggest, in a pattern where "jobs follow households" (Steinnes, 1982).

Moreover, urban real estate represents between 45% and 75% of the wealth stock and expenses flows in both rich and poor metropolitan areas (Ibbotson, Siegel *et al.*, 1985). DiPasquale and Wheaton (1996) show that residences represent the major item of consumption in citizens' budgets. In an urban context, where land is a differentiated good, economies and diseconomies of agglomeration can be viewed as the competition for better-located space, offset by price that increases with competition. Competition for urban space is reflected in real estate prices and in further agglomeration or disagglomeration. Real estate prices are crucial to determine the repelling factors that act against agglomeration (Krugman, 1996). In other words, the dynamics of urban spatial configurations are the result of positive and negative feedback mechanisms (Batty, M., 2005b) that are reflected in and influencing urban prices.

Urban real estate markets also exhibit specific characteristics such as durability, slow and costly adjustment processes, spatial rigidity, limited information and inconsistencies in supply (Whitehead, 1999, p. 1565). These peculiarities of the market and its environment lead to valuation practices that are heavily dependent upon an individual's knowledge of market idiosyncrasies. The Brazilian literature (Liporoni, Neto *et al.*, 2003) and technical legislation (Abnt, 2004) on this theme are based on everyday practice in which "market values" which are determined by experienced brokers, regulate the valuation system. The international literature also shows the importance of this discretionary approach (Kauko, 2002).

Identity of places: the roles of neighborhoods

A crucial factor to consider when analyzing urban real estate prices is the notion that prices reflect the perception of the population in terms of what the city and its parts represent. This is relevant for real estate brokers. Most studies in economics include a large number of variables without considering the notion of identity of places. The most commonly applied variables, such as distance to highways or to a healthcare center, do not contain information on the spatial unit of observation. People, on the other hand, learn about space and location cognitively, which is equivalent to saying that they attach meaning to portions or segments of the city (Tversky, 2003). We argue that neighborhood identity effects on real estate price determinations are not addressed optimally by previous research, and our research thus tries to contribute to this deficiency.

Remarkably, urban economics or real estate finance manuals do not touch on the subject of neighborhoods in the sense that they represent a community with an identity (Durlauf, 2004). Some work (Bourassa, Steven C., Hamelink *et al.*, 1999) on real estate price determination has focused on sub-markets, aiming to find clusters of dwellings with homogeneous characteristics using data-mining techniques. Megbolugbe (1996) states that this is one of the definitions of neighborhood. We believe, however, that the definition that states that neighborhoods "possess an identity or social cohesion" (also proposed by Megbolugbe) is more likely to influence real estate prices. Indeed, the neighborhood as a socio-spatial construct matters for the economic modeling of real estate.

Social science provides useful insights on why neighborhoods, identity and segregational spatial patterns matter (Blakely, 1997; Galster, 2001). However, they do not relate them explicitly to real estate prices.

The dynamic and actor-based nature of urban economic models

In relation to the cross-sectional analysis usually performed in urban economics, our understanding is that prices of urban real estate in neighborhoods in a given time period influence prices in neighborhoods in the next period, i.e., the market reflects positive and negative feedback. This fact suggests analyses should be dynamic. However, the literature on dynamic models of urban patterns (Benenson and Torrens, 2004) does not focus on the income-level dependent behavior of actors, but instead predominantly focuses on 'residential housing' in general (as opposed to other land-uses such as forestry, industry or agriculture). Even the distinction of urban actors in the encompassing study of Hagoort (2006) includes only one actor, identified as 'residential'. This division does not suffice, as the literature has shown, because the residential occupation behavior of dwellers with different income levels exhibits different patterns (Zietz, Zietz *et al.*, 2008). In fact, an analysis that purports to be intra-urban cannot fail to address internal social disparities of the city and its influence on price determination.

Finally, economic analyses of urban real estate markets, such as those by Dowall (1995), Lucas and Rossi-Hansberg (2002) and Wheaton and Nechayev (2005), employ general approaches and focus on the city in a static manner. Other authors, such as Henderson (1974), Fujita, Krugman and Venables (1999) and Capello (2002) discuss relationships among cities, city-hierarchy and optimal city size. These studies have in common the consideration of the city as a homogeneous object. If one considers Soja's (2000) Postmetropolis, which is heterogeneous and multifaceted, as we do in our research, a larger scale of analysis is necessary.

Insights and motivation for the study

Potentially, the two strands of literature (urban economics and complexity modeling) together can conceptually cross-fertilize each other. The interplay of insights generated by urban economics on the one hand, and the social action-space of actors on the other, is central to our analysis. For this, we focus on the inclusion of cognitively-defined neighborhoods in the economic explanatory models, as well as on better modeling of (local) negative feedback that considers social heterogeneity. This is missing in the current literature. Specifically, Table 0-1, in a simplified manner, highlights the insights and strengths of both approaches and our contributions. First, the essential idea of negative feedback, in general, and land price, in particular, from urban economics is not clearly differentiated, although it is implicitly present in CA. We propose to make prices explicit in CA modeling. Second, local influences are at the core of CA modeling; however, they are present only marginally in urban economics through homogeneous sub-markets and distance-based spatial analysis. We argue that this is an essential feature of the market that should be fully incorporated into urban economics analysis. Third, social heterogeneity is present at the theoretical level in urban economics, because it allows for agents with differentiated preferences, but not at the level of the city, which is usually conceptualized as a homogeneous plain. CA models do not impose homogeneity on actors;

Table 0.1 Insights and essence of both approaches

	Approaches	
	Urban economics	**CA modeling**
Negative feedback (prices)	+	implicitly
Neighborhood (local influences)	implicitly	+
Social heterogeneity	implicitly	-

however, most empirical studies apply homogeneous 'residential' categories. We argue that social heterogeneity is crucial for both approaches.

The view of this thesis is that for a better understanding of the complex urban environment and intervention in it, policy-makers and urban scientists should have a comprehension of the city that stretches beyond urban morphology and one that is intertwined with economic factors. Our work emphasizes: (a) the role of the neighborhood and its identity, (b) the heterogeneous gamma of urban characteristics associated with real estate prices, and (c) the differentiated behavior of socially heterogeneous residential actors. This is analyzed via both static and dynamic models.

0.1 Research questions

This thesis hypothesizes that the urban economic and cellular automata approaches of urban development are mutually related, and that both strands of research gain from incorporating elements from each other. The empirical analyses in this thesis extend models in both strands through the cross-fertilization of concepts and causal mechanisms. The outcomes provide useful information for policy-makers responsible for socioeconomic development in cities. As such, the two research questions addressed herein are:

(1) Do urban real estate price models improve when considering neighborhood identity and social heterogeneity?
(2) Do cellular-automata models on urban development increase understanding of urban dynamic processes when including social heterogeneity and negative economic feedback mechanisms?

0.2 Research design and methodologies

To answer these questions, several theoretical, methodological and empirical issues are addressed. The thesis is divided into two parts. The first part reviews the main theoretical approaches from an urban economic perspective and also presents alternative views. The second part focuses on applications based on theoretical discussions.

This thesis is organized as follows (Figure Introduction.0-1).

The first chapter focuses on how urban economists have historically treated the city and real estate prices and the essence of their approaches. Economic factors are among the main driving

forces of urban dynamics. The chapter presents the urban economics literature starting with Alonso's seminal book (1964). It further describes the contributions of Mills and Muth and the non-monocentric model elaborated by Fujita and Ogawa (1982). In terms of applied work, it briefly describes hedonic price model functions (Sheppard, 1999; Anderson and West, 2006).
We then continue to highlight what the urban economic approach misses in terms of capturing social and spatial heterogeneity in cities. Chapter 2 presents criticisms of the neoclassical economic approach and investigates alternative theoretical views that come mainly from self-organizing and complex systems (Arthur, 1999). This strand of literature is based on disaggregated thinking, and local interactions have been used successfully in economics (Holland

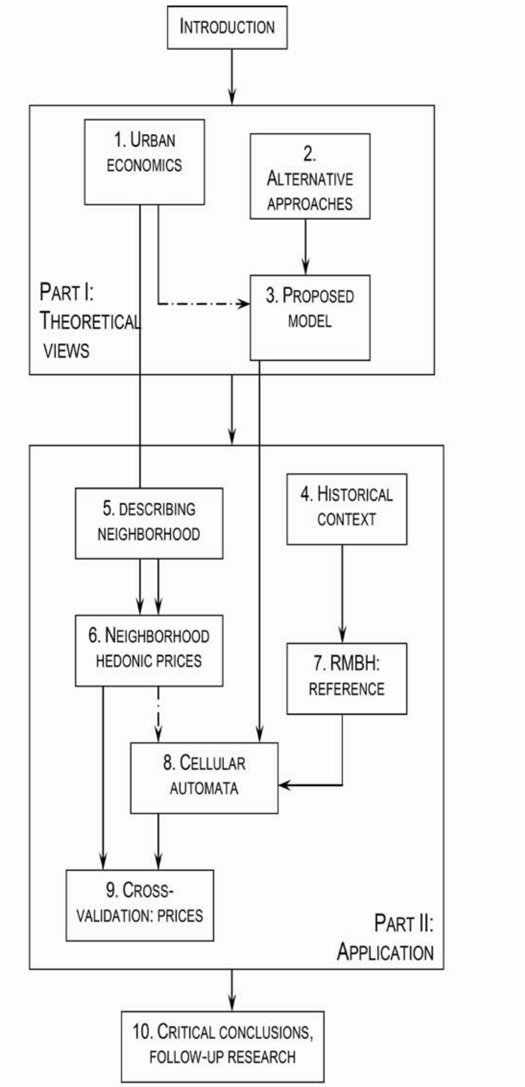


Figure 0.1 Research design

and Miller, 1991) and urban planning (Page, 1999). We focus on the seminal constrained CA model of White and Engelen's (1993), and then propose, in chapter 3, an extension to the original CA model that includes negative feedback (Krugman, 1996; Batty, M., 2005b), and social heterogeneity. This is done to achieve better comprehension of the mechanisms that, together, affect both spatial configurations of actors and relative prices in urban agglomerations.

The second part of the thesis focuses on applications based on theoretical discussions. An empirical analysis, within the rapidly growing, socioeconomic disparate context of the Greater Area of Belo Horizonte (RMBH), Brazil, is conducted to answer the research questions.
To provide context for the case study, it is necessary to describe the specificities of the case of RMBH in chapter 4. This includes a brief description of rapid population growth, with an emphasis on the intensified growth in the second half of the twentieth century. Chapter 4 also introduces the process of urban land production that took place in the 1950s and 1960s and describes the evolution of the boundaries of municipalities of RMBH, the legislation imposed on urban occupation, and present legal arrangements.

As neighborhoods and sub-markets and its associated attributes are central to our analysis, chapter 5 describes why social environments are potential attributes that qualify real estate markets. The chapter introduces the neighborhood as a spatial unit based on the idea of identity as a reference for cognitive perception. A key focus of the approach is that neighborhoods are considered to be representative of a perceived value that is both identifiable by citizens and the result of urban characteristics. This is based on the theoretical discussion of micro-scale urban landscape described by Cullen (1961) and further developed at the city level by Lynch and Camargo (1960), Rossi (1966), Tuan (1980; 1983), and Tversky (2003). Geoprocessing techniques and principal component analysis are used to aggregate spatially distributed socioeconomic data by neighborhood. This produces a detailed description of the urban landscape, at the neighborhood level.

Together, these chapters enable and support an answer to the first question ("do urban real estate price models improve when considering neighborhood identity and social heterogeneity?") in chapter six.

Chapter 6 presents an econometric analysis, based on the information introduced in previous chapters, which can be characterized as static, cross-sectional urban-economic analysis. The analysis aims to explain real estate prices by estate attributes, *vis-à-vis* attributes of the urban tissue. The results also provide empirically determined indications of parameters for the CA model in chapter eight. The city of Belo Horizonte is the empirical case. After a description of the database, the model, results, tests and comments are presented. In terms of methodology, the econometric model in chapter 6, given the nature of the problem, is tested against its spatial component following the approach initiated by Cliff and Ord (1973), which was organized by Anselin (1988) and further developed in the 1990s[1]. Furthermore, because there are indications that the real estate market follows different patterns for different income-groups (Zietz, Zietz *et al.*, 2008), quantile regression analysis is incorporated and tested in the model. To do this, we follow the work of Buchinsky (1997), who advanced the original work of Koenker and Basset (1978), and recommends verifying that the patterns obtained in a regular regression (valid on the mean) repeat in quantiles. If not, the quantile analysis is more appropriate. An Instrumental Variable Quantile Estimation (IVQR) of Spatial Autoregressive Models (SAM) proposed by Su and Yang (2007) is applied to the dataset. The estimation method has been shown to be robust for heteroskedasticity and outliers, performs better than General Moment Methods

(GMM) or Maximum Likelihood (ML), and most importantly, applies quantile and spatial analysis simultaneously. That is, it accounts for the fact that different rules and emphases in economic behavior apply for different income-class members. This is also incorporated in the CA application of chapter 8.

A central question in CA simulation models is its validation. As for the case study of the Greater Area of Belo Horizonte in Brazil, there is no information available for historical actors. It is thus necessary to interpolate this data. Therefore, in chapter 7, census tract data is used to produce a physical configuration of residents in the RMBH, classified by income. The chapter explains the methodology of supervised clustering to produce maps for the years 1991 and 2000. Chapter 7 also presents map quantification metrics that provide a supporting role when comparing the reference map with the result of the CA simulation.

The second research question ("do cellular-automata models on urban development increase understanding of urban dynamic processes when including social heterogeneity and negative economic feedback mechanisms?") is answered in chapter 8. The model proposed and described in chapter 3 is applied to the Greater Area of Belo Horizonte. That is, a CA model of the spatial configuration of income-differentiated actors and relative prices is applied that incorporates quantitative and qualitative results from previous chapters. The model dynamically incorporates spatial interaction among actors and therefore provides further insight into the logic behind price formation. The chapter presents the objective, the dataset used in the model, and the parameters chosen. The explanatory power of the model is also discussed. The results of the application are validated against the spatial configuration of actors (produced in chapter 7). A sensitivity analysis is performed to provide further insight into the various inputs of the model.

Chapter 9 provides an analysis in which the price-distribution derived from and mapped by the CA simulation of chapter 8 is further validated using the econometric data and model in chapter 6.

Chapter 10, the final chapter, reviews and suggests follow-up research. It also provides a summary of the results of both the econometric analysis derived from urban economics with a neighborhood emphasis and the CA model adapted with relative prices, illustrate how each one of these contributes to the understanding of urban dynamics.

To highlight the specificities of each chapter the following summary is presented (Table 0.2).

To summarize the added value of the thesis, five features can be noted. First and foremost, this thesis brings local influences and neighborhoods to the forefront of the discussion. Second, it explicitly considers spatial and social heterogeneity within a rapidly growing urban environment. Third, it provides a model adaptation to further include negative feedback to CA urban development simulations. Fourth, it tests and applies the new methodology of IVQR to a case study. Fifth, it bridges urban study fields and offers policy-makers a specific and more substantial hermeneutic alternative.

Table 0.2 Summary of chapters analyses

chapter	theme	concepts and objectives	methodologies	analyses
Introduction		background, research questions and design		
1	essence of urban economics	residence; Alonso's model; Muth-Mills model	literature review	theoretical
2	alternative approach	city; complex and self-organizing systems; White & Engelen's model	literature review	theoretical
3	adapted model	negative feedback and social heterogeneity		theoretical
4	historical context	historical background; speculative process	literature review	empirical
5	neighbourhood	neighborhood definition; description of urban tissue by neighborhood	principal components; GIS	empirical
6	real estate price determination	price determination	spatial econometrics; quantile econometrics, IVQR	empirical
7	reference data	location of diverse social groups	clustering analysis; GIS	empirical
8	cellular automata application	price and location determination	cellular automata; GIS	empirical
9	CA and econometric analysis' results	cross-validation: prices	spatial econometrics	empirical
Conclusion		critial review; research follow-up		

type of analyses	time-space focus	key explanatory phenomena	case study	specific questions answered
review	static; a-spatial	real estate market		
review, critical	dynamic; spatial	urban morphology; city definition		
model formulation	dynamic; spatial	price and locational determination		Question 1
descriptive	dynamic	urban dynamics	Minas Gerais, RMBH, Belo Horizonte	
qualitative and quantitative analysis	static; spatial	spatial and social heterogeneity	Belo Horizonte	
analytic	static; spatial	spatial and social heterogeneity	Belo Horizonte	Question 1
descriptive	static; spatial	spatial and social heterogeneity	RMBH	
analytic	dynamic; spatial	spatial and social heterogeneity	RMBH	Question 2
analytic	dynamic; spatial	cross-validation	Belo Horizonte	

Part I

This first part of this work is made up of the first three chapters, which delineate the theoretical view of the dissertation. The first chapter relies heavily on urban economics theory and discusses in detail the seminal work of William Alonso, who laid the foundations of subsequent developments in the field. The second chapter reviews criticisms of urban economics and presents self-organizing and complex systems, along with cellular automata methodology as an alternative approach to handling dynamics in urban systems. Chapter 3 highlights the advantages of each methodology, proposes a theoretical balance that incorporates the Brazilian context, and details the model and adaptations that are used in the application in chapter 8.

1 Theoretical considerations regarding urban economics

This chapter starts with a description of the residence as a differentiated good, followed by a review of fundamental authors of urban economics studies. Alonso's (1964) model is described in detail, followed by Brueckner's (1987)[2] synthesis of the models developed in the 1960's and 1970's, and some more recent developments, including those of Fujita and Ogawa (1982), and Wheaton (2004) among others. The last section introduces concepts such as urban amenities and hedonic prices. This chapter establishes the views of urban economics that will be compared with the view in the following chapter.

1.3 Residence as a differentiated good

There are a great number of reviews of the literature on the issue of the residential urban market (Quigley, 1979; Arnott, 1987; Mills and Nijkamp, 1987; Whitehead, 1999), especially in the United States. According to Whitehead, a housing unit is intrinsically a differentiated good, especially because of its characteristic of "durability, heterogeneity, spatial fixedness, slow and costly adjustment" (Whitehead, 1999, p. 1562). Arnott (1987), in turn, highlights residence as (a) a good that fulfills a need; (b) an important item of consumption (most times, the most important item); (c) an indivisible good; (d) a heterogeneous good; and (e) a multi-dimensional good. He also reinforces its characteristics of durability and spatial rigidity.

Other commentaries on the discussion of the residential market indicate considerable potential for model failure (Whitehead, 1999, p. 1565). These remarks suggest that the "housing market is more complex than can be described by simple, single market, competitive equilibrium models" (*op. cit.* p. 1577). According to Whitehead (1999), there is evidence that the typical complexity of the city is associated with its strong temporal structure, demand complexity, and thus a dynamic model.

Furthermore, other features that make the real estate market unique are the notions that (a) the processes of selection and decision involve high costs and uncertainties; (b) the market is intrinsically spatial, enhancing the need to analyze the residence in accordance with its neighborhood, proximity factors, and urban amenities, which create strong positive and negative externalities; and (c) the market is composed of new and already existing residences (Plambel, 1987, p. 26; Sheppard, 1999, p. 1603), reflecting the idea that the price of a residence is not solely related to its present production cost.

Commenting on the longevity of urban structures, either public or private and its historical importance to the organization of the city, Anas, Arnott and Small (1998, p. 1460) write that "the urban structure locks in past forces that may have little bearing today," meaning that prior phenomena can still be determinant in the constitution of the present city.

Finally, the spatial analysis of the market completely changes the object. Because of its importance space should be intrinsically considered in the development of theoretical proposals and in the making of models so that analyses of the real estate market become realistic and robust. Therefore, space should be present in the theoretical framework as a fundamental and seminal principle. This understanding would provide public policy-makers, legislators and other actors with a more comprehensive view of the market.

1.4 Alonso and his followers

As a result of his doctoral thesis, William Alonso published *Location and land-use: towards a general theory of land rent* in 1964. The book may be considered the starting point of the contemporary analysis of urban land markets and is the reference for theory developed since then. The chapters of the book are organized as a classic review of the *economics of urban land*, followed by the analysis of family equilibrium, the adaptation of von Thünen's theory of agricultural areas into an urban firm scope, and the modeling of residential bid price curves. After that, Alonso proposes a market-clearing model and empirical examples. To provide context for Alonso's work, some of his analyses of the works of previous authors are briefly mentioned.

The focus of Alonso's analysis is the residential market and the association therein of land price with land-use (1964, p.2). Before his work, in the nineteenth century, land analysis was bounded to agricultural land, *vis-à-vis* less important and still abundant urban land. This is also true in urban history as defined for instance by Benevolo (1980), who understood that the concept of the city itself and urban phenomena were only socially-culturally effective after the Industrial Revolution.

The theory available at the time was that proposed by Johann Heinrich von Thünen (1826). He described a sort of competition among farmers who produced different crops that competed in the market, with the winner getting to occupy the land. In this model, "rent at any place is equal to sales price less transport and production costs" (1964, p. 4). According to the proposed logic, there is no reference to the neighborhood, but rather to the cost of transport to a dominant center. This emphasis is reinforced in the subsequent development of local theories (Isard, 1956; 1998).

Richard M. Hurd published *Principles of City Land Values* in 1903 and contributed to Alonso's thinking by emphasizing the importance of causality between prices, proximity, and social aspects in price formation. Although these factors remain essential, the development of urban economics highlights social and spatial homogeneity in their theoretical construction. Recently, authors such as Brueckner (1999) and Wheaton (2004) have tried to include heterogeneity in the analysis while using the same general equilibrium framework.

Robert M. Haig (Haig, 1926), also cited by Alonso, contributed to the definition of spatial friction and its association with accessibility and cost of transport. Haig relates spatial influence with spatial permeability. The more accessible a place is in relation to the other the more interaction there is between them.

Simultaneously, Chicago's ecologists, especially Robert E. Park and Ernest W. Burgess (Park, Burgess *et al.*, 1925), contributed to the theory with their sociological approach and the influence of land price as a cause of residential segregation. According to them, residential rent is an indicator that unites land price, the price of relative location (its accessibility), and transport costs. This theory originated with what Portugali (2000) called the "*Social theory of the city,*" which was a distinct approach from that of Regional Science and the New Urban Economics.

Park focused on the transformation of competition rules – taken as intrinsic to an individual's ability to survive in a community – and the construction of community consensus as a constituent element of a society. The ecological sense of this approach was based on the concept of the natural area developed by Park, according to which natural competitive forces tend to produce a balance that shows social adaptation to the urban environment. The establishment of the required processes of competition, domination, succession, and invasion of natural areas led to their most well-known model of organization and expansion of the urban form: the five concentric zone model proposed by Ernest Burgess (1925). The result of this model was a natural segregation by common interests and values that composed the so-called urban mosaic[3].

Alonso also criticized the model of Paul F. Wendt (Wendt, 1958). He observed that Wendt's analysis, which criticized the so-called "Haig-Ely-Dorau-Ratcliff *hypothesis,*" presented little mathematical accuracy and concentrated on a long and cyclic view. Furthermore, Wendt's proposal discussed the relations that are present in the composition of urban land without elaborating exactly on how these relations formed.

Lowdon Wingo Jr. (Wingo, 1961) also contributed to the basis from which Alonso derived his model by adding transport flows to the traditional analysis of urban land markets. His proposal is actually very close to that of Alonso's, but it was independently developed. In summary, in Wingo's model, rents and transport costs are complementary; their sum is equal to a constant, which, in turn is equal to the transport costs to the most distant occupied residential location. Transport costs are taken into consideration as part of temporal and monetary aspects.

Prior to the development of his model, Alonso defined his approach and defined some basic concepts. City is considered to be a featureless homogeneous prairie where land is bought and sold freely by actors who have perfect knowledge of the market Emphasis is given to price determination (and not to the nature and components of its value). The size of the housing unit is explicitly considered. Alonso defines price as "the price that a buyer is willing to pay after considering alternatives" (1964, p. 16).

Alonso then presents his family equilibrium. He assumes that the price of land decreases as a function of distance to the city center, and that a location decision implies a given land price. *Ceteris paribus,* rational individuals always prefer a more accessible location.

According to Alonso, then, an individual presents higher utility when distance (t) to the employment center is closer. If the choice of the individual is made considering consumption of housing surface (q), transport costs $k(t)$, and a given bundle of other goods (z), in any given combination of q and z, a small negative variation in any of them should be compensated by a proportional positive variation of the other so that the utility of the individual is constant. Conversely, small increases in distance (t), if q remains constant, should result in increases in z (Alonso, 1964, p. 21, 26, 27).
Formally, an individual's budget equilibrium is given by:

Equation 1
$y = p_z z + P(t)q + k(t)$, where:

y = income; p_z price of composite good (non-residential); $P(t)$ is the cost of land as a function of distance to the center; $k(t)$ is the cost of transport as a function of distance (*id. ibidem*, p. 31). The problem of the individual, therefore, is to establish levels of z, q and t that satisfy this equation and, simultaneously, guarantee maximum utility:

Equation 2
$U = U(z, q, t)$

The solution of the model is given in the simultaneous resolution of the budget equation (equation 2) and the following equations, which result from the maximization of the utility of the individual:

Equation 3
$u_q/u_z=P(t)/p_z$ and

Equation 4
$u_t/u_z = (qdP/dt+dk/dt)p_z$

The analysis of the results of the proposed model indicates that the marginal rate of substituton between surface (q), and the bundle of composite goods consumed (z), i.e.,: U_q/U_z , is equal to the ratio of their marginal costs. The marginal rate of substitution between distance to the center (t) and z, (U_t/U_z) is equal to the ratio between marginal cost of spatial movement *(qdP/dt + dk/dt)* and the marginal cost of the bundle of other goods p_z. In other words, the numerator expresses the variation of the price of the land as it varies its location (*dP/dt*), multiplied by the quantity of land (q), plus the variation of transport costs (*dk/dt*) (Alonso, 1964, p. 34).

Chapter three of Alonso (1964) uses the solution presented above to propose market equilibrium. In a nutshell, the agricultural framework of von Thünen serves as reference for an adapted model for the urban firm.

Briefly, Johann von Thünen's model is designed for an isolated city with a central nucleus where the market is located and where all products are marketed. A radial transport system covers the entire city and there is no difference between soil fertility. Note that homogeneity is, again, central to the construction of the theory.

The rent of the land according to von Thünen's approach is defined as the payment made to the owner of the land, which is equivalent to the difference between the sales price and the costs. If considered, the assumption of zero profit (or as Alonso puts it, "normal profit"), the bid rent function can be described as:

Equation 5
$p_c(t)=N[P_c\ C\ K_c(t)]$, where:

$p_c(t)$ is the price of rent of one unit of land that is distance t from the market; N is the number of crop units produced by units of land; Pc is price of crop units in the central market; C is the cost of production of one unit of product; and $K_c(t)$ is the cost of transport of one unit of product that is distance t from the market (Alonso, 1964, p. 38). Generalization of the results to many products indicates that every crop (and its derived cost of production) can be associated with a family of indifference curves where it is possible to pay the rent for the land. Market-clearing

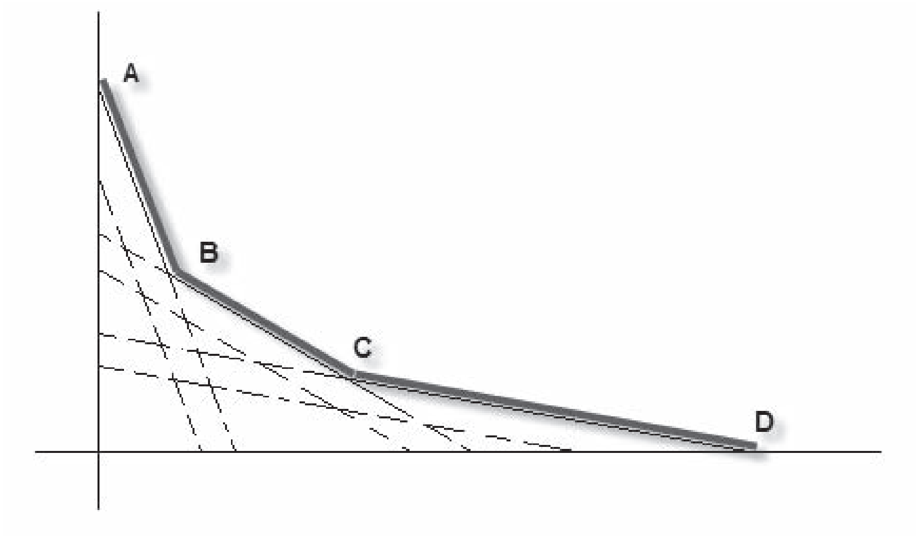


Figure 1.1 Diagrammatic scheme of von Thünen's model for agriculture land Source: Alonso, 1964.

establishes that the winning land-use is that which is able to yield the greatest value for a given location. In Figure 1-1, the winning land-use segments AB, BC, and CD.

Before discussing the equilibrium of urban firms, Alonso mentions the work of Chamberlain (1950). Chamberlain was the first to suggest that an adaptation of the agricultural location theory would apply to urban land-use. An important insight he provided was his suggestion that the rent paid by retail firms is, in fact, derived from monopoly[4].

Alonso's urban retail firm considers that the business:

> "is faced with a given structure of prices for land, according to distance, which is described by the function **P(t)**. [The business owner] will decide on his location and on the amount of land he wishes to occupy in such a way that he will make the greatest possible profits. These profits may be defined as the remainder after operation costs and land costs have been deducted from the volume of business (1964, p. 45)."

Formally, the model (1964, p. 45) is expressed by three basic equations:

Equation 6

$G = V - C - R$, where:

G = profits; V = volume of business; C = costs of operation and R is the land rent;

Equation 7

$V = V(t,q)$

Equation 8

$C = C(V, t, q)$

plus the equation of the land rent (***R***), which is the number of units of land multiplied by the price of the land at a given location ***P(t)***, i.e.:

Equation 9
$R = P(t)q.$

The model solution can be expressed as:

Equation 10
$0 = dt(V_t - C_v V_t - C_t - q\, dP/dt) + dq(V_q - C_v V_q - C_q - P)$ [5].

Because Alonso's proposed analysis assumes static equilibrium, two other equations can be derived from this one. That is, when surface (***q***) is determined, there is no variation in ***q*** (***dq*** = 0) e ***dt*** ≠ 0, so that:

Equation 11
$0 = V_t - C_v V_t - C_t\, q - dP/dt.$

Furthermore, when the locational decision is taken to be and $t = t_0$, the result is ***dt*** = 0. As ***dq*** ≠ 0, the value inside the parentheses is 0, and therefore,

Equation 12
$0 = V_q - C_v V_q - C_q - P$

Equation 11 is named by Alonso the *location equation*, and it indicates that the loss of marginal revenue for having moved far from the center (V_t) is equal to the marginal operational cost ($C_v V_t$). This cost is probably negative once there is a reduction in the land rent (***qdP/dt***), resulting in the variation of the land price. That is equivalent to saying that a movement far from the center reduces the amount of business that can cause an increase in the marginal operational cost and – because of a volume reduction – an increase in the operational cost due to a displacement (of goods and employees), that should be compensated for by a reduction of price land (1964, p. 50).

Equation 12 is called the *size of site equation* and signals that the variation in the sales volume, given by the largest business surface (V_q), is equal to the variation in operational costs resulting from this (indirect) variation in the size ($C_v V_q$), added to the direct variation in operational costs (C_q) and marginal costs in a unit of area (***P***).

Alonso determined the *bid price function* for the urban firm (1964, p. 52) similarly as to that in the agricultural analysis. The aim was to construct a function for urban firms that is indifferent among locations for a given price. Alonso supposed that different actors (households or firms) follow a distinct balance logic. The observed result of the interaction among different actors corresponded to the actors who evaluated a certain location at a higher value than other actors.

The function observed by the firm "is a hypothetical price-of-land-with-distance function, and might be termed an iso-profit curve" (1964, p. 52). This function is denoted in a simplified manner[6] as ***p(t)***. The corollaries used in the derivation of potential price function are demonstrated in Alonso's appendixes and show that (a) at any level of profit (G_0), there is only one value for (p) in any location (t); (b) two corresponding functions at two distinct levels of profit do not cross; (c) the lower the function is, the higher the level of profits, and therefore, this is preferable from the firm's perspective; and (d) the functions usually present a negative inclination.

In the agricultural analysis, the profit of the farmer is constant and there are variations in the price of the crop harvested. In the case of the urban firm, each curve represents a different level of profit. In addition, in the case of agriculture, what determines the winning crop (the one that actually occupies the land) is the sales price of the product, whereas for urban analysis, it is the firm that is able to pay the highest price while maintaining a zero-level profit.

Alonso's chapter 4 (1964) is where he derives the potential price curve and the market-clearing equilibrium specifically for the individual citizen with differentiated preferences. However, potential price is not necessarily the real price paid, but the highest possible to be paid hypothetically.

Individual analysis is conducted in a similar manner to the analysis in chapter 2 (equations 2, 3 and 4), but the individual's utility (u_o) is such that the level of satisfaction remains unchanged. The microeconomic hypotheses remain. They are: (a) a single value for a potential price curve; (b) inferior curves present higher utility; (c) curves do not cross (Alonso, 1964, p. 69). The result of the individual maximization of utility may be rewritten as:

Equation 13

$$\frac{dp_i}{dt} = \frac{p_z}{q}\frac{u_t}{u_z} \quad \frac{1}{q}\frac{dk}{dt} \quad .$$

The right-hand side of the equation depicts two elements that contribute negatively to the potential price function of the individual: the loss of utility, in which distance (t) increases has to be compensated for by higher expenditures on composite goods (z) ($u_t/u_z < 0$), and the cost of transport, which increases with distance ($dk/dt > 0$). Alonso's definition for the potential price curve for the individual is thus expressed:

> "Bid price, then, has been defined so that the income effect of cheap land will counteract the depressing effect of commuting costs on income and will permit the consumer to maintain a given level of satisfaction by substituting land and the composite good for accessibility as distance from the center increases (1964, p. 71)."

Formally, the slope of the potential price curve is expressed by the marginal rate of substitution between the distance to the center (t) and the composite good (z) that equals the ratio of their marginal costs[7]:

Equation 14

$U_t/U_z = (qdp_i/dt + dk/dt)/p_z$

The definition Alonso gives for the individual does not specify any item in relation to preferences related to his neighborhood. Two locations with the same distance to the center have the same utility independent of neighbors or neighborhood. Studies on racial segregation (Rose-Ackerman, 1975), for example, prove, at least for central countries, the importance of this factor in the locational choice.

The land market-clearing equilibrium is proposed by Alonso in chapter 5. His line of reasoning follows a comparison among all potential users of the land in which the user who is able to make the highest bid occupies the land and determines its use. He then concludes that prices at a given location are associated with the prices at other locations through potential price curves. The price actually paid is equivalent to the value of its potential price curve, which, in turn, is determined by the price of the next preferred location (*idem ib.*, p. 81). Formally, for individual *i*, whose equilibrium location is t_i and whose immediately superior location is t_j (at prices P_j), we have:

Equation 15

$P_i = p_i(t_i) || t_j, P_j$[8]

To achieve this result, Alonso applied successive price determinations to find the market equilibrium, starting at the origin of the system – the center of the city – where $t = 0$. Thus, users with higher slopes locate themselves closer to the center. The price of equilibrium paid by a user is equivalent to the potential price that a user immediately in front of him (towards the center) would pay for that same location. For individuals, the curve slope is the result of different preferences. Individuals who are most averse to commuting, for instance, need to be more generously compensated from having to move away from the center and have as a result a more inclined curve. Similarly, firms whose sales diminish rapidly as they move away from the center of consumption (or firms whose employee and goods transport costs increase too quickly) offer better prices to closer locations.

Apart from the type of land-use (individual, firm, and agriculture), the use with the steepest slope of potential curve locates closer to the center and uses with the smoothest slopes locate at the most distant available location[9]. Actual equilibrium is calculated iteratively, preferentially, and starting at the center. The author points out that this solution is only valid for well-behaved curves. He discusses their variations in section E of chapter 5.

Accordingly, regardless of the actor, the dispute for space occurs among all of actors. The potential transitional logic that is used in the development of the model with the actors in chapter 4 also observes this rule. Actors that can afford better locations impose their choices.

Alonso's final chapter amplifies the debate and considers the practical implications of the proposed model. For instance, zoning implies that, in a certain zone, market-clearing is not effective and a sub-optimal level is established. Consequently, the supply of land as a whole decreases and prices in other locations are adjusted upwards[10].

When there is an increase in income levels (y), during economic expansion, for example, results are not generalized, but it may be inferred that individuals with higher incomes (and who own more land) would worry proportionally less about transport costs and would therefore be "location-oriented"; in contrast, individuals with lower incomes would be "price-oriented" (1964, p. 109).

An exogenous increase in population shifts the potential price curve upwards, resulting in a combination of higher prices, greater density, and physical expansion of city boundaries. Technical improvements in the transport system – which reduce transport costs or commuting times – rotate potential price curves which, in turn, determine lower prices in central areas and higher prices at the periphery.

Tax effects, according to Alonso, might generate two results: (a) when tax is charged by percentage, the cost is assumed by land owners and there is no change for users; (b) when there is a fixed uniform tax on all land, it is the user of the land who pays for it, which results in a price above 0 at the most outward boundary (1964, p 116).

Alonso's appendix C (1964, pp. 168) is interesting because it aims to demonstrate or, at least suggest empirically, a theory-proof path. He proposes that "expenditure on land by a family is determined by their income and their distance from the center of the city." Formally, we have:

Equation 16

$$pq = g(y,t),$$

i.e., land rent – established by its price – times quantity (***pq***) is determined by a function that depends on income (***y***) and the distance to the center (***t***). The expected sign is positive for income ($dg/dy > 0$) and negative for distance ($dg/dt < 0$). An empirical test was applied using income data from the smallest unit of U.S. census data (U.S. census tract). The same source provided information about the population density, which was used to calculate land quantity (***q***). Price was obtained by considering real estate sales price (urbanized areas in which residential use is predominant). Where much information on surface price was available, its average was used. In the example presented, there were 52 observations in 32 different census tracts. He applied it only to regions that had grown more than 2% in the previous decade. The results were:

$$pq = -222.65 + 0.4357y - 90.107t; \ (R = 0.6897).$$

Alonso's interpretation of the residuals is that they reflect an "attractiveness measure" of a specific location, not specified in the model. This interpretation is also used by Hermann and Haddad (2005). In the context of a heterogeneously social, spatially diverse city, understanding how the characterization of urban space – represented here as a residue – impacts land price would be theoretically central.

Alonso's final commentary, which is relevant to this thesis, is that "there is, to be sure, the weight of opinion of many writers, who attribute considerable importance to **land values as the basic distributing factor in the spatial structuring of the city** (1964, p. 125 emphasis added).

1.4.1 Muth-Mills model

In the late 1980's, Jan Brueckner (1987) synthesized the models of Muth (1969) and Mills (1967), derived from Alonso (1964), in what he called as "unified treatment of the Muth-Mills model". The objective of the described model is to capture the observed regularities of city's spatial structure that are the most significant to the variation in intensity of urban land-use. The main motivation for the construction of such a model is the empirical observation that financial and time costs in urban travel must be compensated for by different housing prices (Brueckner, J., 1987, p. 821). Different from Alonso, Muth and Mills' proposals use land as an intermediate input for housing, which is the final consumed good. According to Brueckner, Wheaton's contribution of 1974[11] is also incorporated in the synthetic model.

Brueckner's stylized city[12] has a single center business district (CBD) to which all inhabitants commute. All consumers have the same income (***y***) and preferences and pay the same transport cost (***t***), which varies depending on the distance to the center (***x***). The well-behaved utility

function of consumers varies according to the choice of a bundle of composite goods, which does not vary spatially, (*c*) and housing, which is qualified by a single attribute, its surface, (*q*). The price of rent (*p*) varies depending on the location. Because consumers are identical, their utility (*u*) is fixed. This leads to the determination of rent price.
The model resolution indicates that:

Equation 17

$$\frac{\partial p}{\partial x} = \frac{-t}{q} < 0$$

which reveals that price is a decreasing function of distance from the center. The author confirms the intuition that "consumers living far from the CBD must be compensated in some fashion for their long and costly commutes (otherwise, no one would live voluntarily at great distances)" (Brueckner, J., 1987, p. 824).

The analysis of partial derivatives contributes to yield the results of the model.

Equation 18

$$\frac{\partial p}{\partial y} > 0 \text{, and } \frac{\partial q}{\partial y} < 0 \text{; and } \frac{\partial p}{\partial t} < 0 \text{ and } \frac{\partial q}{\partial t} > 0$$

It is easy to see that an increase in income causes an increase in the price of housing[13], whereas augmenting transport costs has the opposite effect: a reduction in prices and an increase in surface consumed for housing.

If one considers housing as a normal good, then, according to Brueckner (1987, p. 826), we have:

Equation 19

$$\frac{\partial p}{\partial u} < 0 \text{, and } \frac{\partial q}{\partial u} > 0$$

which indicates that utility is a decreasing function of prices and an increasing function of housing surface. The analysis of the supply made by Brueckner assumes a production function that depends on land (*l*) and capital (*N*) that is "perfectly malleable", which implicitly disregards the low technical elasticity of substitution of the structures.

Once again, in the context of this work, it is believed that there is an irreversibility of structures that are historically built, and that determine the evolution of the urban structure (Anas, Arnott *et al.*, 1998). In addition, numerous authors present residence as a differentiated good (item 1.3).

Brueckner introduces two new parameters: the rent of the land (*r*), and a structural density index (*S*) that denotes the ratio of capital to land (*N/l*).
The meaningful result of the analysis, according to the emphasis given by the author, is that:

Equation 20

$$\frac{\partial r}{\partial x} < 0 \text{, and } \frac{\partial S}{\partial x} < 0 ,$$

which suggests that "land is cheaper and buildings are smaller the farther they are from the CBD" (Brueckner, J., 1987, p. 827). The intuition behind these results is that producers demand cheaper land prices to compensate for the lower price paid for constructed surface at more distant locations. The analysis of population density (*D*), which is the result of the decision of consumers and producers, and which is derived from the model also indicates that there is a reduction of density towards the borders of the city where buildings are shorter (smaller *S*) and have a larger surface (*q*).

Having presented the model, Brueckner makes a static comparative analysis and highlights its explanatory power for different spatial structures in different cities. There are two cases: the closed-city, in which the population level (*L*) must adjust to a maximum distance ($\overline{x}$)[14] and, therefore, into which there is no migration; and the open-city, in which costless migration guarantees equal utility across cities in the economy. The author evaluates changes in each case. For the closed-city, for example, an increase in population (*L*) reveals that:

Equation 21

$$\frac{\partial \overline{x}}{\partial L} > 0 \text{ and } \frac{\partial u}{\partial L} < 0 \text{; and } \frac{\partial p}{\partial L} > 0, \frac{\partial q}{\partial L} < 0, \frac{\partial r}{\partial L} > 0 \text{ and } \frac{\partial S}{\partial L} > 0$$

there is a physical expansion of the city, lower levels of utility considering fixed transport costs (*t*) and income (*y*), higher prices (*p*) and smaller housing units (*q*), and higher prices of land (*r*) and higher densities (*S*). The intuition of the model is presented by the author:

> "When the city starts in equilibrium and population increases, excess demand for housing is created at the old prices: the urban population no longer fits inside the old X. As a result, housing prices are bid up throughout the city. On the consumption side of the market, this increase in prices leads to a decline in dwelling sizes at all locations. On the production side, the price increase causes land rents to bid up everywhere, and higher land rents in turn lead producers to substitute away land, resulting in higher structural densities (Brueckner, J., 1987, p. 831)."

Static comparative analysis further demonstrates that increases in rural land prices (r_A) reduce the surface of the city and levels of utility, i.e., according to the example provided by Brueckner, a city located in a competitive agricultural region would have a smaller size when compared to a city in the middle of the desert (*op. cit.*, p. 834). An increase in income (*y*), in turn, increases the demand for housing and consequently, an increase in the surface size of the city and of the utility of its inhabitants. Furthermore, the raise in income (*y*) reduces prices of housing units (*p*), land, and structural density (*S*) in places near the city center and it does the opposite in more distant areas. This implies that individuals with higher incomes are able to bid for larger housing units that can be found in more distant places. The opposite effect is seen when there is an increase in transport costs (*t*).

In the open-city model, utility is no longer determined endogenously, which leads to direct effects on the static comparative. Increases in the rent of agricultural land (r_A) reduce the population of the city that is concentrated in a smaller area, yet with the same structure. The increase in income (*y*), in turn, would increase the price of housing (*p*), land (*r*), and the density of spatial structure (*S*) and it also reduces the size of housing (*q*). Finally, an analysis of the open-city model suggests that cities with higher transport costs (*t*) have cheaper housing (*p*), but with smaller surface (*q*) and population (*L*).

At the end of his proposed synthesis, Brueckner (1987) reminds readers that the model refers to an extremely stylized city but the synthesis has laid a common path in which other authors have thrived in making more realistic assumptions. Fujita and Ogawa (1982), for example, work on a model in which employment is decentralized and endogenous. Wheaton (1976) presents another model in which there is a differentiation of income among consumers, which results, according to Brueckner, in no change in the essence of his proposal. The assumption of surface as the only attribute of housing is dealt with by the empirical literature of hedonic price functions (Brueckner, J., 1983; Sheppard, 1999). Other models consider the issue of housing durability by taking capital as non-malleable (Miyao, 1987). Yet, other authors such as Yang and Fujita (1983), include an analysis of public goods as spatially heterogeneous and how this fact influences the real estate market

1.4.2 Advances in urban economics

More recently, a number of authors have dedicated themselves to making models more realistic by allowing for a more heterogeneous environment. Wheaton (2004) analyzes a model with mixed use; Brueckner and colleagues discuss the weight of urban amenities in market equilibrium (Brueckner, J.K., Thisse *et al.*, 1999; Hermann and Haddad, 2005); and yet others (Glaeser, E. L., Gyourko *et al.*, 2003) analyze the impact of regulation and zoning on housing prices.

Fujita and Ogawa (1982) define a model of urban land-use based on the framework drawn by Alonso (1964), but they do so with a choice of parameters that enables multiple monocentric and polycentric equilibria. Furthermore, they emphasize the notion that cities may go through structural changes (significantly and rapidly) when from one equilibrium state to another. Even though the model is based on rather strong and simplifying assumptions, firm and household locations are endogenous, a notion that contributes to theoretical cohesion.

The justification of Fujita and Ogawa is rather solid and aims to explain a city's formation based on agglomerative and disagglomerative forces[15]. According to them,

> "The spatial configuration of the city is treated as the outcome of these interactions between business firms, which favor concentration by reason of agglomeration economies, and households, which follow closely the employment distribution (because of the costs of commuting from residences to job sites), with the consumption of urban land as the mediator of the balance (Fujita and Ogawa, 1982, p. 163)."

The city in their model is considered linear, i.e., it unfolds along a straight line and it is homogeneous with the width of the unit. From our perspective, the remarkable contribution of the model is the introduction of a potential location function, given by:

Equation 22

$F(x) = \int b(y) e^{\alpha d(x,y)}\, dy$, where:

$F(x)$ is the potential location function at x; $b(y)$ is the density of the firms at y; α is the parameter of potential that captures the importance of the firm interaction with its surrounding agglomeration at point y; and $d(x,y) = |x\ y|$ is the distance in module between x and y (1982, p. 166).

Afterwards, Fujita and Ogawa describe the equilibrium of the model using the household *(x)* and firm *(x)* functions that simultaneously determine the spatial configuration of the city.

Their numerical approach demonstrates that different values of the parameters of the model, lead to different results in configurations. They highlight that the ratio between the commuting rate and production and potential parameters are the two dominant factors in the various configurations presented.

Brueckner, Thisse and Zenou (1999) add to the Neoclassical Spatial Synthesis the concept of urban amenities. Their work aims to explaining the location of groups of residents with different levels of income at distinct spatial locations of urban amenities. They define urban amenities as a combination of three basic factors: (a) natural amenities, such as beautiful landscaping or lakes; (b) historical amenities, such as historical heritage centers, historical monuments, and landmarks; and (c) modern amenities, such as consumption and leisure areas, restaurants, and theater districts. Brueckner, Thisse and Zenou stress that modern amenities might be endogenous, especially because they are correlated with the average income level of residents (who choose their place of residence partly because of available modern amenities in nearby locations). According to Brueckner *et. al.*, "the virtue of the theory is that it ties location by income to a city's idiosyncratic characteristics. It thus predicts a multiplicity of location patterns across cities, consistent with real-world observations" (1999, p. 92).

In relation to the Alonso-Muth-Mills synthesis, Brueckner (1999, p. 93) argues that their suggestion that the net effect of the ratio between commuting time (*t*) and housing surface consumed (*q*), i.e. ***t/q***, is determinant to the location of groups with different income levels[16] is scientifically unsatisfactory. According to Brueckner, it is based on an ambiguous theory. He reinforce this commentary by citing studies that conclude that this ratio is relatively constant (Wheaton, 1977), and that a change in the means of transportation would lead to a change in the ratio (Leroy and Sonstelie, 1983).

The conclusion presented by Brueckner *et al.* (1999) is that the spatial configuration of groups with different levels of income follows the distribution of amenities in the city. Historical sites, which usually offer with a great variety of quality services, such as theaters, restaurants, and urban attractions, drive individuals with higher income levels to locate near those areas. The opposite occurs in cities where the center has low levels of amenities. The basic assumption used by the authors to develop the model is that the marginal valuation of amenities increases strongly with income (1999, p. 93). Indeed the econometric analysis of this thesis in chapter 6 confirms this observation.

The advances proposed by Fujita and Ogawa (1982), followed by the introduction of amenities, dismiss the theoretical excessive dependence of CDB and introduce the concept of neighborhood and its characteristics to the analysis.

A primary point in their analysis is that they do not establish causality. That is, they do not affirm whether urban amenities constitute the main cause for location preferences of high income individuals, or whether urban amenities are a consequence (1999, p. 94).

Section two of Brueckner's article describes the model using the variable ***a(x)*** which denotes the level of amenities ***a*** at ***x***, that is, the distance to the business district center. Presently, consumer utility includes the quantity of space consumed (*q*) the quantity of the composite good (*e*) and a given level of urban amenities.

The solution of the model indicates that:

Equation 23

$$p\,(x)=-\frac{t}{q(x)}+\frac{u^{a}}{q(x)u^{e}}a'(x)$$

or similarly,

Equation 24

$$p\,(x)=-\frac{t}{q(x)}+\frac{v^{a}\left[y\quad tx,p(x),a(x)\right]}{q(x)}a'(x)$$

Compared to the traditional model, Brueckner, Thisse and Zenou's proposal adds the second term to equation 23. In other words, the slope of the potential function (bid-price rent) *p'(x)* equals the traditional result (the *t/q* ratio), plus the marginal rate of substitution (u^a/u^e) between amenities and the consumption of the composite good in the numerator (*e*). Equation 24 denotes the same slope, but is represented by function v^a, which is the indirect utility function that would be null in the traditional model and which indicates the marginal valuation of amenities after optimal adjustment to housing consumption (Brueckner, J.K., Thisse *et al.*, 1999, p. 96).

Afterwards, the authors propose an analysis for two income groups and conclude that, within this model's framework, "location by income is then linked to a city's idiosyncratic features, and the multiplicity of observed location patterns around the world becomes explicable" (1999, p. 98).

The authors have then demonstrated that the results of the model are flexible enough to predict high income location of households both in the city center and in the suburbs, when modern amenities are relatively inferior compared to the center. Furthermore, the results suggest that multiple equilibria are possible[17].

The expansion of the model makes urban amenities endogenous (because there is no established causality). Amenities are then considered to be dependent upon neighborhood income (*z*). The result of the simplified analysis leads to similar conclusions when the variable is considered exogenous. Thus, they conclude that, "the theory demonstrates that the relative location of different income groups depends on the spatial pattern of exogenous amenities in a city" (1999, p. 105).

Recently, Wheaton (2004) criticized the traditional neoclassical approach, especially its monocentric configuration[18], and suggested a change of emphasis away from the traditional trade-off between accessibility and land to an trade-off between agglomeration and congestion. In the proposed model, mixed land-use is possible so that it can be occupied proportionally to land rent and not simply in total by the highest bid (2004, p. 418). The author's motivation is the empirical certainty that employment is as dispersed as the population. From a worker's perspective, the choice to work near home means both less time and financial resources lost to commuting. From the firm's perspective, according to the traditional neoclassical equilibrium, this allows for a reduction in the salary paid. This residence-service attractiveness scattered throughout the city is theoretically incorporated in the modeling based on actors in chapter 8.

In Wheaton (2004), the introduction of the fractioned use of land is made through the function *F(t)*, which varies from 0 to 1 for commercial use, with *1 – F(t)* for residential use. This allows for mixed density (residential and commercial) in the model through a cumulative function of number of employees *h(t)*:

Equation 25

$$h(t)=\int_0^t 2 \ x\frac{1-F(x)}{q_h}dx$$

where q_h is the quantity of land in the residence (2004, p. 422).

Wheaton's assumptions include the assumption that transport only occurs towards the center, that the outermost location is always a residential one, and that there is a higher dispersion of families relative to firms.

Having designed the model according to the traditional equilibrium in the neoclassical urban economic solution, "for households at a fixed place of residence, all alternative workplaces must yield identical net income. Because rent is fixed by residence, the choice of workplace impacts net income through commuting" (2004, p. 426). The result implies that when there is total dispersion of employment places, transport costs are nearly zero.

1.5 Hedonic prices and urban amenities

One of the approaches of empirical analysis derived from urban economic theory is hedonic price, which was originally mentioned by Griliches (1961) and organized properly by Rosen (1974). Rosen calls the hedonic price function the function that relates the price of a heterogeneous good (residence) to its attributes. Put simply, hedonic prices may be described as a function of a closed package of attributes over which one estimates marginal prices for each feature from analyzing observed values of the heterogenic good and its quantity of attributes (Dale-Johnson, 1982; Kauko, 2002; Hermann and Haddad, 2005; Fávero, Belfiore *et al.*, 2008).

This approach is necessary because there is not an explicit market for urban amenities such as air quality or the presence of green areas and parks. This way, the usage of hedonic price analysis allows the incorporation of positive amenities (externalities) as well as negative amenities (such as violence, pollution or criminality) in the price of housing. Thus, there is a need to estimate its marginal impact through the price of a multidimensional composite, in this case urban residence.

Sheppard (1999, p. 1599) presents two scenarios in which to use hedonic price functions in the real estate market: (a) in the construction of a price index that considers changes in the quality of consumer goods – here, the residence built in a different location has different quality characteristics, and (b) as an informational attribute in the construction of prices of heterogeneous goods.

Though theoretically consistent, practical difficulties often arise, according to Bartik and Smith (1987, p. 1211), in (a) the results interpretation and inference possibilities, given the non-linear functional form of hedonic price functions; (b) the evaluation of changes at the level of amenity attributes associated with the property, i.e., in capturing its variations and varieties; and (c) the assessment of public policies that impact levels of amenities.

The discussion of the empirical work by Bartik and Smith states that the most common functional form is the semi-logarithmic form with the natural log of the price of the residence as a linear function of its characteristics. This implies that a change in those features results in a constant effect on price; thus, every actor can perceive amenities in a homogeneous way. Other authors use box-cox transformations or Taylor-series expansions. Examples of this transformation applied to the Brazilian real estate market can be found in Macedo (1996), Biderman (2001) and Hermann and Haddad (2005).

In some applications, Sheppard comments, the average income by census sectors are used as approximate measures of amenities (1999, p. 1239)[19]. The results of these studies should be interpreted with caution given their positive correlation with omitted variables of amenities. The correct procedure would "summarize the information contained in various measurements in a single variable", for example, with the use of principal component analysis (1999, p. 1240 and footnote 19).

The discussion of urban amenities is relatively new to urban economics and it aims to detail how locational amenities influence individuals and firms' choice and vice-versa. Bartik and Smith (1987) define these amenities as "location specific characteristics with either positive or negative contributions" (p. 1207)[20]. They highlight the idea that there is a multiplicity of tangible dimensions when considering amenities, because they vary regionally (in a location such as New England in the United States or the Dutch Randstad) or intra-urban areas (such as 5th avenue in New York City, The City in London or Amsterdam Zuidoost in The Netherlands), and because it is rather difficult to measure amenities such as charm. In the literature, the usual description of amenities include (a) education expenditures, (b) pollution, (c) violence levels and police preparedness, (d) historical districts, (e) zoning and land-use, and (f) levels of development (1987, pp. 1210-1211). At any rate, whether tangible or not, amenities play a central role if the objective is to describe a realistic model of housing consumption.

1.6. Concluding remarks

The inclusion of urban amenities is made *a posteriori* to the discussion of urban economics theory. If the neighborhood and its characteristics and attributes are considered essential elements of a real estate market, then these factors should be present in the conceptual origin of a theory aiming to explain it.

This chapter provides a brief overview of urban economics, which provides the basis for the comments in the following two chapters.

2 Criticism of urban economics and alternative approaches

2.1 Introduction

This chapter starts with some criticism of Alonso and his followers. We then present a brief description of the city as a working object and we introduce the approach of complex and self-organizing systems. This line of work directs us to other authors (apart from urban economists) who discuss the city and are founded in critical reviews made by Allen (1997), Portugali (2000), Benenson and Torrens (2004) and Batty (2005b). With this literature as a theoretical support, cellular automata (CA) models are introduced including the seminal article by White and Engelen (1993).

Because actors and their local interactions are thought to be fundamental to the understanding of urban dynamics, a description of these features is provided. They are important when considering the city as heterogeneous and irregular at the intra-urban scale of analysis. Therefore, relevant actors and their competition for space, subject to socioeconomic conditions, are described.

2.2 Criticism of urban economics

The criticism of Alonso and urban economics comes from different sources, including mainstream economics, and is summarized in this section with four central points: (a) the focus on transport and location within a perfect equilibrium market, which, in turn, (b) leads to the exclusion of heterogeneous agents and disharmonic urban agglomerations; (c) the lack of explicit spatial influence; and (d) the a-historical, static perspective.

The first criticisms are whether markets spend most of the time near equilibrium rather than at equilibrium (Arthur, 1990; Anas, Arnott *et al.*, 1998; Arthur, 1999; Abramo, 2001; Arthur, 2005) and whether the hypothesis of perfect competition and complete availability of information is too strong. The latter hypothesis is especially appropriate when one considers the characteristics of the residence as a differentiated good (see item 1.1). This is certainly true in the Brazilian case, where medium-sized and larger cities are diverse and heterogeneous, form an idiosyncratic prairie, have imperfect markets where information is limited to neighborhoods (not the whole city), and choice is severely restricted by income constraints. Furthermore, the process of land supply, construction, property supply, and diachronic demand for residences constrains the possibilities for instant market-clearing.

The neoclassic logic of urban analysis also does not include city users whose income is not great enough to own a residence. Unconventional or alternative solutions are disregarded. A full analysis of the real estate market from a developing country perspective cannot fail to address those facts. The possibility for construction or acquisition of a sub-normal property should be added as a fundamental of the market (Plambel, 1987). In other words, the rigidity of the supply – which leads to assumptions that there is land for everybody and that everybody can afford it, excludes the possibility of irregular settlements, slums, or tenement houses, which have been present throughout the twentieth and twenty-first centuries in Brazil (Bolaffi, 1979; Maricato, 1979; Somarriba, Valadares *et al.*, 1984; Santos, Milton, 1994).

A second relevant factor left out from the urban-economic analysis is spatial relationships. The spatial importance is reduced to "the distance to the business center, expressed economically by the dislocation cost" (Brueckner, J., 1987; Abramo, 2001, p. 58). However, flexibility of jobs and activities, alternative transport means, and multicentric configuration (Garreau, 1992), suggest decentralization of points of attraction points in a city that is not encompassed in the idea of a center business district (CBD) and the distance to it (Wheaton, 2004). From an empirical perspective, establishing specifically where the CBD is in a large-scale analysis, so that the distance to it may be measured, is by itself a broad undertaking.

A third criticism of urban economics models derived from Alonso's theory concerns the lack of historical dynamics. "The occupational process … and the occupational mechanisms" (Plambel, 1987, p. 26) are not present in urban economics models (Arthur, 1994; Allen, 1997; Anas, Arnott *et al.*, 1998; Parker, Manson *et al.*, 2003; Brown, Page *et al.*, 2005). When one neglects historical influences and path-dependencies, one is not taking into account the strong trend of spatial rigidities, or that a number of family, employment, school, or friends, inhibits decision factors that have to be coordinated when moving residences.

A fourth criticism of the urban-economic analysis is that land and property markets are taken to be essentially speculative (Plambel, 1987). There are peaks in the property value that are not a consequence of its intrinsic nature, but rather, they are the results of composition of the neighborhood or a result of exogenous market influences. A strong indication of this in the city is when land and buildings are vacant or unoccupied where infrastructure and services are ready and available and, conversely, when low-income class residents are located in distant locations that lack both infrastructure and services (Santiago, 2006)[21]. In other words, people have to commute long distances when there are huge tracts of land developed with infrastructure remaining unoccupied in between. Abramo (2001, p. 126) summarizes the reach of the neoclassical synthesis, adding that a non-conflict space should emerge as a result:

the orthodox speech was successful in interpreting the rationality that rules spatial choices of economic actors, as well as the balance of individual family location; in saying the actors were coordinated by spatial competition and that the process was conceived as a set of simultaneous statements of intentional residential location with the absent owners mediation; saying this competition determined the actor's location as well as the price to pay for the occupied space, a certain composition of expenditures on location … but also the quantity of space consumed in each place in the city, i.e. the configuration of urban densities. Moreover, we highlight that the introduction of different families' income did not indicate denial of the idea of harmonious spatial organization[22]; that the market coordination, in this case, triggered a segmented order, contrary to what one might imagine (spatial conflict), it structures in an efficient way the residential space".

Some of these criticisms are present in recent studies that have an urban neoclassical strand. The distance to the center – a key issue in the equilibrium analysis – proposed by Alonso's model (1964) loses its importance when Fujita and Ogawa (1982) introduce potential density, in which a cumulative agglomeration of jobs is what matters. Wheaton (2004) goes beyond this and spreads the CDB, i.e., by spreading employment and residence throughout the city, and demonstrates that the residence's relative location is essential in the understanding of residential and localization economics.

However, the propositional guidelines still remain distant from one another. Whereas Brueckner (1987) searches for the identification of the "regularities observed" (Glaeser, E. L., 2007), Soja shows that the city conforms to an "uneven intra-urban pattern" (Soja, 2000, p. 234).

The following items establish that a distinct understanding of the city, combined with an emphasis on its spatial relationships and neighborhoods, is also widely described in the literature and can contribute to real estate market analysis in the Brazilian case.

2.3 Defining city

The definitions of city are varied and depend considerably on the background and discipline of the author. In order to study the real estate market within an urban context, a clear understanding of what comprises the city is essential. This is the objective of this section. Therefore, before addressing a differentiated approach relative to the approach started by Alonso, stressing the concepts of urban and city that concur with Brazilian urbanization particularities, which are also present in some "post-modern" metropolis in developed countries (Garreau, 1992; Soja, 2000), is worthwhile.

Many are the concepts of "city", both in the academic and scientific environments, as well as in common sense. Broadly speaking, the main definitions of city include its permanence in time, its reasonable size, and the development of activities that lead to economic and social reproduction.

In the functional corbusiana strand[23], for example, Serra (1987) takes the city as a space in which one can realize four basic individual needs: production, consumption, trade and management.

In the more contemporaneous strand, Monte-Mór (2006, p. 6-8) recognizes the end of the city-country dichotomy (used by functionalists of urban sociology to define city), and based on Henri Lefebvre (1965; 1976; 1991) he describes the city as the expression of a social-spatial division of work. Monte-Mór argues that the city is the space of power, the *locus* of the surplus, and the place of social reproduction, where influence on the country is fulfilled. Consequently the urban landscape expansion, which qualifies the suburbs – the physical expansion of the political city – segments and formalizes the space, but it does not make it homogeneous. On the contrary, the political spaces of control become segmented and segregate physically and spatially. Thus, urban sprawl is different from the political city (*civitas*) that organizes the space.

One other factor that makes the city complex is that each actor's perception or cognitive understanding of the city might be distinct. This alone might impose an intangible evaluation of urban tissue, which results in differentiated valuations of land that otherwise should be considered the same.

Historically, there is also an evolution of the concept of the city. Benevolo (1980) comments that our modern understanding of the city was only consolidated after the French Revolution and the Industrial Revolution. In his book *The History of the City*, Benevolo proposes four segmentations of the city that follow an historic social evolution.

The first segmentation, the Liberal City, is a result of social changes of the period, which may include the following: (a) an increase in population due to a decrease in mortality; (b) an increase in productivity due to technological innovations; (c) a redistribution of the population in the territory; (d) development of mass media; (e) speed in transformation processes; and (f) trends in politics and society (Benevolo, 1980, p. 551-552). The convergence of these factors transform the city into a new object – one that had not been conceptually described – in which the novelty is in the change of scale and functions. The medieval city (*op. cit*. p. 565) was composed by a palace and the church, and no longer bore the mass of new inhabitants. The effect is chaos, which is often described in the novelistic literature[24].

Because the concept of the city itself is new, the Liberal City had yet to divide the public and private spheres.

> "This disorderly and uninhabitable environment … [is] a consequence of the overlapping of many public and private initiatives, which were not regulated or coordinated. Individual freedom, demanded as condition for the development of the industrial economy, reveals itself to be deficient in regulating the transformations of construction and urbanism, produced by the economic development (Benevolo, 1980, p. 567)."

The second denomination of Benevolo (1983) was the Post-Liberal City, which is placed in the post 1848 revolution context and is a consequence of the reorganization of a society that restrains freedom and a society in which public administration acquired its modern configuration. The remodeling of urban form is acknowledged by functional and hygienic pleas (after the cholera epidemic in the beginning of nineteenth century), allowing for the groundwork for the development of urban society in the following expansion period[25]. Its main features are regularity, uniformity, predictability, and monumentality. Camilo Sitte criticizes the planned city that he pictures as lacking naturalness (hence beauty) in his classic work "City planning according to artistic principles" (Sitte, 1889).

The third typology of the city proposed by Benevolo (1983) is the Modern City. The planning of the city detaches itself from bureaucracy and technocrats and becomes the task of architects and artists. They jointly with innovation strands in the painted arts (e.g., impressionism, post-impressionism, and cubism), and in literature and the arts in general, use new technological alternatives in the beginning of the twentieth century, such as steel dissemination, the invention of the lift (along with vertical circulation possibilities), and new construction systems (1983, pp. 615-616), to "reinvent" the city.

At the same time, mass production and assembly line (Ford) concepts are reflected in the production of houses as "living machines" which was proposed by Le Corbusier (1946). The man and his ergonomics and comfort became fundamental concepts in design development, led by the Bauhaus School and Walter Gropius. The starting point when building the city is that of the scale of the man. From that scale on, in an additive way, the surrounding areas are built and then the building itself is considered to be the basic element of the city. Moreover, the city's basic function is to be the place to live and "work, [to] cultivate body and spirit and to circulate" (Benevolo, 1980, p. 637). The categorized zoning of distinct land-use and "garden-cities", coupled with planning and scientific treatment, are also products of this period.

Another way to discuss the evolution of the city is by observing the variations in its primary functions. Monte-Mór (1994) describes the (a) Political City with its original function of Power Headquarters, which is succeeded by the (b) Market City, where the bourgeois incorporate the market into the city, overcome the limitation of medieval city walls and conform to a beginning of structuring of market nets, and finally, (c) the Industrial City, which represents the subordination of the country to the city.

Currently, the discussion of what is the city is still open to various interpretations and is not yet segmented and defined historically.

The definition that we use in this thesis is the heterogeneous city described by Edward Soja. He builds on the concepts discussed by David Harvey (1973; 1992) and Manuel Castells (1999), after his initial pieces on the importance of spatiality in the social sciences (Soja, 1990; 1996), Soja published "*Postmetropolis*" (Soja, 2000) which proposed a review of the role of the city in development and an interpretation of the contemporaneous reality of large urban agglomerations. The city is described as a discontinuous, fragmented, and polycentric entity that do not exhibit Fordism regularities (Soja, 2000, p. 234).

Soja's describes the city as

> "residual, abandoned, [not anymore] easily identified, discrete; increasingly anachronistic, [experiencing] substantial changes, unmoored from its spatial specificity, that we can no longer hope to map, because we can no longer assume we know, porous. It is this turbulent restructuring of territorial identity and rootedness amidst a sea of shifting relations between space, knowledge, and power (Soja, 2000, p. 150)."

In this characterization, Soja refers to large urban agglomerations, especially in metropolitan regions, which cannot be studied as independent parts (municipalities), even though they might be politically-administratively and historically-culturally independent.

Thus, the heterogeneous, fragmented metropolitan urban agglomeration is the emphasis of this study. This is in contrast to the 'homogenous prairie' proposed by Alonso (1964), the 'regularities' observed by Brueckner (1987), or the 'indifference across space' advocated by Glaeser (2007).

This definition has an immediate impact on the appropriate choice of models. On one hand, for the econometric exercise of chapter 6, the definition of the city implies a model that captures (1) the nuances of the city, and hence requires an emphasis on the neighborhood as a local influence, and (2) the heterogeneity of actors who behave differently for the same given attributes. On the other hand, the theoretical approach to cellular modeling, as it is clear in the following section, is on local influences that are able to incorporate the rich disaggregated information of both large urban agglomerations and heterogeneous actors.

2.4 Complex and self-organizing systems: an alternative to urban analysis

As an alternative to urban economics theory and modeling the literature suggests other pathways, including the bottom-up approach discussed in Arthur (1994; 1999; 2005), Holland and Miller (1991), Holland (1992), Epstein and Axtell (1996), Batten (2001) and Rauch (2002) models and arguments. They argue that this approach is adequate in dealing with the disparities and diversities described in the understanding of the city (item 2.3). Torrens, for example, is absolute in his comments.

> "[U]rban models suffer from a lack of realism. Bluntly stated, cities don't really work the way that traditional models would have us believe they do. There is a disparity between models and reality on a behavioral level. In particular, traditional models adopt a reductionist view of urban systems. For the most part, assumptions are made that portray

cities as operating from the top down. This implies dissecting cities into constituent local components from aggregate conditions in order to understand them. In many cases, this is appropriate (planning constraints, large-scale infrastructure improvements, etc.); however, in other instances it is inappropriate (housing demand, commuting, etc.). Many components of urban systems do not work in a top-down manner; on the contrary, aggregate conditions emerge from the bottom-up, from the interaction of large numbers of elements and entities at a local scale (Torrens, 2001, p. 8)."

The focus of this approach is that actors interact locally, with preferences and differentiated budget restrictions, and follow decision rules that have, as an informational horizon, a specific socio-spatial neighborhood. Thus, the systems' typical emergent property (Batten, 2001) can be observed. Spatial relation and neighborhoods are thus the center and origin of the theoretical discussion.

This broader view towards observing constituent elements and their interaction, which aimed to understand the system as a whole, allowed Peter Allen, in *Cities and regions as self-organizing systems* (Allen, 1997), to make the transition between the theoretical discussion of complex systems that come from physics and chemistry to social sciences and regions. Self-organizing systems, for Allen, are:

"collective structures which emerge from the interplay between average behavior, and deviations around this which drive the system through successive instabilities. While a structure is stable, then it certainly can be 'described' by the churning of its connected parts. But, when instability occurs, it can change its structure spontaneously, and *afterwards* will be described by the churning of a new set of parts. The system is therefore both the 'structure' that is observed at some aggregate level and the deviations around this which can change the structure observed (1997, p. 18)."

Thus, Allen suggests that the analysis of emergent structures is more able to fully grasp the phenomenon as a whole and that the description of the parts, which is the traditional top-down view, would not, in most cases, be suitable to describe changes and instabilities.

The presence of path-dependence (Arthur, 1988) occurs when there is a historical dependence, and early events generate positive or negative feedback that, together, define present configurations and urban patterns as a result of earlier events. Batty (2005, p. 29) prefers to define this path-dependence as "qualitatively different trajectories that emerge from the application of particular initial conditions".

In other words, small events are perhaps fortuitous, or random in the beginning, t_0, and lead to certain unique future configurations, t_n, that are dependent upon that initial event. Thus, the implementation of a new cultural center may trigger a renewal process and urban revival, while the establishment of a polluting factory may lead to process of decay and degradation in the subsequent years.

Brown and his colleagues remind us that the presence of dependent trajectories in the spatial context shows that small differences in spatial conditions may lead to largely distinctive patterns in future moments that reveal high sensitivity to the initial configuration (Brown, Page *et al.*, 2005).

From an urban perspective, Allen stresses that the objective of understanding the processes expressed by the proposed models is that they should depict "how urban and regional structure evolve and change as a result of the multiple decisions of inhabitants, which are emade according to individual goals and circumstances, yet each changing the circumstances and goals of the others" (Allen, 1997, p. xiv).

2.4.1 Cellular automata

Cellular automata ("CA") represent one of the main ways to apply the self-organizing systems approach to urban models of land-use and transport. Specifically, for urban matters, Batty (1998), Torrens (2001), Pines and Thisse (2001), Capello (2002), Longley and Batty (2003), and Glaeser, Gyourko and Saks (2005) acknowledge CA as a promising instrument to deal with local interactions and social neighborhoods, spatial irreversibilities, cumulative processes, and a variety of behaviors and urban space uses. Specifically, Brown (2005) and Batty (2005a) highlight the usage of CA models in studying processes, as opposed to forecasting.

The concept was primarily developed by the mathematicians S. Ulam and J. von Neumann, in the 1940s and applied by John Conway in 1960 in his famous "Game of Life" (Berlekamp, Conway *et al.*, 2004).

The game of life can be briefly described. There is a number ***n*** cells that compose a matrix of size ***i*** x ***j***. Each cell is in one of two states *n S(1,2)*, representing death or life. The neighborhood is determined by the eight cells immediately adjacent (Moore neighborhood). Two transition rules apply: (a) the cell stays in state 1 (life) if two or three of its immediate neighbors are also in state 1, and (b) the cell in state 2 (death) transforms into state 1 if there are three neighbors in state 1. These simple rules, which were later altered in many different ways, allow for complex and unexpected results, or as Batty says, an "emergence [of a] process whereby 'unanticipated consequences arise from well-defined rules" (2005, p. 51).

The use of this approach led Thomas Schelling (1978, *apud* Batty, 2005) to contend that the emergent property, or the one that is observed, may many times differ from the behaviour that actually rules a system. In his classic example, Schelling establishes that if only one-third of the population shows a segregationist behavior, the resulting urban system presents itself with a higher than one-third degree of tangible segregation[26].

From the 1990s onwards, with technological advances in computing and remote sensing, interest in CA was renewed, which empowered theoretical advances and new applications (Portugali, 2000; Benenson and Torrens, 2004; Batty, M., 2005b). Portugali and Benenson, for instance, built models of urban CA whose fundamental aim was to broaden research of urban segregation that was already validated by Thomas Schelling. Batty (2005), in turn, emphasized the morphological analysis of the urban dynamic.

CA studies usually have in common (a) a certain lattice, grid, or matrix containing cells, (b) the definition of a cell neighborhood, (c) different possible states, (d) transition rules between states, and (e) the imposition of specific restrictions on specific states or cells.

The most general definition to CA for Batty is that: "[C]ellular automata are computable objects existing in time and space whose characteristics, usually called states, change discretely and uniformly as a function of the states of neighboring objects those that are in their immediate vicinity" (2005, p. 67).

More precisely, Portugali interpret CA as:

> "A standard two-dimensional cellular automata (CA) model is a lattice of cells where each individual cell can be in one of several possible states (empty, occupied, etc) and have one out of several possible properties (developed, underdeveloped, poor, rich, and so on). The dynamics of the model are generated by an iterative process in which in every iteration the state of each cell is determined anew by some transformation rule(s). The rules are local and they refer to the relations between the cell and its immediate neighbors. The name of the game is to see how, what and in what circumstances, local

interrelations and interactions between cells entail global structures, behaviors and properties of the system as a whole (2000, p. 66)."

Roger White defines detailed CA as "very simple dynamical spatial systems in which the state of each cell in an array depends on the previous state of the cells within a neighborhood of the cell, according to a set of transition rules" (White, Engelen *et al.*, 1997, p. 323).

The introduction of the transition process between states incorporates a temporal dynamic in the analysis plus the influence of preview effects in present results, i.e., the conception of path-dependence. Moreover, transition rules, when based in the neighborhood, allow the introduction of local effects that are spatially specific.

A key difference in urban CA models in relation to more general models derived from Conway is that they are constrained. This means that, at each interaction, the total number of cells that are allocated to a certain land-use is exogenously determined. Thus, city growth itself is not being modeled. Rather, its land-use intra-urban allocation is (White, Engelen *et al.*, 1997).

2.4.2 Agents, actors and cells

As previously stated, this thesis conceives of the city as a multidimensional object that is the result of the individual actions of various agents (or actors, as preferred by the urban planning literature) in space. Specifically, in terms of the CA model, both Dietzel and Clarke (2006) and Fernandez *et al.* (2005) demonstrate the importance of using distinct agent classes when modeling within an urban environment.

To make definitions clear, however, it is worth noting the difference between agents, who are common in agent-based models (ABMs), actors, who are referred to as the active parties of urban intervention, and cells, which have certain properties or states, that are usually used in CA models.

The literature on ABMs usually labels agents as mobile entities who move spatially on a *grid* or *lattice*. Agent *a* can be at coordinates *x,y* (0,0) at t_0 and move to *x,y* (0,1) at t_1. This specification is suitable for dealing with problems with animals, pedestrians, or cars, for instance.

Portugali (2000) describes agents as "free" in the sense that they "act intentionally on the basis of their personal history, memory, and perception (i.e. cognitive map) of the city" (2000, p.4).

Allen (1997) portrays agents as individuals who control, in terms of decisions, the energy flow. They turn connections on and off, "react, learn, and change according to their individual experience".

Michael Batty (2005b) describes agents as reactive, autonomous, goal-oriented, temporally continuous, communicative, capable of learning, mobile, and flexible, with an authentic personality (2005b, p. 212).

Urban actors can also be easily described in a manner similar to these agents. They are featured as belonging to different typologies, active in urban space conformation process, and not as individuals "who react to a situation", as Celso Furtado (1980, p. 35) states. Urban actors usually include developers, residents, institutions, and economic actors, such as industry and commercial businesses (Plambel, 1987).

The main difference between agents and actors on the one hand, and cells on the other, is that cells are fixed. Cells maintain the same coordinates throughout the process. Nevertheless, they undergo state transformations, as described previously.

In this thesis, the analysis at the proposed intra-urban scale allows for the assumption that one actor occupies the entire surface of each cell. That is, cells are referred to as actors because they have the characteristics and attributes of those who structure the urban space. In the CA process, when an actor is allocated to a certain cell according to transition rules, the cell might be considered to have the attributes and characteristics of that actor. In this sense, one cannot say that "a cell 'acts", but that one actor within the cell is the one who acts. Hereinafter, actors are defined as the individuals who act in the urban environment, but who are spatially fixed, and represent urban agents who are active in molding the city process. To avoid confusion with ABMs literature, after this section, we use only the word actors.

These main actors to select the most relevant ones, broadly follow Plambel's definitions (1987):

a) segmented residential users, who account for social heterogeneity, in at least three sub-divisions of income: high, average, and low;
b) real estate intermediaries or brokers who manage estates that are "in the market";
c) service and commerce units, which provide jobs and services and might be described as local or regional, depending on their reach and influence on the city;
d) the industrial sector, which is also organized on at least two scales, (a large and a small scale), with the large scale provided exogenously and the small scale being more scattered and with less negative impact on the city;
e) other actors who are regarded to be important in the establishment of urban structure are large university campuses or schools, parks and conservational units, airports, train and bus stations, city halls, and other "institutional" buildings.

2.4.3 Transition rules

There are many possible rules that influence the transition of cellular automata states, ranging from simple rules as in the Game of Life, to more complex rules such as the rules proposed in Portugali (2000) and Batty's (2005b) models. Given the researcher or planner's intention, the conceptual and methodological theoretical framework allows for sufficient combinations to comprehensively describe the dynamics of the object of study.

Batty (2005b) presents the following types of transition rules: (a) standard, which is based on a specific configuration of neighboring cells; (b) accountant, which effectively counts the number of neighbors in a given state; (c) statistic, which obeys some statistical formula, such as the state or neighbor features average; and (d) suffrage, in which there is a threshold of neighbors in a given state so that the transition can occur.

These transition rules are further detailed in White and Engelen's model in the next section (and in chapter 3). In short, they add, to the accountant general type described by Batty (2005b), a decaying effect given by distance and a matrix of weighted parameters that mediate the influence of pairs of cells at similar or different states.

2.4.4 White and Engelen's model

Concerning the application of CA to urban, regional, and transport analysis, one of the seminal articles of reference is that by White and Engelen (1993), which served as a conceptual basis to a series of subsequent papers (Engelen, White *et al.*, 1995; White, Engelen *et al.*, 1997; Barredo and

Demicheli, 2003; Barredo, Kasanko *et al.*, 2003; Liu and Phinn, 2003; Parker, Manson *et al.*, 2003; De Nijs, De Niet *et al.*, 2004; Verburg and Veldkamp, 2004; Lau and Kam, 2005).

In their proposed model, White and Engelen (1993) delineate the evolution of urban land in time by assuming a high spatial resolution in which the main interaction happens among actors (land-use) in a given neighborhood. There are different land-use types (empty, residential, industrial, and commercial), and the cells are converted from one use to another at each time step according to transformation rules (*op. cit.*). The growth rate (conversion) of cells is exogenously determined and the neighbourhood considered is a constant radius of 6 cells. The representation of actor behavior that determines the land-use conversion is accomplished with the variable "transition potential". Therefore, given the exogenous demand of a certain land-use type (e.g. additional residential areas needed to accommodate population growth) in a certain time period, one can identify the cells with the highest total potential make the conversion. The total potential is calculated as:

Equation 26

$$Tp_{ij} = (S_I + \Sigma_{h,k,d}\, m_{kd} I_{hd}),$$

where Tp_{ij} is the transition potential of the state "***i***" to state "***j***"; M_{kd} is the weight parameter applied to cell in state "***k***" in distance "***d***"; ***h*** is the cell index within a given distance; I_{hd} equals 1 if the state of cell is $h = k$; otherwise, it equals zero; ***S*** is a stochastic disturbance term[27].

With this fairly simple model[28], White and Engelen (1993) simulate the land-use dynamics that structure urban environments. Essentially, the model and other CA models have been used to achieve an understanding of a process that is basically dynamic and spatial.

In the models developed at the Research Institute of Knowledge Systems (RIKS) (White, Engelen *et al.*, 1997; Engelen, White *et al.*, 2004; Engelen, Hagen-Zanker *et al.*, 2005; Riks, 2005; Delden, Luja *et al.*, 2007; Wickramasuriya, 2007), which are extensions of the seminal work by White and Engelen, the transition potential (***Tp***) is not only determined by the neighborhood effect (*N*), but also by a wider range of factors. The suitability (*S*) is added to make sure that land-use conversions comply with physical conditions of the terrain. ***S*** might vary

> "for each land-use state. Thus, a site might be suitable for industrial occupation, but not for recreational activities, for instance. Also a low suitability value may prevent a transition to a certain state, even if other factors are high. Accessibility (*A*) is defined as access to transport infrastructure, such as arterial streets, or transfer points, such as railway stations. Zoning (*Z*) is a layer of imposed restrictions by zoning or other legislation that prohibits occupation of specific land-use types in certain locations. Finally, as in White and Engelen's (1993) model, a stochastic disturbance term (ε) is incorporated that represents unobserved effects and unpredictability:"

Equation 27

$$Tp = f(N, S, A, Z, \varepsilon).$$

Information related to suitability (*S*), accessibility (*A*) and zoning (*Z*) are usually static unless there is public intervention, though the neighborhood formation (*N*) is updated at each time-step.

The dynamics used in the model are described by Delden and Engelen (2006, p.2): For each land-use function, a set of spatial interaction rules determines the degree to which it is attracted to, or repelled by, the other functions present in its surroundings; If the attractiveness is high enough, the function will try to occupy the location, if not, it will look for more attractive places. New activities and land-uses invading a neighborhood over time will thus change its attractiveness for activities already present.

2.5 Other cellular automata models

Other authors, such as Page (1999), Bell (2000), and Behrens (2005), follow similar methodologies and develop models in which there is a greater behavioral diversity. Portugali (2000) uses a similar logic, but focuses on modeling segregation and migration processes.

Similarly, in Batty's (2005) models the [transition] potential is determined by three factors: (a) the positive feedback effect related to existing potential, (b) the interaction effect between the neighborhoods, and (c) a stochastic disturbance.

Another model of land-use simulation available is UrbanSim (Waddell, 2002; Waddell and Ufarsson, 2003; Waddell, Ulfarsson *et al.*, 2007). The authors' main goal was to simulate the effects of transportation systems on land-use and the subsequent land-use effects on transportation systems. In doing so, they aimed to reduce biases in transportation models that do not account for those feedback effects. To model the entire process, their framework includes a number of sub-models, such as land development, land price, accessibility, jobs and household location. In terms of spatial approaches, the authors subdivide the study area into 150 by 150 meters grid cells. Each is a summation of internal agents that account for individual households, jobs and firms. The database of the simulation (Waddell, Ulfarsson *et al.*, 2007) is vast and includes tax assessor office files, unemployment insurance information, commercial sources, and census data. These are all integrated within a GIS interface. Specifically, the land price model follows the consolidated urban economics literature (Alonso, 1964; Rosen, 1974; Dipasquale and Wheaton, 1996). As a result, the authors regress the aggregated sale price value of the cell on housing characteristics, accessibility and 'environment', which is given by grid cell information (Waddell, Ulfarsson *et al.*, 2007, p. 397). As such, they perform a "cross-sectional estimation"[29], in which the values of independent variables are updated at each year by the other sub-models of the simulation.

There are some differences between the UrbanSim model and the model of White and Engelen (1993), which is extended in chapter 3, and applied in chapter 8.

The essential difference is that the focus of our study is on the differentiated impacts of price on the location of socioeconomically distinct residential actors.

That the application of our model is targeted toward a developing country's perspective imposes two further differences in relation to the UrbanSim framework. The notion that transportation systems impact land-use and vice-versa has to be seen under different focus. Although we do not dispute that this is true, the dynamics might be rather different in a country such as Brazil. The density cities is much higher than in Utah (USA) where UrbanSim is applied (Waddell, Ulfarsson *et al.*, 2007). Apartment-buildings are the rule rather than exception, and the 'suburban detached houses' is not the standard housing type. Due to historical financial constraints on the country's budget, transportation systems are implemented a great deal of time after they were needed and, therefore, are built after occupation is consolidated rather than as an influential factor in urbanization patterns[30].

Yet another model developed for modeling the evolution of land-use patterns is SLEUTH (Silva and Clarke, 2002; Pontius, Boersma *et al.*, 2008), which is an acronym for the incorporated features of the model, i.e., Slope, Land, Excluded areas, Urbanization, Transportation, and Hillshade. This model is an example of how computational power ('brute force method') is applied to derive statistically significant trends, which are then used for prediction. According to the description provided by Clarke (1997, p. 252), the behavior rules of the model "involve selecting a location at random, investigating the spatial properties of the neighboring cells and urbanizing the cell or not, depending on a set of probabilities".

2.6 Concluding remarks

This chapter presented criticisms of urban economics, but also presented alternative views – fragmented, rather than homogeneous – to define the city. The characteristics of cellular automata, which include the neighborhood effect and positive feedback mechanisms in a dynamic representation of the urbanization process, can complement the Urban Economics approach. The fact that models are dynamic allows the modeler to focus on understanding the processes underlying the emergence of the city. Furthermore, given its theoretical framework aimed at explaining local influences as key to urban development, empirical analysis can provide insights into how interactions among differentiated actors actually take place on a local level.

The model proposed in chapter 3 emphasizes understanding the urban development process. This approach is distinct and, therefore, difficult to compare with cross-sectional analysis that establishes correlations for a given *status quo*, as in the case of the urban-economic analysis presented in chapter 6.

However, the core understanding of economic factors and influences contribute to the development of CA models. Positive feedback, for instance, is usually referred to in the economics literature as economies of agglomerations and goes back to Alfred Marshall (1890). From a different perspective, they are also present in Jane Jacobs' work (1970) and it is sometimes referred to as centripetal and centrifugal forces in an urban geography. However, the CA models presented in this chapter do not include negative feedback or the effect of diseconomies of agglomeration explicitly. Although Delden and Engelen (2006, p.2) mention that the rules determine whether a land-use function is attracted or repelled, it does not leave room for both to happen for the same pair of land-use relations.

In the next chapter we argue that one single event, such as the installation of a new industrial facility, brings both positive and negative effects simultaneously, rather than either attraction or repulsion. In the example of a new industry, for instance, firms from the same chain of production might feel a tendency to locate nearby. The competition generated by the new attractiveness of the site, in turn, leads to higher prices for the terrain, which in itself, is a repelling factor. Therefore, both these unfolding effects should be included in any model.

3 An extension to White and Engelen's model

This chapter addresses both preceding chapters from a Brazilian perspective and proposes an extension and adaptation of the model presented in 2.4.4. A general introduction of the extension is followed by a detailed description of the model in which neighborhood rules and accessibility influences are described, along with further modifications. Finally, the system used in the implementation of the model and the description of calibration and validation procedures are presented.

3.1 Theoretical compatibility and the Brazilian case

This section examines the two theoretical approaches discussed in the previous section to determine which is most suitable for the Brazilian context.

The theoretical discussion of urban economics often takes place within the North-American context and follows typical characteristics of that country, such as well-established central and sub-central areas (whether cities of older urbanization processes like New York, Boston, and Chicago or recently developed metropolitan areas such as Dallas-Fort Worth and Houston, Texas), and suburbanization, i.e., residential areas far from the center. In Europe, despite differences in urbanization when compared to the United States (Anas, Arnott *et al.*, 1998, p. 1428-1430) a similar, though less pronounced, hierarchy in urban functions can be observed.

The Brazilian case is different for various reasons[31]: (a) accelerated, rapid, intense and recent urbanization processes have taken place since the 1950s (Santos, Milton, 1994, p. 29); (b) the urbanization process has resulted in dense and numerous metropolitan slums and associated social problems (Maricato, 1979; Somarriba, Valadares *et al.*, 1984); (c) people from different social classes do not have similar access to either the public or private transport systems (Gouvêa, 2005); (d) enormous disparities of income are observed (Barros, Foguel *et al.*, 2007; Soares, 2008); and (e) intervention by public policy is scarce (Gouvêa, 2005).

Finally, an important factor that differentiates the Brazilian perspective in relation to developed countries is the lack of a consistent and sufficient social housing program or housing subsidy schemes. The construction and financing system implemented in the 1960s collapsed in the 1970s due to internal inconsistencies (Bolaffi, 1979). Since then, government aid has not been sufficient (Santos, C., 1999) and sub-normal housing has prevailed for a large part of the population.

Even though a number of Brazilian authors, such as Macedo (1996), Biderman (2001), Hermann and Haddad (2005) and Fávero (2005; 2008), apply urban economic theory to describe real estate market processes, we believe that the analysis can be enriched when it explicitly considers the differences in the Brazilian urbanization process explained above (see also item 2.2).

On one hand, the theoretical neoclassic approach does not seem to fully describe the observed city described in 2.3. In addition to the criticisms made by scientists within mainstream economics itself, it is important to add (a) the fundamental aim of searching for regularities (Brueckner, J., 1987) when the observed city is highly heterogeneous (Soja, 2000); (b) the lack of attention to how space and the neighborhood influences the real estate market; (c) the imposition of equilibrium and market-clearing; and (d) the static nature of the analysis. These facts together suggest that the urban economic theoretical scope, without further adaptation, is not fully sufficient to handle the urban differences observed in Brazilian cities.

On the other hand, the bottom-up approach brings some advantages in terms of analysis of the dynamic urban structure defined by Anas and colleagues (1998, p. 1431). It is also better capable in explaining how polycentric urban structure (the so-called "*edge cities*" (Garreau, 1992)) can emerge. Due to a more flexible conceptualization of how urban growth takes place, the bottom-up approach can describe centralized urban patterns resulting from economies of agglomeration (Anas, Arnott *et al.*, 1998, p. 1459), but also typical urban discontinuities and irregularities (*op. cit.*, p. 1427)[32]. It considers the importance of history and stochastic factors (path-dependence) (*op. cit.* p. 1459), but also of the temporal organization of the city formation and the "first-mover advantage" concept (*op. cit.* p. 1428).

The bottom-up approach encompasses concepts of "chaos, complexity, fractal, [and] dissipative structures of self-organization … out of equilibrium" (Anas, Arnott *et al.*, 1998, p. 1451). In addition it assumes that the presence of some urban functions may stimulate the emergence or growth of similar or other urban functions, and that the strength of this influence depends on distance. This implies that the city is defined as a self-organizing system in which the neighborhood concept takes a central position.

Nevertheless, a criticism of this approach is that the economic influence is not explicitly part of the analysis. Anas and colleagues call the resulting models *noneconomic dynamic models* (1998, p. 1451). Batty corroborates this by saying that the approach is more physical and morphological than really economic-like (2005, p. 458). Hence, a crucial question is the extent to which pricing mechanisms are in reality determinant of urban development and what this implies for the applicability of the bottom-up approach.

> "Anas asks whether: are spatial interactions [are] mediated through prices that are more important than unpriced spatial influences and externalities. Since unpriced externalities probably play a dominant role in shaping urban spatial structure, the challenge posed by the noneconomic models cannot be easily dismissed (Anas, Arnott *et al.*, 1998, p. 1455)."

Furthermore, we believe that the different behavior observed among diverse social groups should be included as a key element of urban development processes. Social heterogeneity also enables a proper treatment of differentiated price-influence.

Hence, what is a proper approach to modeling urban development? Is it the approach that is too stylized that deviates from urban reality and does not consider the essence of the influence of neighborhood, proximity, and spatiality? Or, is it the approach that focuses almost exclusively on spatial relations, path-dependence, and the resulting morphology, but ignores the existence of prices as an influence on urban development?

A comprehensive approach would be an approach that has an origin in the essence of the bottom-up approach, with all the listed advantages, but that, beyond the result, could distinguish the economic influences more explicitly in the process. This alternative is proposed in the next section.

3.2 The extended bottom-up model, including social heterogeneity and negative feedback

The intended extension to the standard CA model, which is based on local influence *per se*, should account for the shortcomings mentioned above, and should therefore include: (a) socioeconomic heterogeneity, and (b) negative feedback of real estate prices, which is mediated by the spending affordability of each actor.

These factors are not explicit in White and Engelen's model (1993), in which only locational quality, as defined by neighboring land-uses, determine land-use conversions. Lau and Kam (2005) acknowledge the need for the inclusion of negative effects. However, they implement this with what they call a "retarding effect", which is operationalized as inertial (gravity) forces. They consider residential property value as a driving force. Heterogeneity of agents is urged by Fernandez (2005, p. 799), who believes that within agent-based models should reflect both the "heterogeneous nature of preferences" of agents and "diversity in the population". Heterogeneity of agents is also used in a CA non-urban context by Dietzel (2006).

Because of the changes proposed, the model not only simulates the process of occupation of land by various actors, but also urban land value.

Furthermore, the extended model strengthens the theoretical foundations of the original model because it incorporates concepts of classic urban economic theory, such as centripetal and centrifugal forces (Colby, 1933; Krugman, 1996), that influence location behavior. In other words, agglomerative and disagglomerative forces can be simply described as, on one hand, the attractiveness of a location in the center (or near any point of interest) or the centripetal force, weighted, on the other hand, by the diseconomies resulting from increasing prices due to competition for space (the centrifugal force).

Including both effects in the basic CA model requires that the conversion of land from one function to another be determined by (a) the presence of other urban functions in the neighborhood, weighted by distance (the neighborhood effect) and (b) the price (land value) of the location. At the same time, both the neighborhood effect and prices will change due to changes in the urban structure from these conversions. This suggests an iterative process, in which land-use conversions on the one hand and neighborhood effects and prices on the other hand, are mutually influential.

In a dynamic sense, actor locations at time *t* determine the land-use configuration and the associated land prices. Actors may generate a positive agglomeration effect for a specific site. However, the agglomeration advantage leads to more competition for the site, which leads to higher land prices (as well as congestion, pollution, etc.), thus generating disagglomerative effects. Both effects are taken in consideration at the next time-step in which there is allocation of new-entry actors. Competition among actors for locations following the transition rules (detailed below) leads to a new configuration of actors at *t+1*, which, in turn, establishes new land prices (Figure 3-1).

An example is provided to clarify these processes.

Suppose a new high-income residential actor is located at a given site. Its location makes nearby locations more attractive for residential, commercial, or service actors, but more strongly so for other high-income residential actors. At the same time, the occupation of a site by this kind of actor causes a price increase that is stronger than in the case of a low-income residential area, for instance. Although all actors have a preference to locate near the high-income residential actor,

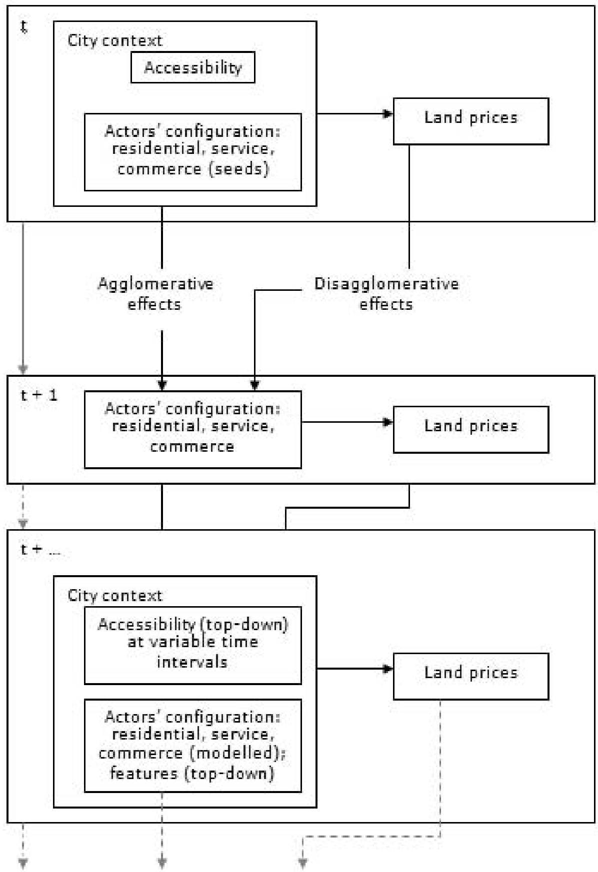


Figure 3.1 Illustration of processes

the price increase may prevent some actors with lower spending capacities (such as low-income actors) from actually doing so.

Thus, prices reflect the attractiveness of locations (expressed in neighborhood effects and accessibility) as the outcome of competition between actors for these locations. This implies that prices are related to the locational characteristics of the neighborhood and its accessibility. In the proposed model extension, price will therefore be implemented as such.

This way of calculating prices is similar to the approach used by Waddell and Ulfarsson (2003, p. 7; 2007, p. 396), in which they regress the value of the land based on structural characteristics of the estate, its neighbors and accessibility factors. Furthermore, both Waddell (2007) and Dipasquale (1996) indicate that real estate is a summation of two factors: the estates' unique attributes and its locational influence determined by the urban environment.

In the proposed extended model, the influence of price on transition potential is weighted by a parameter [β ε (0, 1)] that expresses the impact on prices of factors not included in the model, such as the estate's attributes and characteristics, its surface area and its quality.

In addition, the impact of each actor's location and the neighborhood on land prices is given by μ_s. The impact of price on transition potential is dependent upon the spending power of each class of actors, which is given by τs (0, 1). Both parameters are necessary because of the steps considered in the model. In Figure 3-1, μ_s is applied when the city context configures land prices. τs influences the location of new comers differently, depending on their spending power.

One other important point to highlight (see also Waddell (1996) and Dipasquale (2007)) is that we will ignore cyclic fluctuations or long-term growth (e.g. due to inflation or demographic developments), but we will instead assume that relative prices remain relatively stable.

Generally speaking, equation 27 becomes:

Equation 28

$$Tp = f(N, S, A, Z, P, \varepsilon),$$

in which (***P***) is the price of real estate. However, as zoning and suitability are not present in the model (see item 3.3.4, p. 68), equation 28 can be written as:

Equation 29

$$Tp = f(N, A, P, \varepsilon),$$

More specifically (equation 30), the total potential is now expressed as the result of agglomerative effects[33] minus disagglomerative effects[34]. Agglomerative effects include a random effect and the influences of neighborhood and total accessibility, whereas disagglomerative effects consist of the influence of location on price, weighted by spending power.

Formally, we have:

Equation 30

$$^{t}Tp_{s,c} = [(1+e)^{t}N_{s,c} \cdot {}^{t}TA_{s,c}] - [\beta \cdot e' + (1 - \beta)^{t} P_{c} \cdot \tau_{s}],$$

in which, $^{t}TP_{s,c}$ is the potential transition value of each actor (s), for each cell (c) at each time step (t); $^{t}N_{s,c}$ is the neighborhood effect; and $^{t}TA_{s,c}$ (0,1) is the total accessibility value, both at each cell (c), at a certain time (t) and for each actor (s), which is further detailed in item 3.3; $^{t}P_{c}$ is the price of cell (c) at time (t), influenced by the neighborhood; β and $_s$ are parameters; and **e** and **e'** are random values.

Additionally price (***P***) is calculated endogenously as:

Equation 31

$$^{t}P_{c} = f\,(^{t-1}N_{s,c}, , ^{t-1}A_{s,c}, e),$$ or

Equation 32

$$^{t}P_{c} = \Sigma_{C' D(c)\ s} F\,(.)$$

where $^{t}P_{c}$ is given by the summation for a given radius ($D_{(c)}$) of the actual configuration of actors in the previous time step, weighted by a parameter μ, which varies for each actor (s). This is necessary because each actor impacts land prices differently. Note that each actor has a full map of transition potentials. Although price is the same for all actors, different μ values imply different transition potentials.

Occasionally, there are changes in accessibility as a result of policy implementation and funding availability, which is therefore top-down. There are also actors who are implemented exogenously – called features – in specific years given historical information.

For each time step t, the above model will predict the occupation of each cell (i.e. which actor is located in the cell) as well as the price of the cell.

3.3 Detailing the model

3.3.1 The neighborhood effect (N)

The neighborhood effect is essential in models of cellular automata because it provides the tools to actually implement local spatial relations.

Note that the concept of neighborhood used in CA models is slightly different from the concept applied in chapters 5 and 6. Whereas the emphasis in the cross-sectional analysis is on 'identity' and 'cognitive perception', as proposed by Galster (2001), CA models emphasize neighborhood in the sense of homogeneous areas, posed by Megbolugbe (1996). In the scale analysis of Guo (2007), neighborhoods are conceived as 'the immediate vicinity of one's house' (see chapter 5), so that identity is treated implicitly rather than explicitly.

Combining these notions with the mechanism of CA models, the behavioral interpretation of CA models would be that households from different socioeconomic groups perceive the characteristics of different neighborhoods and the extent to which these match with their preferences, including the price. This perception eventually determines where particular socioeconomic groups decide to reside.

This raises the question of succession of actors in space and spatial fixity. On one hand, succession has been described by Burguess (1925, p. 50) as "the main fact of expansion, namely, the tendency of each inner zone to extend its area by the invasion of the next outer zone". On the other hand, this effect is counter-balanced by spatial fixity of actors, also referred to as inertia, which is the tendency of an early occupant of a certain site to remain at the same site despite influences of all sorts on the contrary[35]. Lau and Kam (2005) refer to inertia as a 'gravity' effect which is modeled locally and separately ('single-pass') from the rest of their effects (attributes/accessibility and heterogeneity [of neighborhood], multi-pass). In our model, inertia is modeled by the influence parameter at distance 0 – that is, the cell itself, as detailed in following paragraphs.

Thus, the neighborhood effect represents the impact of the perceived surroundings on the attractiveness of a cell for a certain use. The neighborhood effect is given by ${}^{t}N_{s,c}$. More specifically[36]:

Equation 33

$${}^{t}N_{s,c} = \Sigma_{c'D(c)} N_{s(c),s'(c'),d(c,c')}$$

where $D(c)$ is the neighborhood of cell (c); $(d(c,c'))$ is the Euclidean distance between cell c and c'; $n_{s(c),s'(c'),d(c,c')}$ is the influence parameter (n) that expresses the strength of the influence of a cell of a given actor s on another actor s' for each distance $d(c,c')$.

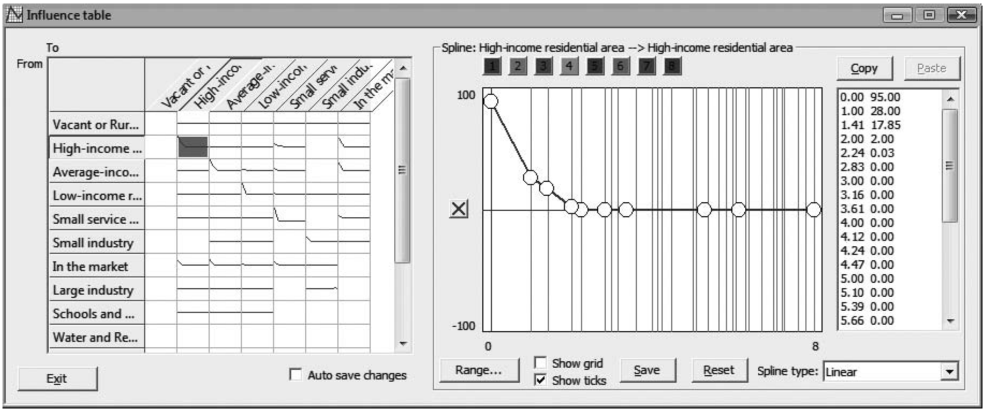


Figure 3.2 Illustration of influence table effects

These parameters are determined iteratively and manually according to the calibration procedures discussed in item 3.5. They enter in the model via the influence table that describes, for each pair of functions how the impact on the transition potential depends on distance. See the example in Figure 3-2.

In Figure 3-2, the highlighted square on the left-hand side represents the influence of high-income residential areas on high-income residential areas. The graph and the specification to the right indicate that at distance 0, the impact is 95. This represents the impact of the cell on itself, which is, in fact, the inertia effect. The high value implies that a high-income residential area has a high probability of remaining as high-income area. The influence on distance 1 is 28, suggesting that cells close to high-income residential areas are also likely to convert to (or remain) high-income residential areas. The next number for distance, 1.41, accounts for diagonal neighbors for whom, in the example, the effect is lower at 17.85. The same effect is depicted in a two-dimensional configuration in Figure 3-3.

The size of the neighborhood is also tested iteratively and it may vary from 0 – no influence – to a maximum distance of 8. In most cases, a small neighborhood with a distance of 3 suffices to represent the impact (see also Hagoort (2006))

3.3.2 Accessibility (A)

Accessibility is a factor that brings the infrastructure of transport and its evolution into the framework of the system. It accounts for proximity to the nodes and links of the transport network and their impact on different types of actors. For every feature of accessibility and each of the modeled function, the modeler provides two inputs: (a) a distance decay function and (b) a relative importance for a certain type of network or node to each specific state ($W_{y,s}$). This is because different actors might value transport networks differently (such as bus systems or metro lines which are more important for low-income residential actors and arterial roads for private car users).

Formally, the equation for accessibility[37] can be expressed as:

Equation 34

$$ {}^{t}A_{c,y,s} = \frac{a_{y,s}}{{}^{t}d_{c,y} + a_{y,s}} \ , $$

where ${}^{t}A_{c,y,s}$ is the accessibility of cell c in relation to a certain type of node or transport link y, for an specific actor s at time t; $a_{y,s,s}$ the accessibility distance decay factor for actor s and type of node or transport link y, such as the importance of highways to small industry activity; and $d_{c,y,t}$ is the distance d of a certain cell c to a transport node or link y at time t.

The total accessibility value, which takes into consideration distance to all nodes and link types ($s \in S_y$)-, is given by ${}^{t}TA_{s,c}$, as

Equation 35

$$ {}^{t}TA_{s,c} = \frac{1 - \left[\prod_{s \in S_y} \left(1 - w_{y,s} \, . \, {}^{t}A_{c,y,s}\right) \right]}{1 - \left[\prod_{s \in S_y} \left(1 - w_{y,s}\right) \right]} \ . $$

In practice, this means that every cell has, for every actor, a total accessibility value that varies between 0 and 1.

3.3.3 Constraints and competition between actors

The model used in this study assumes, in accord with White and Engelen (1997), that what is being simulated is the spatial allocation of urban actors rather than their growth. Thus, the number of actors (cells) of each type is exogenously provided for each time-step of the simulation as the outcome of municipal growth (see item 8.3.2, for details on exogenous demand for the application of the model to the Greater Area of Belo Horizonte (RMBH) case).

The allocation of user types to cells and the transition of cells into a new state (actor) occur as follows:

(1) In every time-period of the simulation, the transition potential of all cells for all actor types is calculated, according to equation 30. Thus, at every step there are as many transition potential maps as there are actor types being modeled.
(2) The cell with the highest potential value among all transition potential values, for all actors, is allocated to the corresponding actor.
(3) Once one transition occurs, the demand of that actor for that time period is reduced accordingly.
(4) Then, the second highest potential cell among all cells is allocated for the correspondingly actor.
(5) Then, this process is repeated successively until all demands for a certain time-period are met.

This allocation procedure implies that actor types compete for locations in a way that that is governed by the price mechanism and spending power of the actors. If, ignoring land prices, a location is equally attractive to two actor types, differences in price sensitivity (as expressed by parameter τ_S) will cause a higher preference for the location of the actor that is least price sensitive. Consequently, in such cases the actor with the largest spending power will occupy the location.

As an illustration, consider the transition potential of a low-income residential area. Given its preference rules, the area enjoys being allocated near commerce and services and average-income residential actors, for example. Thus, cells near these other actors have a high transition potential to low-income residential actors. However, average-income residential actors might have similar preferences and present a transition potential just as high for the same cell. Among all 'interested' for a specific cell, the actor with the highest value and the cell for which there is still demand to be fulfilled makes the transition and "occupies" the referred cell. Most likely, the higher price sensitivity of the low-income residential actor will result in a lower transition potential for this actor when compared to an average-income actor, so that the average-income actor will occupy the land.

3.3.4 Further adaptation: zoning and suitability

The model described in this study further differs from the original model presented by White and Engelen (1993; 1997) in item 2.4.4 in that it is applied without *suitability* and *zoning*. The reason for doing this is that including *suitability* is appropriate where geographical conditions are important for the locational choices that actors make. In the present context of Belo Horizonte, where one can observe different types of actors located at various levels of steepness of the terrain, *suitability* is not considered an essential factor when explaining urban structure and developments. Preliminary outcomes of the econometric model also confirmed the statistical insignificance of this factor (see chapter 6).

Zoning, in turn, as explained in chapter 4, only plays a stronger role in Belo Horizonte starting in the second half of the 1990s, and even then, zoning is restricted to parameters of preferred density per area. This differs from other applications where certain land-uses are either permitted or prohibited or where future developments are planned or written in master plans.

3.4 Implementing the model: the METRONAMICA System

The simulation is implemented using METRONAMICA[38]. METRONAMICA is a spatial dynamic land-use modeling toolkit developed within the Research Institute for Knowledge

0	0	0	0	0	0	0
0	0.03	0	2	0	0.03	0
0	0	17.85	28	17.85	0	0
0	2	28	95	28	2	0
0	0	17.85	28	17.85	0	0
0	0.03	0	2	0	0.03	0
0	0	0	0	0	0	0

Figure 3.3 Illustration of influence of a high-income cell for high-income transition potential on its neighborhoods

Systems (RIKS), based on White and Engelen's (1993) seminal paper. The system is defined as "a modeling and simulation package for development of high-resolution land-use models" (Riks, 2007b). METRONAMICA allows for the import and export of possibilities that make it easier to use in conjunction with usually static commercial GIS software.

The package enables modeling both supply and demand-side forces that drive spatial development using quantitative or qualitative social, economic, demographic, environmental, and physical information on different spatial scales (Riks, 2005).

3.5 Calibration procedures and the importance of parameters

Calibration is the process of adjusting the parameters of the model so that actual land-use developments are best reproduced by the model.

More specifically, in the case at hand, the objective of the calibration is to establish the sensitivity of the urban area for different factors and to find the parameters that properly express this sensitivity. Thus, we need to quantitatively determine the importance of inertia effects, accessibility, etc. for different actors.

Because in a complex system as described here, there could be many different (and possibly conflicting) sets of parameters that lead to an approximate replication of current land-use, we feel that calibration should start from initial parameter values that are based in theoretical considerations and effects reported in the literature, which are then tested to see if the results (emerging structures) resemble observed data. This approach differs from statistical 'brute force' methods in which the aim is to mimic land-use change trends through machine learning-processes such as SLEUTH (Silva and Clarke, 2002).

3.6 Validation

The parameter values resulting from the calibration can only be accepted for use in scenario analyses after validation. That is, using calibrated parameters, the model should produce outcomes in a variety of scenarios that are realistic with respect to the literature as well as historic information.

Validation can be performed by comparing simulated results for a certain year with empirical data. Batty and Torrens (2005) emphasize that calibration and validation should use different datasets. Brown *et al*. (2005) also stress that validation may be achieved either by replicating processes or by comparing results.

For this thesis, the emphasis is on the underlying processes and validation is undertaken both by comparison with the general structure of the actual map of actors of 2000 (Figure 7-3) constructed in chapter 7 (validation part I, item 8.7), and by comparing outcomes to the econometric analyses presented in chapter 6 (validation part II, item 9).

Part II

The second part of this thesis applies the concepts and models discussed in part I to the city of Belo Horizonte and its metropolitan regions, further developing the definition of neighborhoods and differentiated income-level influences on urban development.

The first chapter of part II provides a historic background for the case study, and is followed by a definition of neighborhood and an empirical application to Belo Horizonte. Then, an econometric analysis is presented, followed by a cluster analysis that defines the reference and validation cases for the CA model for the Greater Area of Belo Horizonte (RMBH).

4 Historical contextualization

This chapter has two main goals. First, it reinforces an acknowledgement that Brazilian urbanization has been through concentrated and accelerated growth and its conformation is heterogeneous[39]. Second, it describes the rapid urban growth seen in Brazil, the state of Minas Gerais and the Greater Area of Belo Horizonte (RMBH) in the second half of the twentieth century and the specific characteristics of RMBH.

4.1 Brazilian rapid urban population growth

According to the 1940 census – the first reliable census to distinguish the rural from the urban population – 31.24% of the population is urban (Table 4-1). The urbanization index, which is the ratio of urban population to the total population, is constantly growing, and reached 81.25% in the last census (2000). The rural population, however, reached an apex of 41 million in 1970, and has decreased since.

Although intensification of Brazilian urbanization started in the 1940s and 1950s, and was reinforced in the 1960s and 1970s, growth rates continued to be very high, reaching almost 3% every year in the 1980s and remaining at 2.5% in the 1990s. This represents an added population of over 57 million in just the last two decades of the twentieth century. The urban population consistently grew faster than the total population. This secular trend continued from 2000 through 2005, with stable numbers observed for the rural population and continuing growth of urban areas.

The high rate of urbanization is further expressed when considering southeast metropolitan areas, in which the urbanization index reaches 91.8%. For the city of Belo Horizonte, urbanization is 98.7% (Brasil, 2006).

These numbers refer to the official national statistical foundation's definition of a census tract. They include all legally urban areas, described in the documentation as an "urban area of city or town; non-urbanized areas of city or town; isolated urban areas" (IBGE, Documentação de setores censitários). For Belo Horizonte, at the municipal level, Decree 10925 of 2001 reported that 100% of the city's territory was urban.

Given these numbers regarding population growth, it is easy to conclude that Brazilian cities did not have adequate time, resources, or political institutions (Gouvêa, 2005) to accommodate the incoming population in an appropriate manner[40]. This fact has also been observed in other countries in other periods, such as London at the beginning of the Industrial Revolution (Benevolo, 1980).

4.2 Historical description at the state level

4.2.1 Establishing Belo Horizonte as Minas Gerais' capital

Ouro Preto, a municipality that was legally established in 1711 (Costa, J., 1970), prior even to the creation of the Province of Minas Gerais in 1720 (Figure 4-1), was basically an administrative and political center of power. Given the economic structure of the Colony at the time, produced minerals went straight to maritime ports in Rio de Janeiro (Singer, 1977, p. 199), which played the role of commercial center. Space was scarcely populated and presented an "archipelago" of independent islands, as it was described by the famous Brazilian economist Celso Furtado (1956). Furthermore, the concentration of people near the recently discovered mines promoted occupation of distant locations in large steps, rather than as the continuous progression that is usually the case in an agricultural frontier expansion. As a result, virgin unoccupied land remained between settlements. This also required the considerable extension of the supply lines because food production was limited (Singer, 1977, p. 200). As such, as early as the eighteenth century, a precarious network of pathways connected: (a) the Province of Bahia to Minas, providing cattle and mules to Minas, (b) the Provinces of Goiás and Mato Grosso, which were areas of ongoing mining expansion, (c) the Province of Rio de Janeiro, and (d) São Paulo, with the newly constructed pathway called "New Path"[41].

The end of the peak of the gold cycle at the end of the eighteenth century left behind a legacy for the Province of Minas Gerais of an ample economy of subsistence and self-sufficiency after the second half of the century (Singer, 1977. p. 203). This province was comprised of an incipient network of cities that was able to sustain part of the population all the way into the next economic growth cycle of coffee in half of the following century. The relative isolation of the region, given the distance to (800 km to Rio de Janeiro) and precarious accessibility of other regions, in conjunction with the available worker pool freed from the mining chores, triggered an initial and important process of substitution of imports, which was readily suppressed by the Portuguese Crown (*op. cit.* p. 205).

The coffee economy boomed after 1840, reaching 7% of Brazilian total exports by 1850. This boosted the construction of Emperor Pedro II's railway line that reached the south of Minas Gerais in 1880 (Singer, 1977, p. 208). At the same time that a better connection promoted an outgoing channel for coffee exports, it broke the natural monopoly of goods observed at the time, which was derived from the poor accessibility of Rio de Janeiro from Minas Gerais.

Processes focused on exports, be it extracted minerals or coffee production, contributed to a strong connection to the outside world, so that the archipelago configuration (independent regions connected to outside nodes) remained in place, which resulted in weak internal links among established cities and towns. Within this framework, the capital of the province, Ouro Preto, did not benefit directly from the new booming coffee industry, and remained exclusively as political-administrative center that lacked significant economic power[42]. This effect was such that some separatist movements – following the movement that detached the Province of Paraná from São Paulo in 1853 – emerged[43]. At the time that a proposal for a change of capital was made and vetoed by the President of the Province, an idea came forward into the political discussion at the end of the Empire and Proclamation of the Republic on 15 November, 1889.

As ordered by the newly sworn-in Governor Domingos Rocha, engineer Ferreira Pena suggested a new location for the capital with main justification being the development of a central node that would enable and foster communication among regions in the state (*op. cit.*

p. 216). The town of Curral del Rei was originally a local producer of agriculture and cattle that supplied the still-operational mines of the valley of Rio das Velhas and that was chosen as the place for the new capital instead of traditional competitive economic centers of the time, including Barbacena, Paraúna, Várzea do Marçal and Juiz de Fora. The site was soon named Belo Horizonte. In 1895 it was connected to the developing railway network called the Central do Brasil.

The construction of the new, planned capital was approved and the means to do so were made available.

4.2.2 Urban-economic evolution

The beginning of the expansion of Belo Horizonte coincided with coffee's super-production crisis and the ensuing decline in prices around the turn of the century, which depressed the Republican economy as a whole. Minas coffee exports decreased from a staggering 81% of total exports in 1888 to a mere 38% in 1908 (*op.cit.*, p. 228). The export of subsistence products, in turn, grew during the period.

The cotton industry on the northwest border of Belo Horizonte, especially in Santa Luzia and Sete Lagoas, also grew in the first two decades of the twentieth century. By 1920, Belo Horizonte had become a regional economic center with a number of educational institutions, with better accessibility to other centers due to the implementation of a 'West Railway' that connected the city to Divinópolis and the southern region of the state, as well as to the west and the regions of Triângulo, Campos das Vertentes and Alto Paranaíba.

Belo Horizonte was then the third leading industrial municipality of the state, with the clear leader being Juiz de Fora. Some mining municipalities close to the capital, such as Conselheiro Lafaiete and Nova Lima, were also important at the time (Singer, 1977, p. 236).

The First World War, along with new mining projects, such as the Itabira Iron Ore Company, which was installed in 1911, and Companhia Siderúrgica Belgo-Mineira, which was started in 1921 in Sabará, a neighboring municipality, accelerated the import substitution process, especially in metallurgy and siderurgy, reaching 43.5% of national production in 1929.

Table 4.1 Brazilian urban and total population growth Sources: Estatísticas Históricas do Brasil/ volume 3 – Rio de Janeiro: IBGE, 1987; Anuário Estatístico do Brasil/IBGE – Rio de Janeiro, volume 56, 1996; IBGE, Censo Demográfico 2000; IBGE, Síntese Indicadores Sociais 2005

Years	Total	Urban	Rural	Urbanization index	Rate of urban pop. annual growth	Rate of total pop. annual growth
1940	41,236,315	12,880,182	28,356,133	31.24%		
1950	51,944,397	18,782,891	33,161,506	36.16%	3.84	2.34
1960	70,070,457	31,303,034	38,767,423	44.67%	5.24	3.04
1970	93,139,037	52,084,984	41,054,053	55.92%	5.22	2.89
1980	119,002,706	80,436,409	38,566,297	67.59%	4.44	2.48
1991	146,825,475	110,990,990	35,834,485	75.59%	2.97	1.93
2000	169,799,170	137,953,959	31,845,211	81.25%	2.45	1.63
2005	184,388,620	152,711,363	31,677,257	82.82%	2.05	1.66

In 1926, the railway lines extended all the way to the north of the state to the city of Montes Claros, and in 1930 another rail line 'Estrada de Ferro Vitória-Minas' connected the central area of the state to the port of Vitória and the Atlantic coast in the east. In 1939, the production of iron ore and steel was concentrated in the central area of Minas Gerais and it was four times higher than its imported value (SINGER, 1977, p. 240-249).

Despite its upward trend up to the 1930s, the national economic importance of the city was small due to the lack of other cities and centers of economic activity in the north of the state, which left the city quite isolated. However, the function of railway nodes, along with the beginning of an internal consumption market, which was also boosted by the relatively high salaries paid to public workers, who accounted for one sixth of the active population in 1912, slowly gave more prominence to the city.

Construction and textiles were the first industries to pick up pace, with the Belo Horizonte textile industry reaching second place within Minas Gerais in 1908 (*op.cit.*, p. 222-223). In the national comparison, the scattered 'Mineiro'[44] production reached 75% of São Paulo's total production and a bit over half of Rio de Janeiro's.

The mining boost of neighboring municipalities and agriculture growth in Rio Doce's valley, especially in the 1930s and 1940s, augmented the total wage paid in industry and the demand for industrialized products which were further concentrated in the capital.

The Second World War, similar to what happened in 1914, fostered a further process of import substitution and the continuous growth of production of siderurgy and with other industrial sectors, such as timber, textile, and dairy products. Singer suggested that by this time, Belo Horizonte had started a process of rural-urban specialization in which the city increasingly became a center for redistributive activities (1977, p. 253).

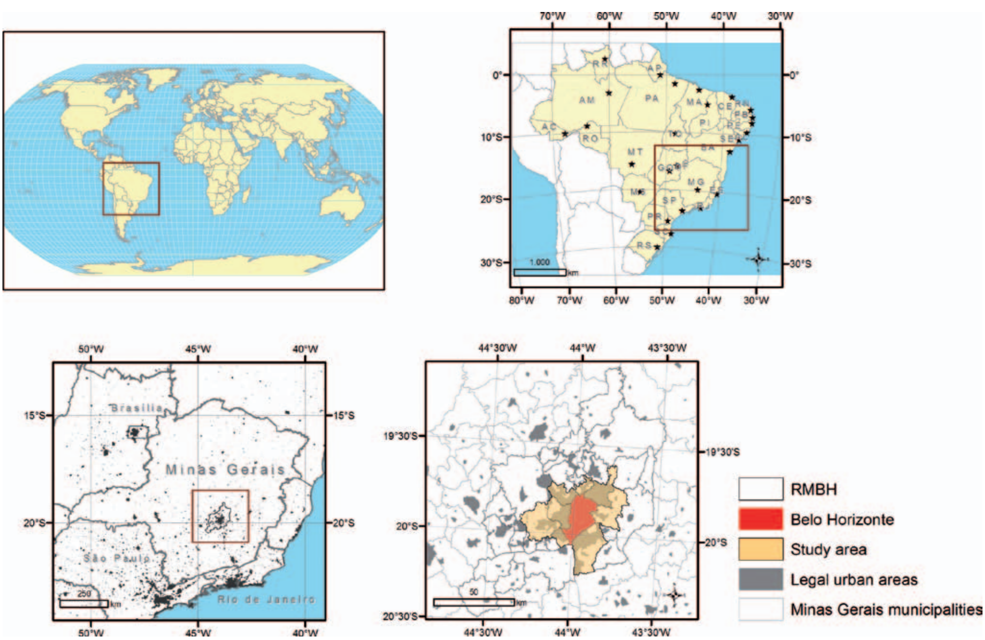


Figure 4.1 Location of Minas Gerais and RMBH within Brazil

The decision to build a national siderurgy company in Volta Redonda, Rio de Janeiro, and its effective beginning in 1946 threatened the thus-far 'Mineiro' supremacy in siderurgy. This, in turn, strengthened the resolve of state legislators to diversify industrial production and the consequent approval of a decree institutionalizing the municipality of Contagem, 10 km from Belo Horizonte, as an industrial city.

Contagem was chosen for this designation because it had (a) road and rail accessibility; (b) a water supply; (c) a reliable energy supply, planned with the construction of Gafanhoto's plant; and (d) cheap and available land. The actual supply of energy became effective in 1955 which facilitated the consolidation and implementation of a number of industries.

In the broader state perspective at that time, there was also the implementation of a great special steel plant, Acesita (Aços Especiais Itabira S.A.), founded in 1944 in Timóteo and Usiminas (Usinas Siderúrgicas de Minas Gerais S.A.), which was rapidly built and brought into operation in 1956-1962 in Ipatinga.

4.2.3 1950s and 1960s: "urban land production" and speculation

Heloísa Costa, a geographer and professor of architecture who studied and wrote on Belo Horizonte, characterizes urban land production (in Brazil, popularly referred to as urbanization) as the process of transforming rural areas into urban ones through parceling the land (Costa, H., 1994)[45]. She differentiates the dynamics of producing urban land (and putting it forward for sale) from the effective occupation of land by the population, which occurred many times some decades later.

The logic described by Costa suggests the triad of production-appropriation-occupation, in which the production of land that usually occurs in distant localities without any infrastructure. The strategy from the perspective of the price is described by Costa:

> The central point of the land producer is based on the fact that the final product – the parcel – has to be the lowest value possible so that the potential market be the greatest possible. That means that the price to be charged for each parcel determines not only the process of production but also the quality of the delivered product. Therefore, a number of costs are passed on to the buyer or simply eliminated altogether and an efficient financing scheme has to be put forward … the first is [that] of the construction itself [which is] totally borne by the buyer … Second, [is] the cost of the land (Costa, H., 1994, p. 64, our translation).

Sales take place slowly and continuously as part of a long process by which land is "appropriated"[46]. Once residents are in place and have this connection with their neighborhoods, they start processes of self-organization and political demand arises for development and improvement of infrastructure (which usually include closer bus lines, schools, health-centers, and kindergarten in public schools). Once the area has been provided with these public services, actual occupation takes place with real estate brokers selling the reserved stock of land that is now much more valued given responses from authorities to demands of earlier residents. Then, a newer cycle takes place.

This process was applied successfully to the Greater Area of Belo Horizonte, and it is important to understand the process of expansion of urban occupation.

The production-appropriation-occupation triad continued to proceed from the 1950s to the 1970s, until it was strongly curtailed by the effect of federal legislation in the late 1970s. The Law of Urban Land Parceling – Law 6766 of 19 December 1979 – imposed on the land producer a number of demands, specified in Chapters II and III (Of urban requisites for parceling and Of parceling projects). Article 37 prohibited the sale of non-registered parcels.

Although the law can be considered successful, illegal parceling still occurred. By the time the law was passed, much of the rural land in a large area around the metropolis was already "urbanized".

4.2.4 Population growth of RMBH

The population growth of the RMBH (Table 4-2) clearly demonstrates that municipalities that were contiguous to the capital, especially on the northern boundaries (Santa Luzia,

Table 4.2 RMBH's municipality population growth 1920-2000. Source: Estatísticas Históricas do Brasil/volume 3 – Rio de Janeiro: IBGE, 1987; Anuário Estatístico do Brasil/IBGE – Rio de Janeiro, volume 56, 1996; IBGE, Censo Demográfico 2000.

Municipality	1920	1940	1950	1960	1970	1980	1991	1996	2000
Baldim			9819	10264	9362	7567	8383	7935	8155
Belo Horizonte	55563	211377	352724	693328	1235030	1780839	2020161	2091371	2238526
Betim		19930	16376	29960	37815	84193	170934	249451	306675
Brumadinho		10836	13018	14313	17874	18018	19308	24336	26614
Caeté	17917	20872	21911	28131	25166	30630	33251	34869	36299
Capim Branco				3134	4147	4930	6344	7070	7900
Confins									4880
Contagem	12140		6022	28065	111235	280470	449588	492214	538017
Esmeraldas	15399	10850	14311	15310	15698	16215	24298	33934	47090
Florestal					4430	4809	5053	5363	5647
Ibirité					19508	39967	92675	126627	133044
Igarapé					7675	16561	27400	31063	24838
Itaguara			7107	8325	9030	9763	10671	11225	11302
Itatiaiuçu					5330	5426	7366	8243	8517
Jaboticatubas		24874	16357	16536	12159	11569	12716	12409	13530
Juatuba								12306	16389
Lagoa Santa		6838	7738	11559	14053	19499	29824	35026	37872
Mário Campos									10535
Mateus Leme		11165	11676	13338	11929	18659	27033	20720	24144
Matozinhos			9768	9109	8674	16201	23606	26722	30164
Nova Lima	17448	29714	21932	28223	33992	41217	52400	56960	64387
Nova União					3958	4066	4865	5148	5427
Pedro Leopoldo		17821	15729	16382	20670	30007	41594	47342	53957
Raposos			6411	8402	10133	11801	14242	14630	14289
Ribeirão das Neves				6391	9707	67249	143853	197025	246846
Rio Acima			5276	5099	5118	5073	7066	7556	7658
Rio Manso					5214	4445	4461	4276	4646
Sabará	8045	11060	13310	23098	45149	64210	89740	100539	115352
Santa Luzia	58376	18321	10875	12753	25301	59893	137825	153914	184903
São Joaquim de Bicas									18152
São José da Lapa								12201	15000
Sarzedo									17274
Taquaraçu de Minas					4034	3455	3383	3369	3491
Vespasiano			5610	8335	12429	25046	54868	60952	76422
Total	184888	393658	565970	990055	1724820	2681778	3522908	3894796	4357942

Vespasiano and Ribeirão das Neves) and on the southern boundaries (Ibirité), combined with the industrial development vector to the west (Contagem and Betim), were the areas that most extensively suffered the effects of the rapid occupation described by Costa (1994). Most other municipalities, which are non-contiguous, experienced growth similar to the national average. Nine municipalities (26% of the total of 34 municipalities) had fewer than 10,000 inhabitants in the 2000 Census.

4.3 Specific intra-urban evolution of RMBH

As introduced in the beginning of the chapter, Belo Horizonte is a planned city that represents an intentional policy outcome.

In terms of urban design, a proposal of the planned city based on L'Enfant's design of the American capital was put forward by Aarão Reis, the appointed chief-engineer. It consisted of the superposition of two squared grids: one with "generous and costly" local streets of 20 meters width; and another grid 45 degrees from the first with avenues of 35 meters width (Singer, 1977, p. 219). The city was organized in a succession of three rings, in which the official urban area would amount to 8,815,382 m^2 which is nearly 832 hectares surrounded by a Central Ring (Figure 4-2). The central area would then be enclosed by a sub-urban *boulevard* of 24,930,803 square meters which would, in turn, be encircled by a rural area[47].

According to economist Paul Singer (1977, p. 220), initially, 3,639 parcels (from 480 to 600 m^2) were put on sale. This represents not more than 25% of the total urbanized area. The region that was effectively occupied was concentrated between avenues Cristóvão Colombo and Araguaya (presently, Francisco Sales) (Villaça, 1998) (see annex 11.1), and at the time of inauguration on 12 December 1897, the city contained about 12,000 inhabitants.

The sale of the first parcels was made through public auctions, which suggests that an active and speculative real estate market was already in place (Costa, H., 1994).

The initial occupation of the urban space occurred in the neighborhood now known as Funcionários (which means *workers*) (see Figure 11-1). The imposition by the City of time-limits to actually finish the construction of the houses in the parcel restricted initial expansion to individuals who could afford fast-paced construction. Commerce and service pursued their own interests because there were no specific zoning or planning laws to follow.

They first occupied Rua da Bahia, which was the street that connected the railway station to the most valued, reserved areas near the Governor's Palace and State Departments (Villaça, 1998, p. 120) (Figure 4-2). The western slope (where today is the most valued neighborhood of the capital) was only offered for sale later when the number of high-income domiciles in the city increased (see Figure 4-3 to Figure 4-5) (Villaça, 1998, p. 200). Since its beginning the city presented a polarization of occupation based on income-class divisions.

In 1936, Barro Preto's industrial area, which runs along the railway lines, was installed with an initial number of 20 companies. The population already occupying the area is moved into other neighborhoods (Costa, H., 1994, p. 55) (Figure 4-6). This began the process of space succession that occurred intensely throughout the 1950s (Figure 4-7), specifically in relation to slums (Somarriba, Valadares *et al.*, 1984; Furtado, B.A., 1995).

The high price of the input infrastructure helped to maintain the extremely high price of the parcels, and as a consequence, "we still have in 1940, whole blocks vacant within the Central Ring" (Villaça, 1998, p. 124) (Table 4-4).

This fact occurred again in relation to the development of a neighborhood known as Pampulha, where up to the 1960s, most of the parcels put up for sale in the 1940s remained unoccupied (Villaça, 1998, p. 201). Part of the reason for this was the preference of high-income families to be located in the southern area of the city. At the same time that parcels in Pampulha, in the north, remained vacant, new development near the mountain of Serra do Curral, in the southern limit of the municipality, was launched. Those neighborhoods, such as Belvedere or Mangabeiras, approved in 1970 and 1973, respectively, were nearly totally occupied 20 years later. The justification presented by Villaça for this preference for the southern vector coincided with the proposal of this work. He stated:

Table 4.3 Population growth of Belo Horizonte 1897-2006 Source: Estatísticas Históricas do Brasil/volume 3 – Rio de Janeiro: IBGE, 1987; Anuário Estatístico do Brasil/IBGE – Rio de Janeiro, volume 56, 1996; IBGE, Censo Demográfico 2000; (Brasil, 2006)

Year	Belo Horizonte's population	Annual rate of growth	Belo Horizonte's percentage in relation to total RMBH
1897	12000		
1900	13472	3.93	
1912	40365	9.58	
1920	55563	4.08	30.1%
1940	211377	6.91	53.7%
1950	352724	5.25	62.3%
1960	693328	6.99	70.0%
1970	1235030	5.94	71.6%
1980	1780855	3.73	66.4%
1991	2020161	1.15	57.3%
1996	2091371	0.70	53.7%
2000	2238526	1.71	51.4%
2006	2399920	1.17	

Table 4.4 Evolution of urban occupation in Belo Horizonte. Source: elaboration of the author based on the maps of Figure 4-2.

Year	Urban stain aprox. area in hectares	Annual growth	Percentage within planned central ring	Vacant area within planned central ring
1900	502		46.8%	73%
1910	793	4.68	53.8%	52%
1920	1676	7.77	32.8%	38%
1930	2935	5.77	23.6%	22%
1940	4850	5.15	18.2%	0%
1964	13042	4.21	6.8%	0%
2000	33080	2.62	2.7%	0%

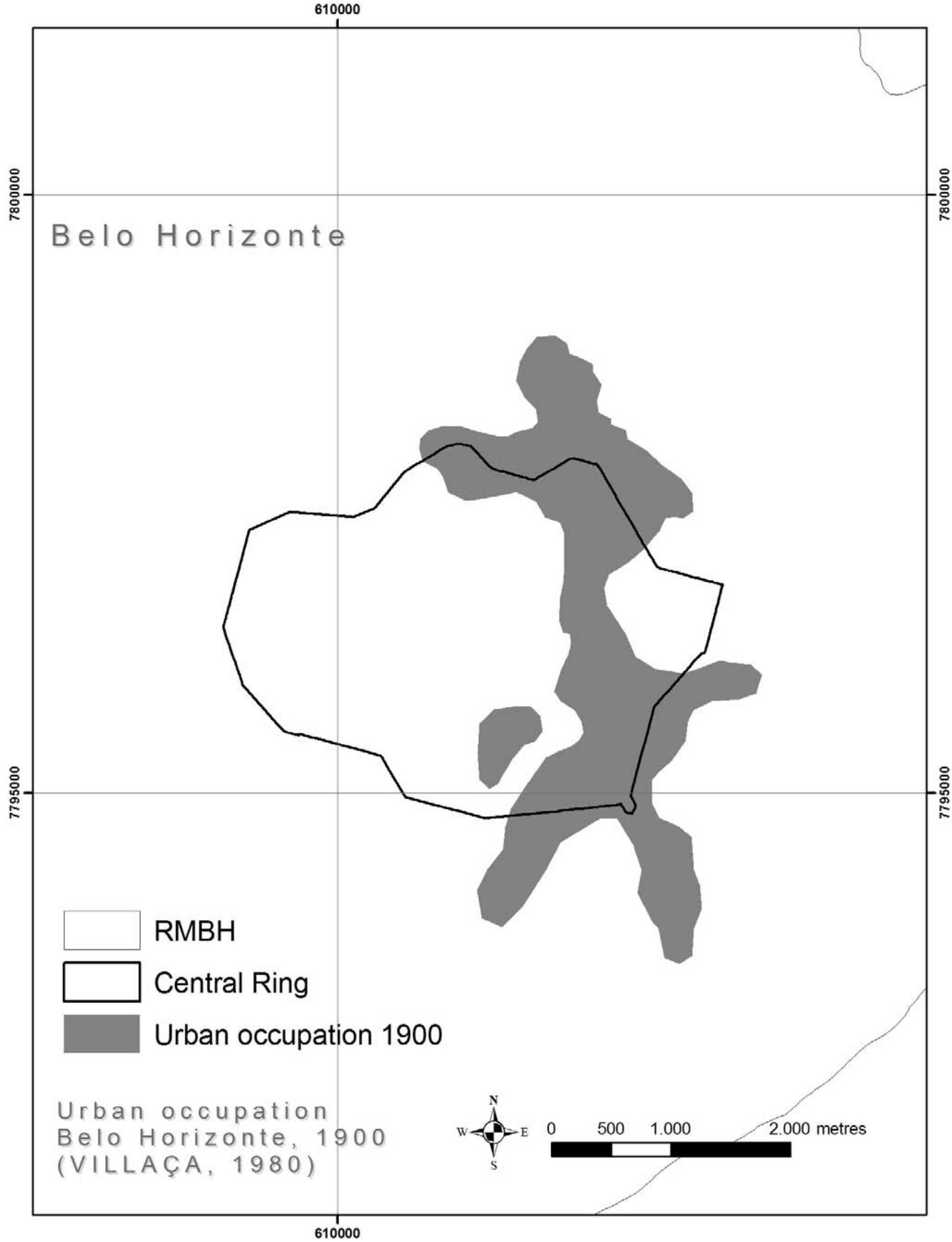


Figure 4.2 Urban occupation in Belo Horizonte 1900. Source: Villaça (1998)

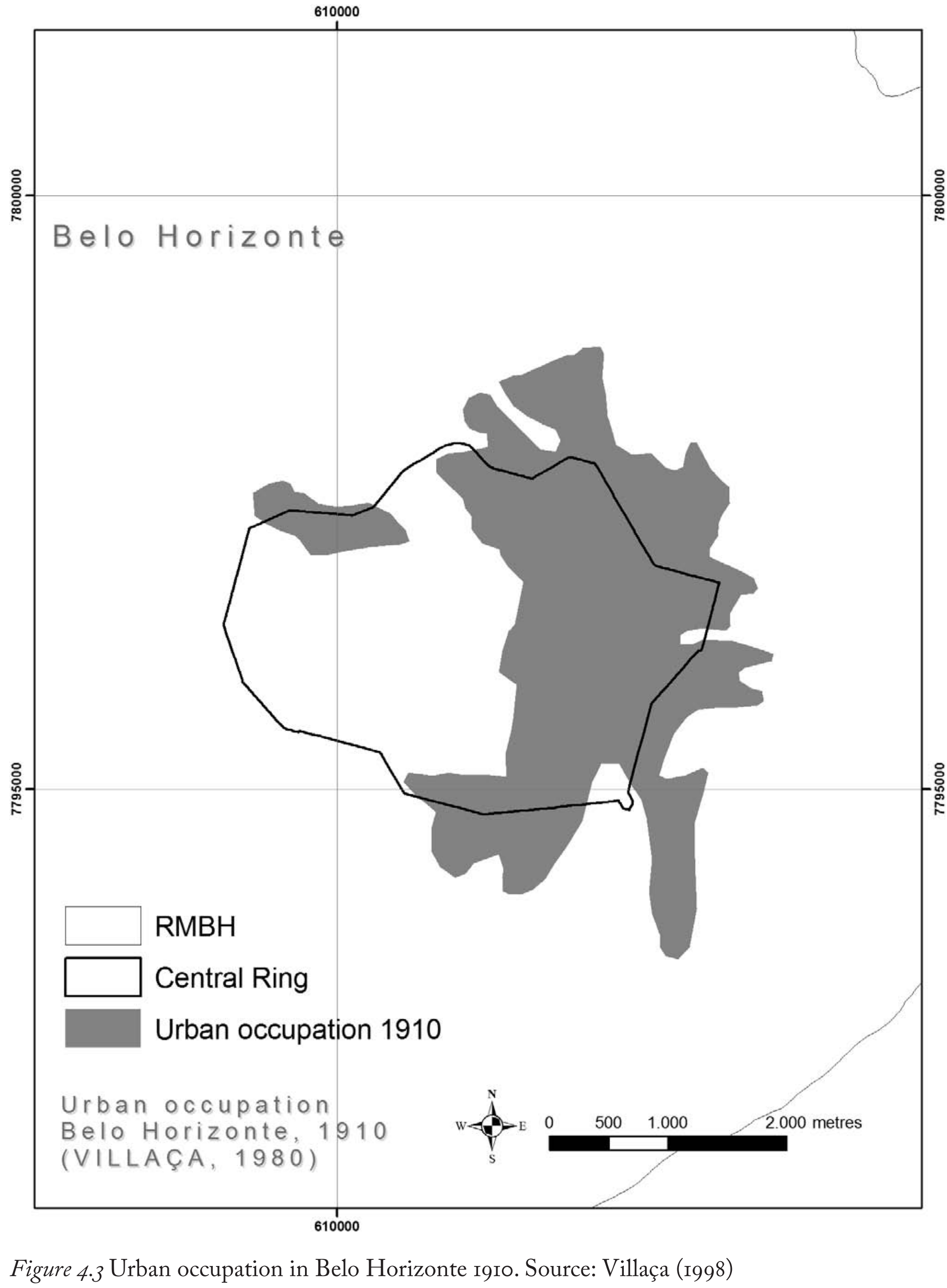


Figure 4.3 Urban occupation in Belo Horizonte 1910. Source: Villaça (1998)

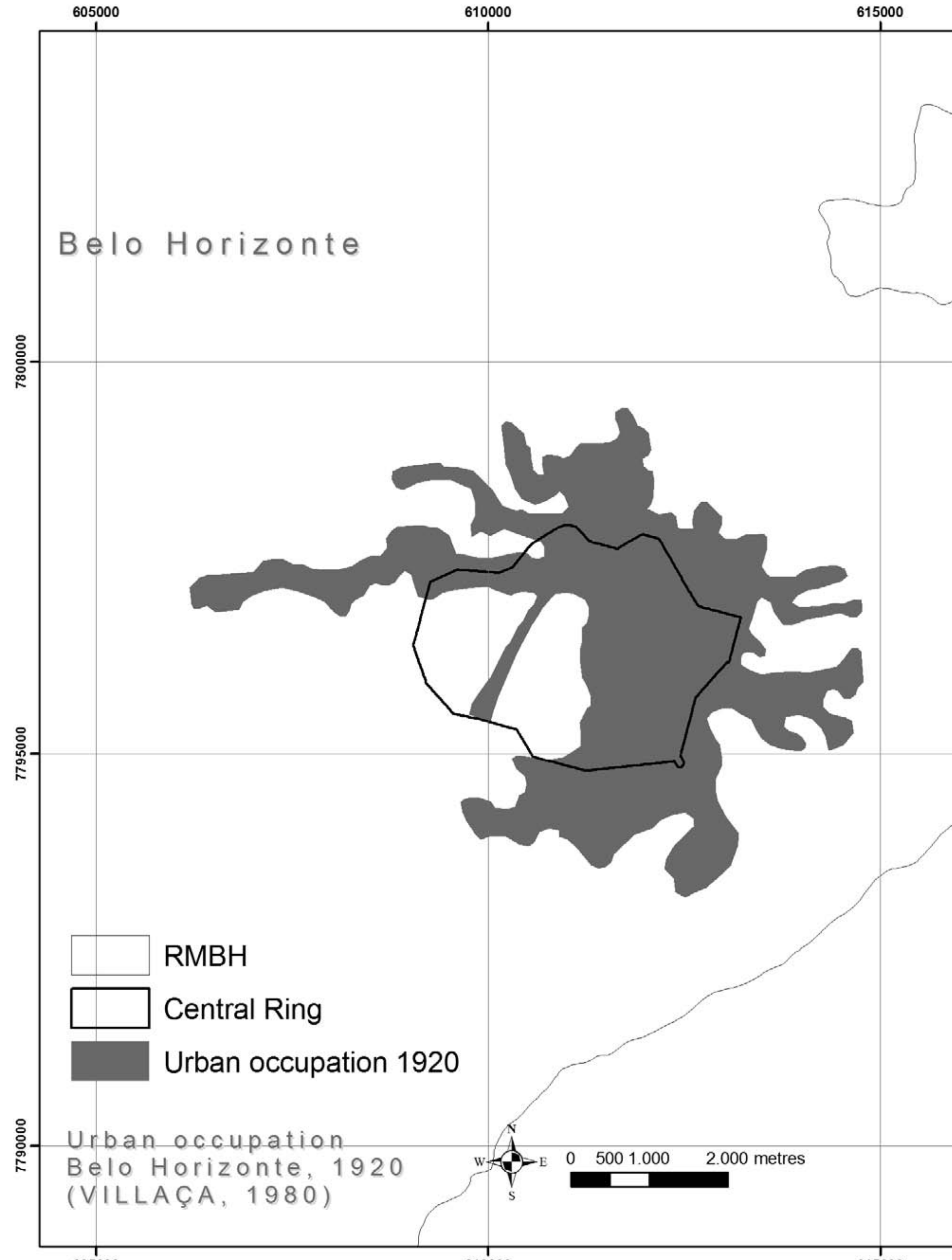


Figure 4.4 Urban occupation in Belo Horizonte 1920. Source: Villaça (1998)

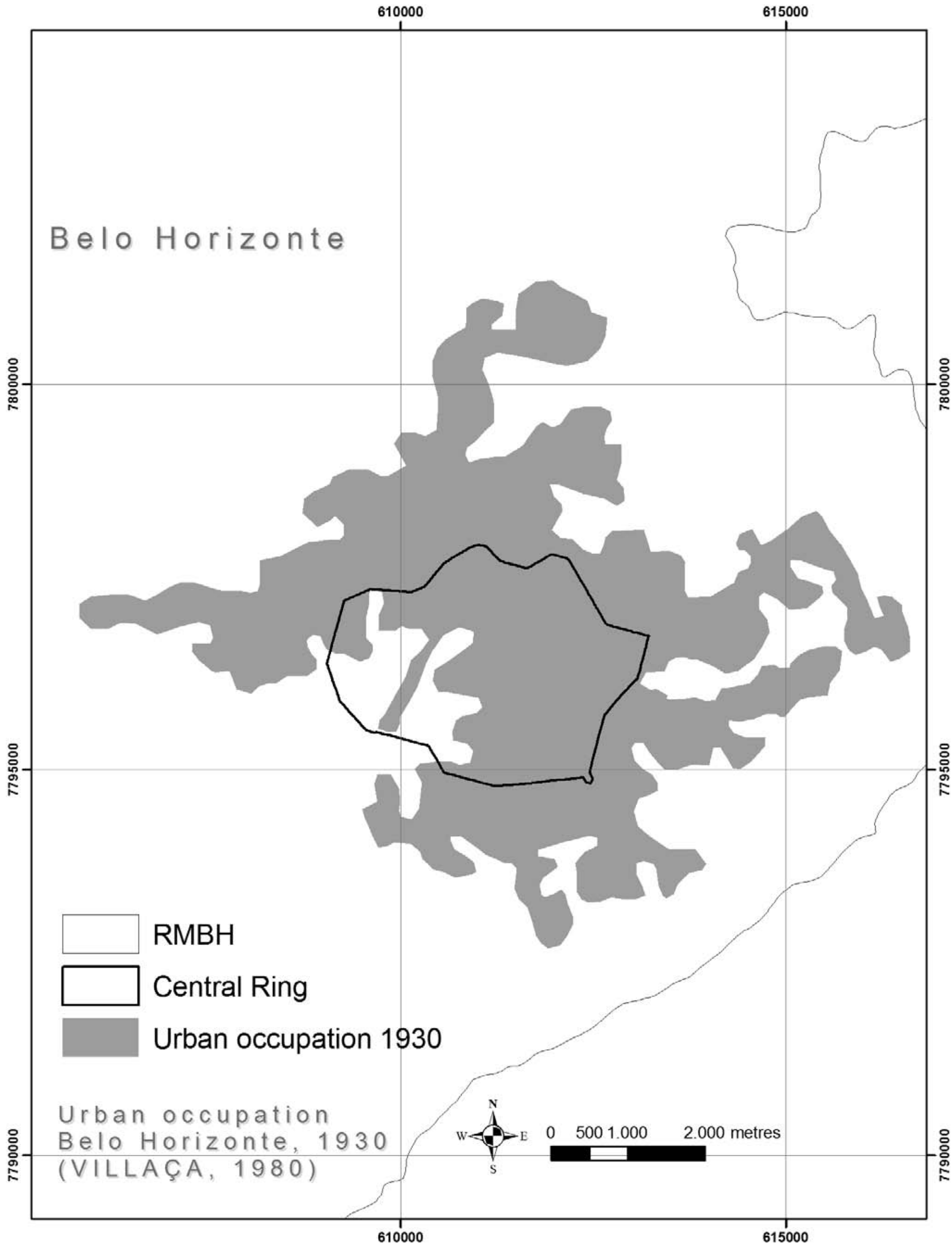


Figure 4.5 Urban occupation in Belo Horizonte 1930. Source: Villaça (1998)

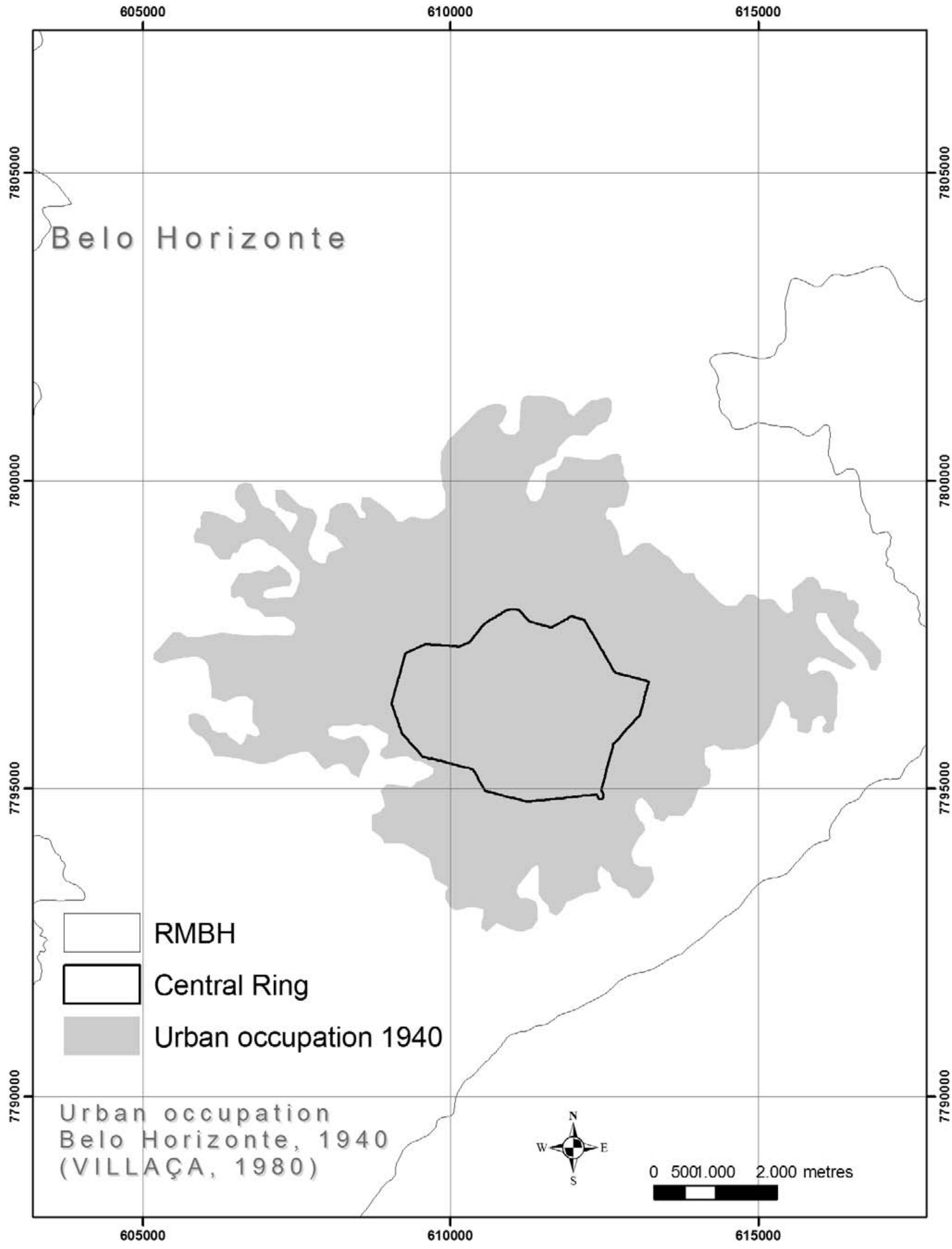


Figure 4.6 Urban occupation in Belo Horizonte 1940. Source: Villaça (1998)

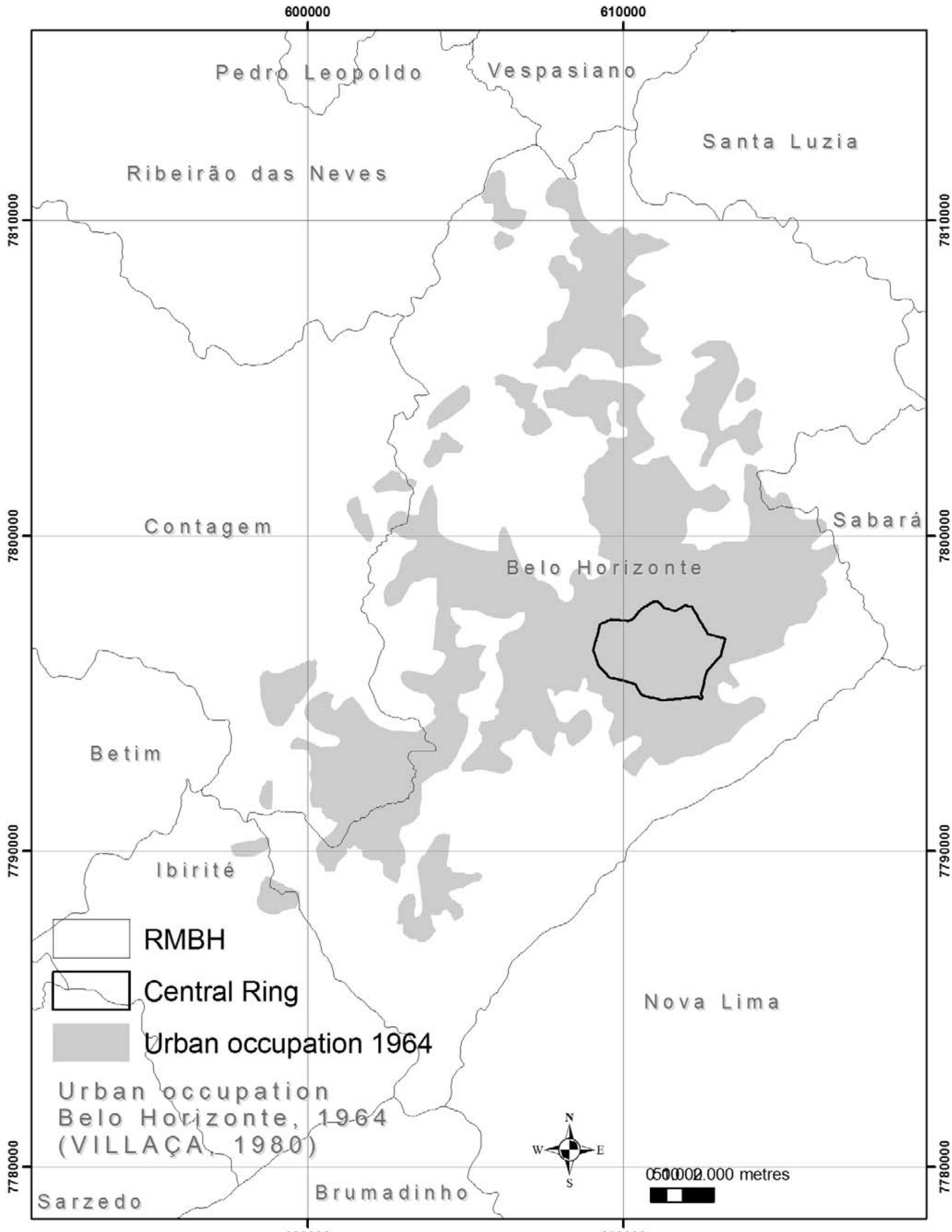


Figure 4.7 Urban occupation in Belo Horizonte 1964. Source: Villaça (1998)

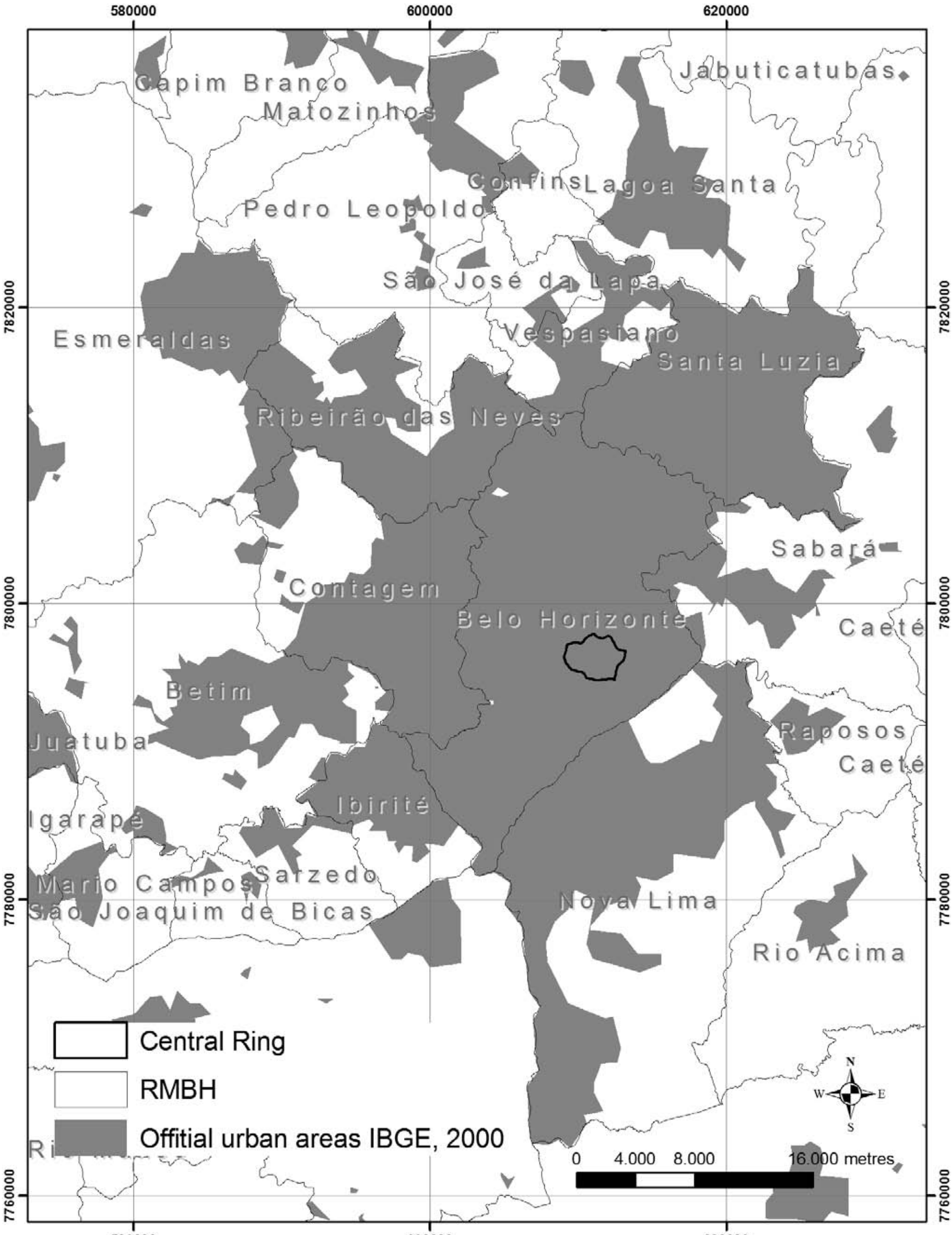


Figure 4.8 Official urban areas RMBH 2000. Source: IBGE, 2000 (Brasil, 2003)

> This happens due to the spatial inter-relations that develop (one should be reminded that we are studying a structure) enhancing further the 'tidiness' within communities and neighborhoods (especially high-income ones), and other vital elements in defining spatial structure, namely shopping places, sub-centers. … a movement toward moving out, abandoning the high-income segregated areas would mean 'staying away from everything that is meaningful and urban-like' (Villaça, 1998, pp. 202-203, *our translation*).

In terms of population growth (Table 4-3), rates as high as 4% are observed in the first decade after the incorporation of the city. In these early stages, the urban occupation area accelerated, reaching nearly 800 hectares in 1910 (Table 4-4) and was concentrated along the main north-south avenue, Avenue Afonso Pena (Figure 4-2). However, roughly half of the total occupation was outside bounds of the Central Ring, which remained partially unoccupied.

Urban occupation continued its growth (Villaça, 1998), and more than doubled in the next decade, with population increases of 15,000 inhabitants per year, which represents a growth rate a bit higher than 4% of annual growth.

Consistent with the rest of the country, the capital of Minas Gerais showed constant and accelerated growth from the 1920s onward, with rates of almost 7% between 1950 and 1960. Urban occupation, in turn, expanded faster than the population and also reduced its pace sooner, showing less steep rates as early as 1940 (Table 4-4). About this time, urban occupation filled the planned area within the central ring; however, many vacant parcels remain today without a construction building (Santiago, 2006). Also from the 830 hectares planned for the original city, urban occupation now stretches over 13,000 hectares, which is an increase of over 15 times the planned area.

Urban occupation, if defined by its legal status established by IBGE (Brasil, 2003), was nearly 100% of the municipality of Belo Horizonte in 2000 (Figure 4-8). On the other hand, neighboring municipalities conurbated with the capital are still surrounded by rural areas that separate the main urban occupation from the capital itself (Vespasiano, Sabará, Nova Lima, and Brumadinho).

In terms of the number of inhabitants, the city of Belo Horizonte increased to a peak in the 1970s when it held 70% of the total population of RMBH. Then, this number started decreasing, representing a bit over half of the total population in 2000 (Table 4-2). That is, the beginning of the 1970s confirmed the suggestion of Costa (1994) concerning land appropriation (after it has been produced by early inhabitants), followed by a proportional concentration of poorer inhabitants moving into contiguous municipalities, especially Ribeirão das Neves, Santa Luzia, and Ibirité.

4.4 Administrative and legislation issues

4.4.1 Municipalities boundaries

In terms of administrative boundaries, the evolution of the legal status of municipalities is in harmony with the intensification of economic growth. The fragility or strength of the municipal body is thought to play a singular role in the level of permissiveness for new land development and land speculation.

Belo Horizonte was separated from Sabará in 1893, and at the beginning of the twentieth century, it was bordered by Esmeraldas in the west, Santa Luzia in the north, Sabará in the east and Bonfim and Nova Lima in the south (Figure 4-9). In 1916, the municipality of Contagem was created on the western border. An administrative configuration in 1938 subdivided the new municipalities of Brumadinho in the southwestern area; Betim and then Mateus Leme, west of Contagem; Pedro Leopoldo in the northwestern area; and a great part of Santa Luzia and the newly created Lagoa Santa and Jaboticatubas in the north.

Vespasiano was separated from Santa Luzia in 1948 (along with 70 other municipalities in Minas Gerais) and Ribeirão das Neves was separated from Pedro Leopoldo in 1953 (along with other 97) (Furtado, B.A., 2003). The municipalities of Raposos and Rio Acima, Capim Branco, Matozinhos and Baldim, and Itaguara were separated in 1953. As noted by Furtado (2003), this profusion of new municipalities occurred happens mainly in periods of democratic rule.

The creation of these new municipalities via succession is relevant because they are characterized by an administrative configuration in which politically weak municipalities were more likely to allow the development of urban land in the 1950s (see item 4.2.3).

In 1962, 237 separation processes in the state of Minas Gerais occurred (Furtado, B.A., 2003, p. 39), elevating the total number of municipalities from 486 to 722, an increase of nearly 50%. This process was concentrated in the vicinity of RMBH earlier, and therefore, only the municipality of Ibirité was separated from Betim. The rest of the separations occurred in non-

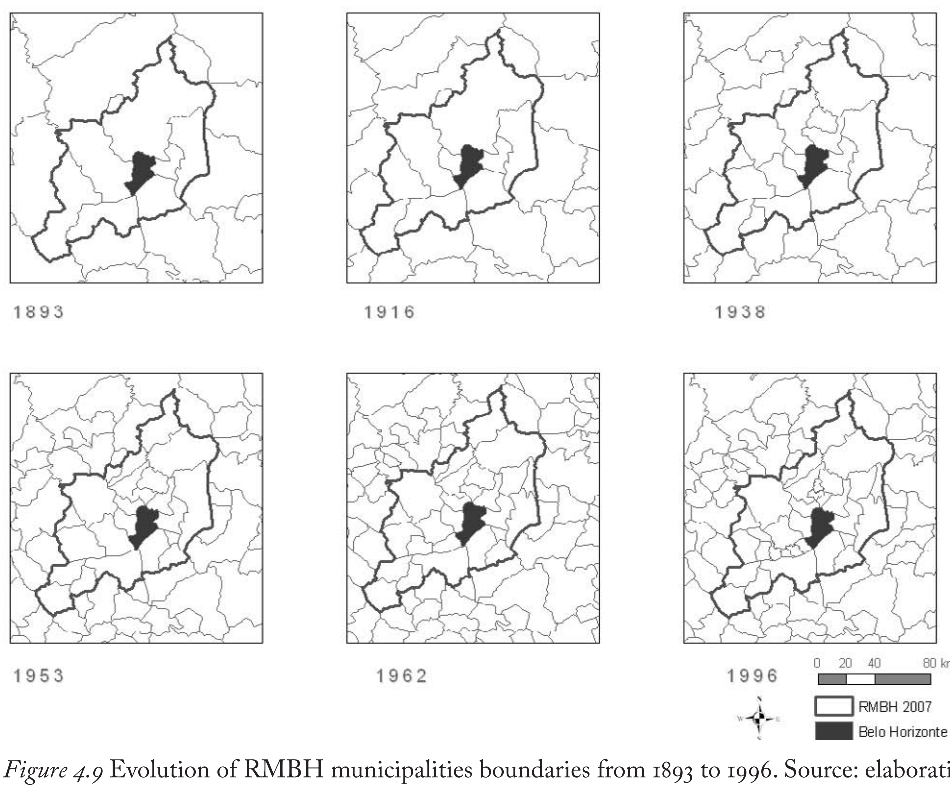


Figure 4.9 Evolution of RMBH municipalities boundaries from 1893 to 1996. Source: elaboration of author, based on Furtado (2003)

neighboring municipalities, where the following municipalities were created: Igarapé, Rio Manso and Itatiaiuçu, Florestal, Taquaraçu de Minas and Nova União.

Finally, in 1996, the municipalities of Confins (the location of the international airport since 1985), São Joaquim de Bicas, Juatuba, Mario Campos and Sarzedo were created.

4.4.2 A brief history of Belo Horizonte's land-use legislation

Legislation applied to the use and occupation of the land in Belo Horizonte was not a concern at the time of its creation. It was only with the unfolding of the Chicago School and Louis Wirth title *Urbanism as a way of life* (Wirth, 1938) that zoning became a factor in coping with city organization. The first norm to establish parameters for use occupation in Belo Horizonte was Decree 084 of 21 December 1940, which dealt basically with the construction itself, not with its use or location within the city (Canuto, 2005, p. 76). In the 1970s, Law 2662 (1976) was passed, but once again it was permissive and had little influence on the dynamics of the city. The law was passed in the context of the planning proposal of the Military Government, then in power, and the broader assumption of planning as a discipline. The following piece was Law 4034 in 1985, which again is evaluated as generating a low impact, according to the analysis of Canuto (2005, pp. 95-97). This series of weak legislative acts was due to the economic and political power being constantly held by real estate market actors since the origins of the city (Costa, H., 1994).

Along with the planning policies of the time and the centralized power held by the Union, the Greater Area of Belo Horizonte was created in 1973[48], along with its planning and think-tank, PLAMBEL, in 1974[49]. The original 14 municipalities that constituted RMBH were Belo Horizonte, Betim, Caeté, Contagem, Ibirité, Lagoa Santa, Nova Lima, Pedro Leopoldo, Raposos, Ribeirão das Neves, Rio Acima, Sabará, Santa Luzia and Vespasiano (Figure 11-2).

4.4.3 Description of the present administrative legal situation in RMBH and its trends

The extinction of PLAMBEL, which was very much associated with centralized, top-down government impositions, and was out-of-tone with the localization implemented after the 1988 Federal Constitution, had its functions taken over by the State Department of Planning and a State funded think-tank called Fundação João Pinheiro.

In Brazil, the post-military Federal Constitution of 1988 assigned to the states the ability to create metropolitan regions. The Minas Gerais State Constitution approved in 1989 suggested a model of management that did not prove to be successful. It was replaced by Addendum 65 to the state Constitution[50] on 25 December 2004, which proposed a new model.

This new model establishes a metropolitan management system comprised of (a) a Metropolitan Assembly, defined as "a high-order decision body" with five members from the state government and two (one from the executive and one from legislative bodies) from each municipality, (b) a 16-member Deliberative Board of Metropolitan Development which leads and coordinates tasks recommended by the Assembly[51], and (c) a Metropolitan Development Agency that aims to provide technical support and planning aid. This proposed system is supported by a Metropolitan Development Fund which is funded, among others, by the municipalities with an equal amount of capital from the State.

If this proposed "new metropolitan order" comes to fruition and public services provided at intra-metropolitan areas become more homogeneous, the incentives for citizens to locate in suburban municipalities will decrease. As it is, the level of services offered (healthcare, education, sanitation) is highly heterogeneous, and in turn, affect relative prices within the metropolis.

What has been demonstrated is that lower-income residents look for places with poorer availability of services because such places are obviously cheaper (Furtado, B.A., 2006).

4.5 Concluding remarks

This chapter demonstrates that urbanization processes in Brazil occurred rapidly in an interval of decades in the second half of the twentieth century[52]. Moreover, it shows that the pattern of concentration of population in a few scattered cities in the state, especially in the capital, which is, itself, isolated from other urban agglomerations in the state and in the country. Furthermore, institutional bodies are weak with regard to land control and planning. Development of new areas has historically shown a strong speculative trait.

This situation has led to an urban occupation that is essentially different from the densely populated, nearly continuous, constantly constructed, heavily planned and controlled urban developments usually observed in Europe and the United States.

5 The neighborhood

The integral urban analysis that is envisaged requires a thorough description of the urban fabric including socioeconomic factors in greater detail than usually available at the municipal level. A number of studies have shown the importance of neighborhoods in the housing market (Schnare and Struyk, 1976; Palm, 1978; Lee, Oropesa *et al.*, 1994; Cheshire and Sheppard, 1995; Archer, Gatzlaff *et al.*, 1996; Galster, 1996; Lee and Campbell, 1997; Bourassa, Steven C., Hamelink *et al.*, 1999; Kauko, 2002). However, no consensus has been reached on how to define or measure this phenomenon (Guo and Bhat, 2007). Neighborhood definitions in different studies vary in terms of spatial, methodological and substantive foci. The analysis of the impact of neighborhoods on urban housing markets and urban development processes requires a precise definition of the neighborhood, its characteristics and its relationships with other neighborhoods. Establishing this is the focus of this chapter.

Item 5.1 discusses the concepts of neighborhoods and sub-markets in the literature and proposes a definition of neighborhood that would be adequate for studies of real estate valuation. Item 5.2 proposes a spatial unit of analysis for the empirical case study that incorporates aspects of the established definition. Then, the data and methodologies are presented in items 5.3 and 5.4, followed by an analysis of the four indices that describe the neighborhoods of Belo Horizonte in Brazil (item5.5). The chapter ends with concluding remarks.

5.1 The literature and the definition of neighborhood

Segmentation of the housing market at an intra-metropolitan or intra-urban scale has been the focus of economists for some time (Dale-Johnson, 1982), specifically after the development of the hedonic price function by Rosen (1974). More recently, increased heterogeneity, due to the spatial dispersion of employment urban sprawl and mixed land-uses (Wheaton, 2004), has made urban economists refer back to sub-markets and neighborhoods to control for locational influences on real estate prices (Durlauf, 2004; Wheaton and Nechayev, 2005). As Bourassa, Cantoni and Hoesli (2007) demonstrate models of real estate prices and development gain in quality when the specification includes explicit sub-markets.

The most common definition of sub-markets describes a sub-market as a "set of dwellings that are reasonably close substitutes for each other, but relatively poor substitutes for dwellings in other submarkets" (Bourassa, Steven C., Hamelink *et al.*, 1999, p. 161). Other scholars emphasize the internal homogeneity of sub-markets (Goodman and Thibodeau, 1998), which led Stevenson to describe sub-markets as "largely distinct communities" (2004, p. 139). Apart from the difference in scale, this is not much different from the classical definition of homogeneous regions in Geography (Castro, Correa *et al.*, 2001).

Although sometimes used interchangeably, it is important to stress the differences between sub-markets and neighborhoods.

Megbolugbe (1996, p. 1787) defines neighborhoods as units: (a) that are homogeneous areas; (b) that possess an identity or social cohesion; (c) in which residences are close substitutes for each other; and (d) that are "small areal units that do not necessarily have any of the above". Whereas the sub-market definition of Grigsby and his followers is similar to (c), the remainder of this section discusses neighborhoods as a concept defined by (a) and (b), which are thought to add to the understanding of real estate markets. This section concludes by establishing our understanding of both terms.

When discussing neighborhoods, urban researchers usually emphasize the concept of self-contained identity-space. This concept was proposed by urban morphologists who built on the work of architect Kevin Lynch (1960) and the subsequent field of Urban Landscape research (Rossi, 1966)[53]. These authors argue that physical and imaginary or symbolic elements[54] (such as highways, rail lines, rivers, squares, corners or monuments) fragment space and allow individuals to identify spatial units.

This implies that perception of space is important in establishing boundaries, especially because larger urban areas are perceived as collages (or wholes), rather than "piecemeal", as described by Barbara Tversky (2003, p. 8). Tversky reinforces Lynch's arguments by arguing that "the critical elements of the space of navigation are landmarks and paths, links and nodes". At the same time, the notion of collages weakens the Cartesian notion of distance and limits it to places in which every pair generates a distance in space independent of the subset it belongs to.

Coulton (2001) argues that perception varies for individuals and that studies that do not explicitly specify the neighborhood as perceived by inhabitants might have biased results. Guest (1983) and Lee (1997), on the other hand, claim that a majority (over 80%) of residents on average agree on locality names, suggesting that they have the same perception of a given neighborhood.

The identity space approach implies that inhabitants perceive neighborhoods as well-defined entities, to which they attach positive and negative attributes (qualities). This follows the work of Galster (2001, p. 2112), who defines neighborhood as a "spatially based [bundle of] attributes associated with clusters of residences, sometimes in conjunction with other land-uses". Based on the work of Golab (1982), Galster (2001, p. 2112) emphasizes the importance of the neighborhood as a socially constructed entity with both physical and symbolic features. In particular, the neighborhood possesses characteristics associated with (a) the residences themselves, such as structural qualities (material, type, design), (b) demographic and social aspects (income and education), (c) the quality of public service offered within, (d) environmental issues, (e) proximity (such as accessibility to shopping or employment centers), and (f) political and symbolic (or sentimental) importance.

Guo (2007), following the earlier work of Suttles (1972), established four scales within which neighborhoods can be defined. The smaller, local scale is the immediate vicinity of one's house, where boundaries are egocentric, fuzzy, and not always agreed upon. The second scale has a "known entity" recognizable by residents and outsiders. A third, even larger scale, is mostly bureaucratically imposed. This would coincide with the fourth "areal unit" definition of Megbolugbe (1996). The largest scale is defined as part of large sections of the city. Although this level of spatial division has been used before (Bourassa, Steven C., Hamelink *et al.*, 1999), it is too large for the detailed proposal of this study.

Given the preceding discussion, neighborhood is defined as follows. As proposed by Megbolugbe (1996), the neighborhood has both a homogeneous trait and social identity. However, residences in the neighborhood are not considered to be close substitutes for each other; rather, *the locational influence the neighborhood has on dwellings* is comparable. In terms of the attributes that are attached to and can be quantifiable by neighborhood, we follow the proposal of the six aforementioned items suggested by Galster (2001).

Based on this definition, the spatial unit to which an urban amenity belongs rather than the distance to other amenities defines the salient characteristics of the amenity. As Galster describes, “it is precisely these *perceptions* of boundaries that are most critical in constructing theories or predictive models of neighborhood change” (2001, p. 2114, original emphasis).

5.2 The spatial unit chosen

Spatially-localized information comes from various sources that present a number of different spatial, methodological and emphasized foci. A full description of the city should highlight what distinct information is present in singular locations. The analysis of the complex urban space needs thorough detailing of its urban tissue that includes socioeconomic factors at a greater scale than is usually informed at the municipal level – that is, it demands analysis at the neighborhood level. Therefore, in accordance with the definition posed in the previous section, this section establishes the spatial unit of reference for empirical analysis.

Various spatial units of analysis have been considered in this study. In Belo Horizonte, the City administrators apply their analysis to nine regional areas called Regionals and 81 planning units (UP). These, however, would fall into the ‘too large’, ‘bureaucratically imposed’ categories discussed above, and are therefore not adequate for our purposes. The 2,564 census tracts used by the national statistical foundation, on the other hand, have abstract boundaries (‘areal units’) that are not recognized by citizens as geographical, legitimate units and are thus also inappropriate.

A more suitable proposal made by a Minas Gerais state’s foundation – Fundação João Pinheiro – is the Human Development Unit (UDH). The UDH is defined as a close approximation of neighborhoods (Pinheiro, 2006) and is defined by the following criteria: (a) a UDH must be an aggregation of census tracts; (b) the UDH must have a minimum sampled population of 400 domiciles; (c) the UDH must be socially, economically and environmentally homogeneous; (d) the UDH must have a recognizable identity expressed by its widely recognized name; and (e) the UDH must be contiguous. These characteristics make UDHs the best approximation of neighborhoods as defined in the previous section. In addition, the UDH definition follows the classification suggested by the United Nations Program for Brazil (PNUD), which would make them comparable to other Brazilian metropolitan areas[55].

Thus, the UDH definition is chosen as the smallest possible scale that can allow for the use of the official data available, and at the same time, a scale that can be considered to be recognized by the population as an entity with a high degree of homogeneity. Note further that given the heterogeneity of urban features observed in the study area[56], these spatial units might vary in size depending on their internal degree of homogeneity. In practice, an area such as a slum, even if extremely close, would never be classified within the same neighborhood as a formal (non-slum) adjacent occupied area.

5.3 Data: describing urban space

This section describes what information is used to characterize the urban fabric (
Figure 5-1). The approach that is proposed aims to capture the central functions of the city as well as the positive and negative amenities that are likely to influence land prices in economic models. First, the indices are presented (this section) and then they are operationalized using principal component analysis (see 5.5). The four indices discussed in this chapter represent only part of the description of urban fabric. The other data not used in the indices are useful for the econometric exercise on chapter 6.

Urban space associated with neighborhoods uses the six main attributes proposed by Galster (2001).

5.3.1 Neighbourhood quality index

The two first attributes of Galster (2001), structural characteristics of the dwellings and demographic and social aspects, are captured by the Neighborhood Quality Index.

This index is based on three existing complementary indices: (a) the housing inadequacy index (HII), calculated using a multivariate technique by a team from the Federal University of Minas Gerais (CEDEPLAR/UFMG) and technicians of the City of Belo Horizonte (Lemos, Ferreira *et al.*, 2004); (b) the Municipal Human Development Index produced by Fundação João Pinheiro, which is affiliated with the State of Minas Gerais (Pinheiro, 2006), and constructed according to directives of the United Nations Program for Brazil (PNUD); and (c) a social occupational category proposed by Mendonça (2002).

The Municipal Human Development Index is calculated as the arithmetic average of the following three indicators: I) education, which is subdivided into two measurements, (i) the rate

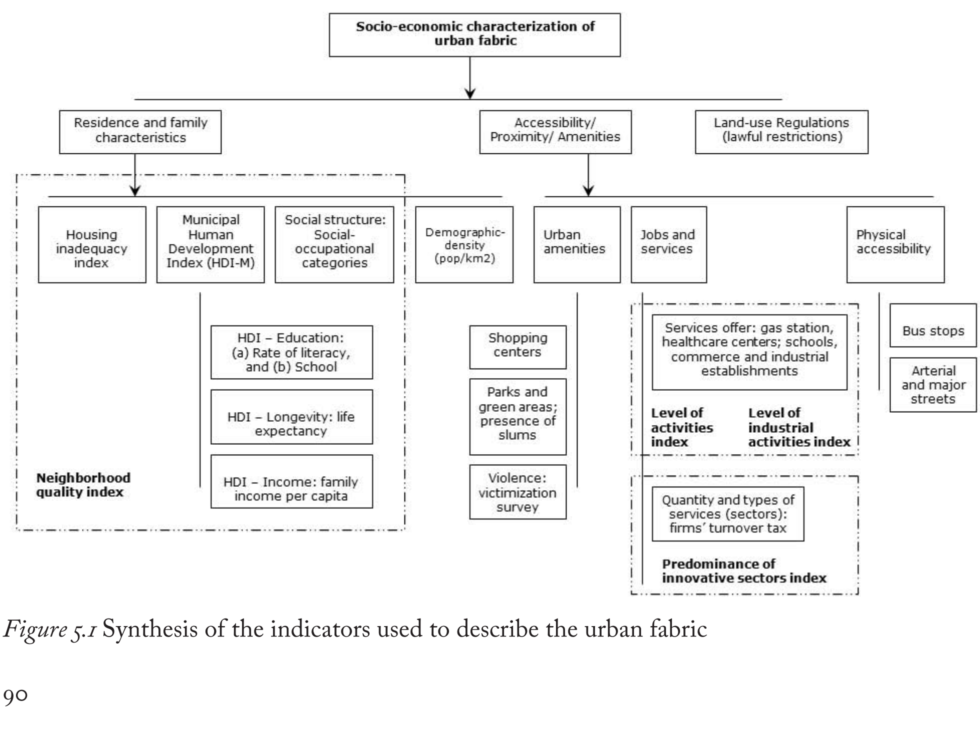


Figure 5.1 Synthesis of the indicators used to describe the urban fabric

of literacy and (ii) the rate of school attendance, II) life expectancy at birth, and III) family income.

The social occupational category proposed by Mendonça (2002) is used to represent social aspects that go beyond income. It brings social stratification to the forefront, including further segregation of parts of the city based on social hierarchy. According to the description provided by Mendonça and Godinho (2003), the following was observed:

> The construction of this social hierarchy is based on the notion of the central importance of jobs on the structuring and functioning of society. The occupational information provided in census questionnaires was the main variable for the analysis of social space. The intention was the construction of a system of social hierarchies that served as a *proxy* for the social structure … the first action was the classical division of classes: those with or without capital (2003, p. 123).

The categories used by the author are hierarchical: superior; average-superior; average; average-inferior; working-men; working-men popular; popular; popular, working-men and farmers; farmers and popular; and polarized.

5.3.2 Level of activities index

The third item proposed by Galster (2001), 'the quality of public service offered', is presented with item (e), 'proximity' and accessibility, in the Level of activities index. In Brazil, services such as healthcare or education are provided both privately and publically. Therefore, they are aggregated in this index, which contains information on the quantity of a number of services offered and commercial establishments. Note that, in contrast to the education systems observed in other countries, there are no rigid school districts in Brazil. Therefore, house location does not necessarily mean that choice of school is restricted.

5.3.3 Level of industrial activities index and Predominance of innovative sector

To further characterize the presence of services, two other indices are calculated: (1) a specific Level of industrial activities index, and (2) an index of the distribution of sectors of activities, with an emphasis on innovation. The latter is measured by a company's turnover tax, which is a proxy for quantification of jobs available in services and presence of diversified commerce. The results show that it is possible to establish six large types of sectors, which were broadly named: (1) heavy industry; (2) innovative sectors; (3) transports; (4) security (5) manufacturing and (6) health-related service. Item (2) is used to construct the Predominance of innovative sectors index, which is thought to complement the other three indices and enhance the proposed detailed description of the city.

5.3.4 Other data

Negative attributes of the urban fabric such as crime and violence, are also thought to be important and are included in the general characterization of urban tissue. Given their greater influence, large pieces of urban infrastructure, especially large shopping centers or university campuses, are treated individually.

The fourth item proposed by Galster (2001), 'environmental issues', is represented by the presence of green areas and water.

The last item mentioned is 'political and symbolic importance'. This item is incorporated when the choice of the spatial unit is made[57].

The above data is available in shape format (ESRI, ArcGIS), and is georeferenced at the same system of coordinates that allow them to be transferred into econometric computer programs and into METRONAMICA simulation software, described in chapter 8.

Together, the four indices proposed capture the idiosyncrasies of each of the 162 neighborhoods in an almost continuous way, because all indices vary from zero to one (see Table 11-3 for quantitative results).

Finally, some variables presented in Figure 5-1 are part of the description of the city, but are not used in the applied econometric exercise proposed in chapter 6 because they did not to add explanatory information to the model applied. First, the influence of bus stops, because they are present in nearly the entire city (considering the 400 meters radius used by transport authorities), is mostly homogeneous for all observations. Second, violence was not added because the available data suggests that both commercial areas and residential areas of varied income-levels have similar perceived violence levels. Third, the quantity of services (ISS) was not added because this item is already present in the index of activities and does not need to be considered twice. Finally, parks and green areas are relatively few and are generally inserted in high-density urban areas. Details of the dataset are available on request.

5.4 Methodology

Data standardization is necessary because all information, in any spatial unit originally provided (neighborhoods, census tract, UDHs), must be transformed to comparable (UDHs) units so that it can be used in the econometrics exercise in chapter 6. To do this, principal components analysis was used (see a brief description in Appendix 11.3). The use of principal components serves the purpose of summarizing the information and in the process standardizing the data.

5.5 Results

The application of principal components to the dataset enabled the construction of four indices that were applied in the econometric exercise. This section presents and comments on the results.

Table 5.1 Matrix of correlation for the neighborhood quality index

	Socio-occupational categories (inv.)	**HDI-Income**	**HDI-Longev.**	**HDI-Educ.**	**Housing Inadequacy Index**
Socio-occupational categories (inverse)	1.00	-0.82	-0.71	-0.80	0.67
HDI-Income	-0.82	1.00	0.88	0.95	-0.72
HDI-Longev.	-0.71	0.88	1.00	0.89	-0.71
HDI-Educ.	-0.80	0.95	0.89	1.00	-0.70
Housing Inadequacy Index	0.67	-0.72	-0.71	-0.70	1.00

5.5.1 Neighborhood quality index

The index of neighborhood quality index (NQI) aims to integrate aspects related to the family and the domicile. The NQI is the result of the aggregation of the Housing Inadequacy Index, the municipal human development index, and socio-occupational categories.

The socio-occupational categories are arranged from 1 to 10, with a lower number representing the highest social status, so that in terms of interpretation the smaller absolute value represents a higher hierarchical position. The Housing Inadequacy Index, like the name suggests, is organized so that the higher the value, the greater the inadequacy of the household. Interpretation should be performed with cautiously, considering that lower values represent better housing attributes.

Table 5-1 presents the correlation matrix of the original indices. It appears that, as expected, the lower the socio-occupational category, the lower the income, education, longevity, and quality of the dwelling.

The results of PCA (Table 5-2) suggest that the first component – expressed by its eigenvalues – explains more than 80% of the total variance of the dataset. This component loads high on social-occupational characteristics and HII, suggesting that these two variables explain most of the variation in all variables. The table of final scores for the UDHs (see the annexes, Table 11-3) was normalized to values between 0 and 1, and the scale was reversed to facilitate interpretation

The visual analysis of the first component – the Neighborhood quality index (Figure 5-2) – confirms empirical knowledge that the central and south-central regions of the city contain the best neighborhoods. It also shows large patches of low values that represent old, central slum settlements. Furthermore, visual analysis confirms the heterogeneity and rigidities of urban tissue. The outskirts of the municipality, in general, in the southern, northern, and eastern areas present lower values than more central neighborhoods.

5.5.2 Level of activities and Level of industrial activities indices

The analysis made using information on the number of establishments for industry, commerce, service and autonomous work, added to other quantitative data enabled the construction of two other indices (Table 5-3 and Table 5-4). The first is called **Level of activities,** and it is the

Table 5.2 Eigenvalues and eigenvectors for the Neighborhood quality index

	Eigenvalues				
Absolute value	4.1938	0.8014	0.0043	0.0005	0.0001
Percentage	0.839	0.160	0.001	0.000	0.000
Eigenvectors	Neighborhood quality index	2nd	3rd	4th	5th
Socio-occupational categories (inverse)	0.772	0.635	0.044	-0.005	0.002
IDH-Income	-0.045	-0.007	0.858	-0.449	-0.244
IDH-Longev.	-0.020	0.003	0.407	0.890	-0.203
IDH-Educ.	-0.017	-0.003	0.308	0.075	0.948
Housing Inadequacy Index	0.634	-0.773	0.028	0.004	-0.001

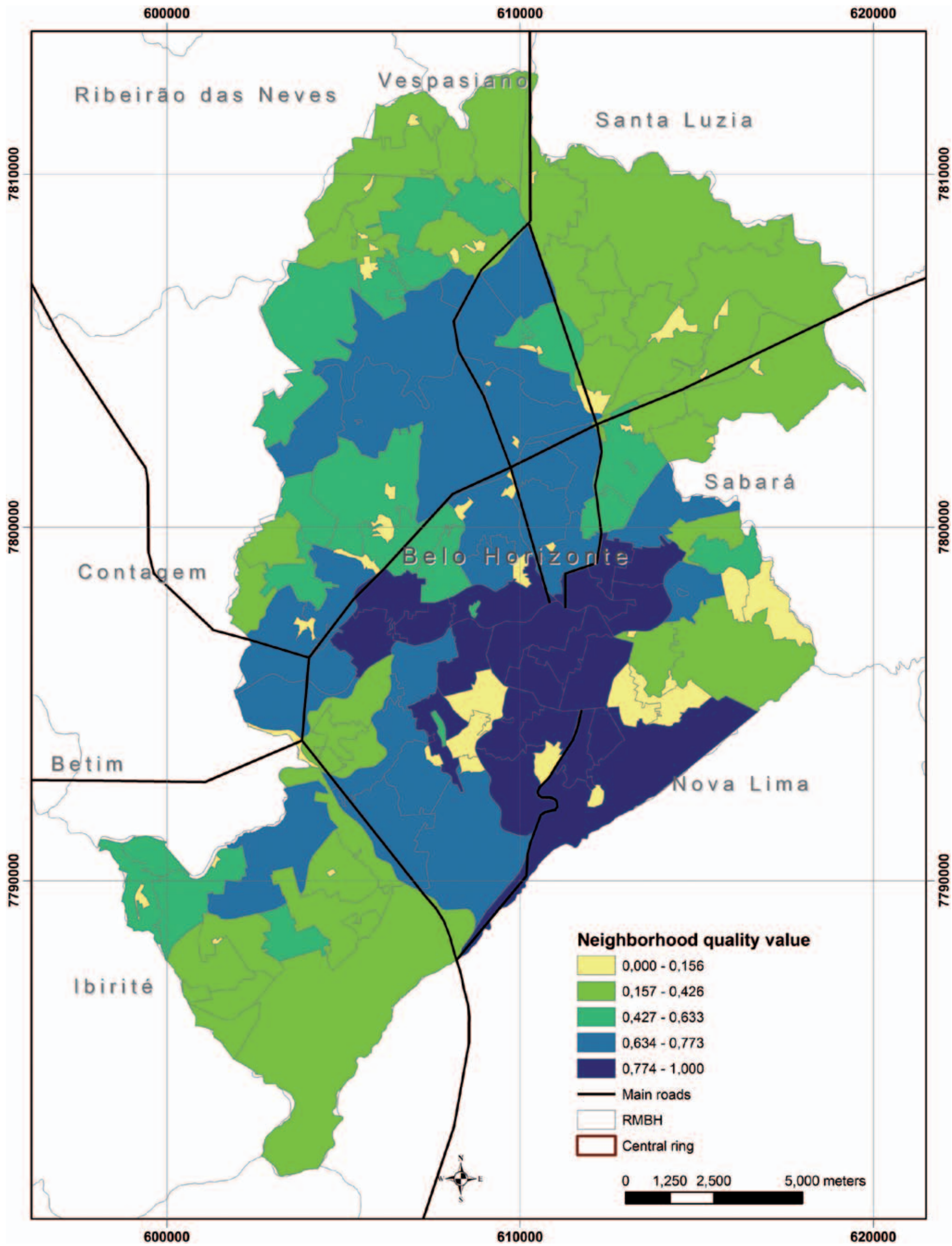


Figure 5.2 Neighborhood quality index for Belo Horizonte

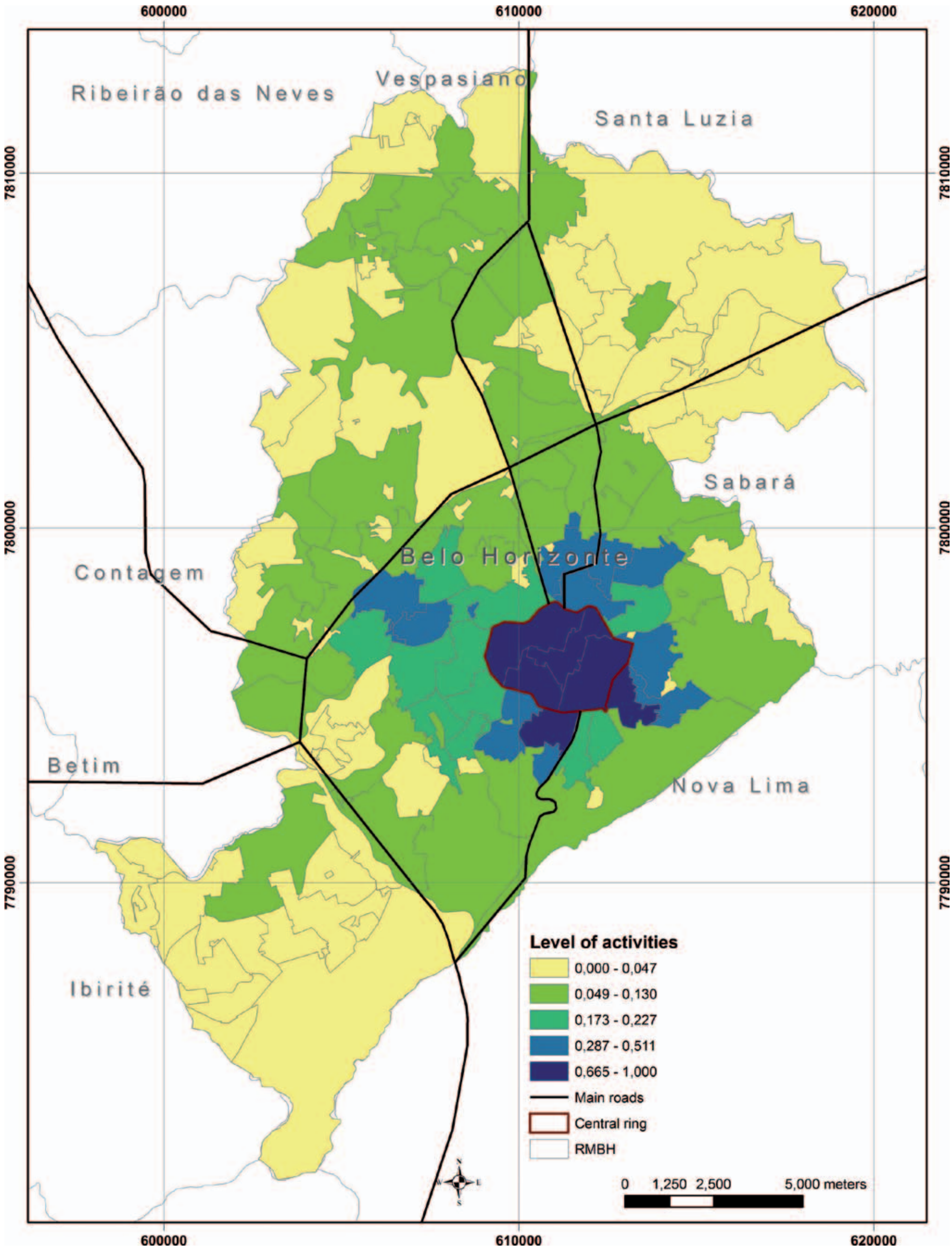


Figure 5.3 Activities index for Belo Horizonte

Table 5.3 Correlation matrix for variables used in the construction of the level of activities and industrial activities indices

	Number of bakeries	**Gas stations**	**Schools**	**Local health centers**	**Industries**	**Bank agencies**	**Commerce establishment**	**Autonomous estab.**	**Service estab.**
Number of bakeries	1.000	0.603	0.610	-0.329	0.596	0.663	0.541	0.570	0.590
Gas stations	0.603	1.000	0.321	-0.248	0.387	0.466	0.321	0.331	0.360
Schools	0.610	0.321	1.000	0.018	0.404	0.770	0.452	0.430	0.466
Local health centers	-0.329	-0.248	0.018	1.000	-0.256	-0.163	-0.208	-0.229	-0.226
Industries	0.596	0.387	0.404	-0.256	1.000	0.539	0.884	0.860	0.886
Bank agencies	0.663	0.466	0.770	-0.163	0.539	1.000	0.625	0.602	0.656
Commerce establishment	0.541	0.321	0.452	-0.208	0.884	0.625	1.000	0.884	0.959
Autonomous estab.	0.570	0.331	0.430	-0.229	0.860	0.602	0.884	1.000	0.961
Service estab.	0.590	0.360	0.466	-0.226	0.886	0.656	0.959	0.961	1.000

Table 5.4 Eigenvalues and eigenvectors for level of activities and industrial activities indices

				Eigenvalues					
Absolute value	1955827,1214	38638,5270	8364,7016	449,8123	88,9286	11,5179	6,0980	1,8797	1,0223
Percentage	217314,1246	4293,1697	929,4113	49,9791	9,8810	1,2798	0,6776	0,2089	0,1136
				Eigenvectors					
Variables/ Components	**Activities level**	**2nd comp.**	**3rd comp.**	**Industrial activities**	**4th comp.**	**5th comp.**	**6th comp.**	**7th comp.**	**8th comp.**
Number of bakeries	0.0015	0.0016	-0.0033	0.0331	0.1649	0.1460	0.8789	-0.3787	0.1860
Gas stations	0.0005	0.0005	-0.0031	0.0182	0.0635	-0.0606	0.3973	0.9128	-0.0305
Schools	0.0019	-0.0007	-0.0049	-0.0052	0.4453	0.8685	-0.1644	0.0948	-0.1063
Local health centers	-0.0002	-0.0003	-0.0001	-0.0070	-0.0011	0.0755	-0.1704	0.1119	0.9761
Industries	0.0310	-0.0115	0.0524	0.9965	0.0308	-0.0170	-0.0425	-0.0044	0.0016
Bank agencies	0.0053	-0.0003	-0.0195	-0.0390	0.8770	-0.4633	-0.1093	-0.0427	0.0224
Commerce establishment	0.4622	-0.7825	0.4147	-0.0452	0.0041	-0.0010	0.0035	0.0006	0.0002
Autonomous estab.	0.3272	0.5874	0.7389	-0.0426	0.0118	-0.0023	0.0014	0.0013	0.0003
Service estab.	0.8236	0.2062	-0.5281	0.0050	-0.0152	0.0028	-0.0017	-0.0005	-0.0003

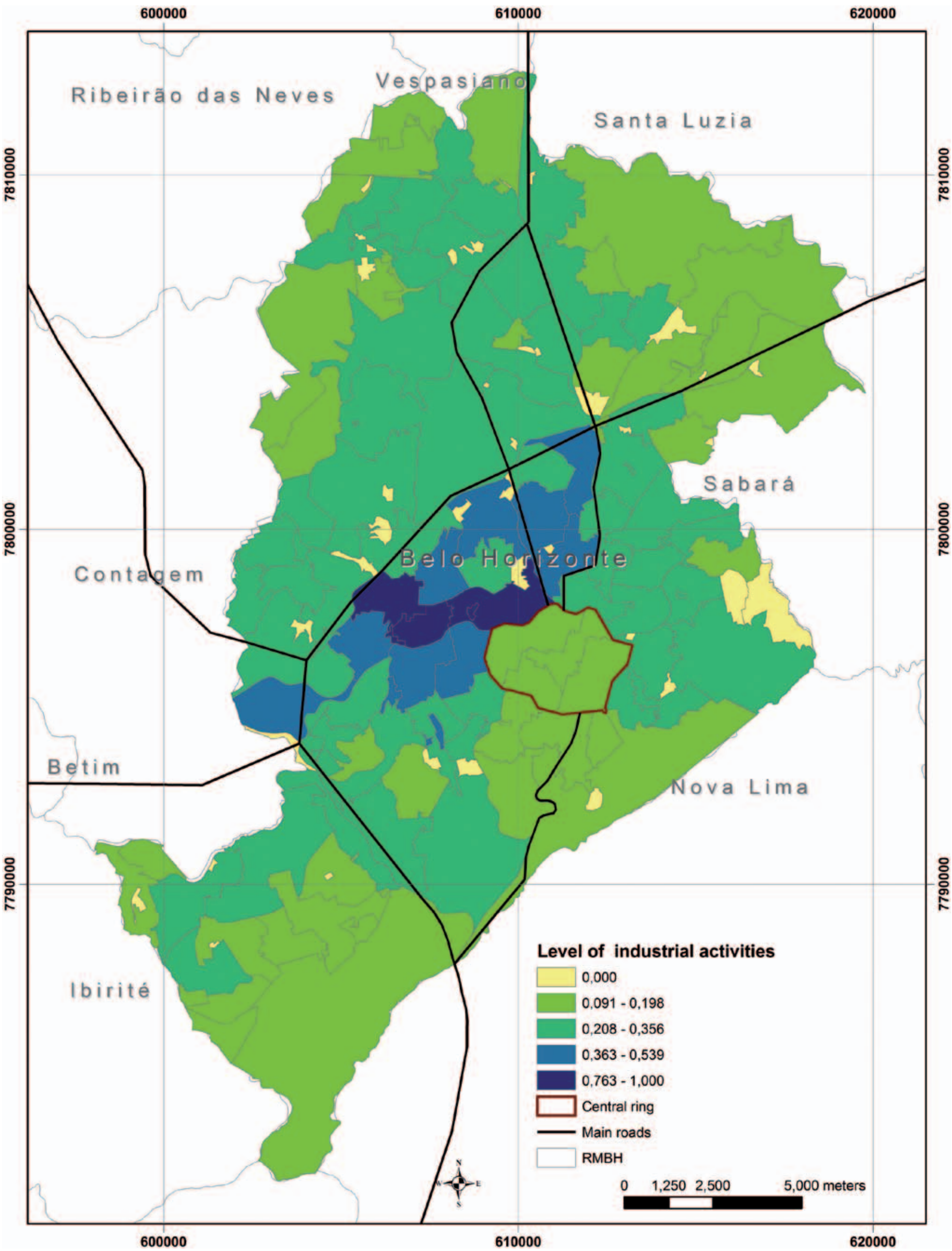


Figure 5.4 Industrial activities index for Belo Horizonte

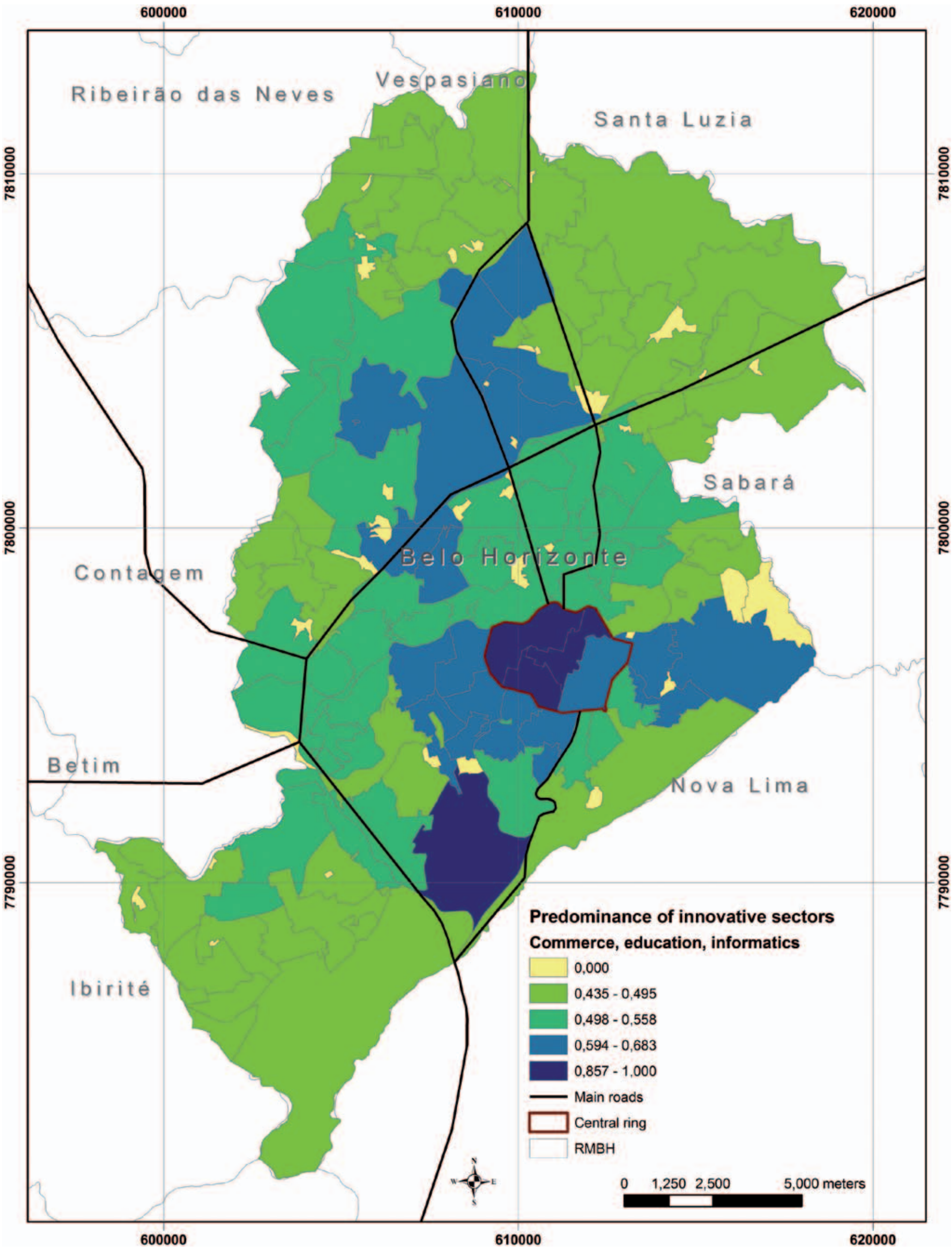


Figure 5.5 Predominance of innovative sectors

Table 5.5 Eigenvalues and eigenvectors for innovative sectors index 104

Eigenvalues						
Absolute value	3.593	0.688	0.524	0.448	0.406	0.340
Accumulated percentage	0.493	0.587	0.659	0.721	0.776	0.823

Eigenvectors						
Sectors/ Components	**Services/ Heavy industry**	**Innovation, commerce and education**	**Transports**	**Security**	**Industries**	**Health services**
Commerce in general	0.145	0.342	-0.104	0.044	-0.077	0.255
Cientific and technological institutions	0.080	0.273	-0.042	-0.098	-0.203	0.225
Publicity and propaganda	0.203	0.272	-0.052	-0.120	0.021	-0.113
Informatics and computering	0.206	0.241	-0.129	-0.024	0.037	-0.224
Automobiles services	0.131	0.235	0.381	0.087	-0.081	0.142
Private education	0.217	0.222	-0.097	-0.110	-0.009	-0.053
Agriculture	0.197	0.132	-0.037	0.191	0.134	-0.364
Entertainment	0.233	0.127	-0.008	-0.047	0.015	-0.001
Consulting businesses	0.235	0.125	-0.044	-0.091	-0.087	0.072
Air transport	0.007	0.107	0.593	-0.008	-0.149	-0.126
Security services	0.028	0.107	-0.031	0.490	0.128	0.215
Health-related services	0.016	0.093	-0.009	0.288	-0.007	0.441
Metalurgical industry	0.007	0.073	0.199	-0.208	0.324	0.056
Transport material industry	0.015	0.069	0.584	-0.017	-0.170	-0.135
Editing and publishing industry	0.197	0.064	-0.014	0.178	0.061	0.064
Mechanical industry	0.044	0.055	0.133	-0.233	0.564	0.107
Personal services	0.017	0.055	-0.009	0.444	0.246	-0.466
Public security	0.248	0.027	-0.023	-0.014	-0.002	0.014
Rubber and plastic industry	-0.003	0.021	0.003	0.144	0.092	0.108
Telecommunications	0.243	0.014	-0.030	-0.006	-0.007	0.032
Eletric and telecommunications material industry	-0.015	-0.014	0.077	-0.240	0.543	0.131
Lodging	0.241	-0.052	-0.049	-0.064	-0.039	0.008
Real estate management	0.247	-0.058	-0.024	-0.078	-0.008	-0.022
Construction	0.243	-0.078	0.045	-0.036	-0.028	0.056
Transports	0.138	-0.088	0.176	0.352	0.218	0.225
Private insurance	0.243	-0.113	-0.029	-0.078	-0.023	-0.031
Eletronical and hospital equipment industry	0.241	-0.117	-0.026	-0.080	-0.023	-0.022
Banks and financial institutions	0.240	-0.150	-0.003	-0.062	-0.029	-0.020
Television and radio	0.210	-0.197	0.030	0.150	0.054	-0.189
Public education	0.213	-0.280	0.045	-0.023	-0.013	0.066
Garbage collecting industries	0.167	-0.369	0.075	0.046	-0.053	0.122
Non-metalic transformation industry	0.157	-0.384	0.073	0.029	-0.052	0.095

first component of another principal components analysis. It explains nearly all the variance in the dataset. The second is called **Level of Industrial activities** and it is the fourth principal component. The explanatory value of this index relies on its ability to highlight observed areas that are more exclusively industrial. Industrial areas (with low-levels of commerce and services) are usually conspicuous to the population, and therefore, negative externalities more tangible. Thus, this index provides additional information because it focuses on the exclusive presence of industries, in comparison to other activities.

The level of activities well represents the notion of centrality, which is largely discussed within urban and regional economics (Figure 5-3). The availability of services and commerce is higher mainly in the central ring. Other sub-centers in the south and in the north are also represented in the results, although at a proportionally lower level.

The results for the Level of industrial activities index show that both peripheral regions north of the municipality and high-income residential areas in the south exhibit low values for this index. The neighborhoods classified at high levels share a perceived image of traditional, industrial areas.

5.5.3 Predominance of innovative sectors

The results for the predominance of innovative sectors also come from the information of total taxes paid from the different 32 sectors and were provided by the Finance Department of the City of Belo Horizonte (SMF/PBH). They represent the second principal component, which explains about 9% of the total variance (Table 5-5). The index concentrates service sectors, which are innovative, with higher positive weight on the coefficients (eigenvectors) and negative weight on sectors of heavy manufacturing or less sophisticated businesses.

This index also complements information that is not present in the other indices produced (Figure 5-5).

5.6 Concluding remarks

This chapter aims to provide a comprehensive description of urban fabric, along with a conceptually strong definition of the term neighborhood and its associated attributes. Consequently, this chapter allows for a coupling of urban space in economic modeling exercises that is richer in detail and more adequate in terms of perception. This is done with an analysis of the literature on neighborhoods from the social sciences, associated with a widely-used and simple multivariate analysis. Thus, this chapter makes procedures easy to replicate, and enables urban analysts or social scientists to discuss the city quantitatively and qualitatively in a more informed way.

The empirical results presented herein showed that the set of four indices provide a description of space that is nearly continuous. One neighborhood might be classified at a similar level to another in one index, but at a totally different level in another, so that taken together the composite classification according to these indices is always unique. To further validate this approach, the results are successfully used in an exercise of spatial econometrics in chapter 6.

The use of dummy variables, widely used in economics, also guarantees this uniqueness. However, the added-value of this approach is that, in addition to singling out the particularities of a certain spatial unit, the modeler can also explain the way in which a specific neighborhood is different from another by simply observing the results from the scale of indices. The use of dummy variables in this empirical analysis, in which there are 162 different units, would make the analysis rather tiresome and confusing.

In general, this chapter shows that quantitative analysis should help enhance the qualitative, and well-defined, discussion of urban space, rather than replace it. We believe that urban economists can gain when including cognitive and perceptive essential facts into their analysis, instead of using only abstract, ad hoc distances. Urban studies can also benefit from data widely available on precise locations within a given urban space.

6 Analyzing the neighborhood hedonic prices function using spatial and quantile regression modeling[58]

6.1 Introduction

Urban real estate price determination has been subject of extensive theoretical and empirical research for quite some time (Sheppard, 1999). Advances in modeling have accordingly emerged (Capello and Nijkamp, 2004a). In the late 1990s, previous work on spatial effects (by notably Cliff and Ord, 1973; and Paelinck and Klaassen, 1979) was systematically organized and it became clear that the inclusion of spatial dependence is a crucial precondition for empirical hedonic price studies (Anselin, 1988). In practice, this has resulted in two types of model specifications: (a) specifications considering distance to a Central Business District (CBD) and to other points of interest, such as concentrations of work as the main spatial determinant; and (b) those that apply weight matrices in its most common form of spatial contiguity. We argue in this chapter that both conceptualizations do not capture the full influence of urban location on prices. The empirical research approach used in this chapter intends to shed light on the behavior of the real estate market, in which we specifically considered neighborhood identity and influences of urban tissue separate from influences of the estate itself. We also observe differences in behavior correlated with prices of estates.

First, in an increasingly mobile (Muhammad, 2007), decentralized (Garreau, 1992) and mixed-use (Wheaton, 2004) pattern of urban occupation, distance to CBD loses explanatory power. Moreover, the scale of analysis hampers the specification of what exactly the center of a large metropolis is, making definition arbitrary[59]. Also, proximity relationships may be more relevant to price forming than the distance to the center.

Second, contiguity or distance matrices usually applied in spatial econometric analyses can also be considered abstract and *ad hoc* in nature. The contiguity matrix usually uses Delaunay triangulation to determine the neighbors of a given locational observation. The distance matrix captures a radius of a certain distance (cut-off point) chosen by the researcher to define individuals within reach as neighbors. This procedure might inaccurately include an observation on the other side of a river or a highway as neighbors.

One way to deal with these issues is to use landscape theory, as proposed by the architects Kevin Lynch (1960) and Aldo Rossi (1966). They argue that the perception of the city by its citizens is spatially circumscribed within a known entity. A person relates positive and negative qualities (attributes) to certain spatial portions of the territory that are collectively recognized by a community in urban space (Tversky, 2003) (see chapter 5). This conceptualization is in accordance with sales practices of real estate agents, who specify the neighborhood when advertising a sale. Accordingly, the approach hypothesizes that if the attributes are relevant, the neighbourhood as a whole is positively (or negatively) valued. In this chapter, we treat the neighborhood as a reference for analyzing the determinants of prices of houses. In our analysis, the configuration of neighbors – given by the weight matrix – explicitly considers social identity and cohesion in the choice of who is considered neighbors.

Our analysis applies other methodological improvements over previous studies as well. First, the determinants of housing prices are captured by a limited number of explanatory dimensions (factors) instead of numerous individual variables. Similar variables of the same dimension are weighed together. In addition to contributing to a better interpretation, this is also helpful in dealing with problems of multicollinearity that arise from the complex nature of urban space and the city when applying regression techniques. Principal component analysis (PCA) is used to construct indices of values that summarize the multiplicity of explanatory factors for each neighborhood.

Second, a homogenous treatment of the population of houses and their prices disguises important population-internal heterogeneity. It therefore has become essential in real estate analysis that different patterns of behavior and preferences for various levels of social characteristics are dealt with. As households at different positions on the social ladder might value a single attribute (such as proximity to a large shopping center) differently, segmentation of households should be internalized in the analysis. Methodologically speaking, we apply a quantile regression estimation (Koenker and Bassett, 1978; Buchinsky, 1997). This helps us not miss variation in responses to variables that are present in the full sample (Wooldridge, 2002). The initial study of Zietz (2008) applies this estimation technique, but it is introduced coincident to a spatially-lagged variable in the model rather than in a full spatial-quantile estimation model, such as in the model proposed by Su and Yang (2007). According to Anselin (2002), that procedure would be reasonable only when one considers spatial autocorrelation to be a nuisance. In our analysis, we want to incorporate spatial autocorrelation and quantile regression simultaneously.

Finally, this chapter also aims to detail extent to which the inclusion of social heterogeneity and cognitively perceived neighborhoods impact real estate prices. Specifically, we want to test the influence of the following factors:

(a) In relation to the surface area of the estate:
    i. Is the positive influence of the estates' surface on prices constant throughout all social levels and neighborhood matrices?
    ii. Regarding the quality of the estate: for whom is it more important?
    iii. Does the estate's age influence housing prices negatively and uniformly?
(b) In relation to the neighborhood:
    i. How relevant is the quality of the neighborhood?
    ii. Is the presence of commerce and services always a positive influence on housing prices?
    iii. How negative are industry-heavy neighborhoods?
    iv. Do neighborhoods with innovative services have a positive influence on housing prices?
(c) In relation to proximity:
    i. Do large shopping malls, proximity to avenues, and arborization all have a positive influence?
    ii. How does proximity to slums affect real estate prices?

6.2 Methodology

6.2.1 Spatial analysis and spatial econometric models

The essential concept to understand when working with spatial analysis is spatial dependence. It is rooted in what is commonly called the First Law of Geography of Tobler according to which "everything is related to everything else, but near things are more related than distant things" (Tobler, 1970, p. 236). Spatial autocorrelation, in turn, is the statistical expression that measures this spatial dependence.

Concerning spatial statistical dependence, Anselin (2005) suggests that one first implements traditional models, tests for the presence of spatial autocorrelation, and then, if necessary, uses a spatial model.

The first indication of the presence of spatial dependence is Moran's I statistic:

Equation 36 – Moran's I

$$I = \frac{\sum_{j=1}^{n} \sum_{i=1}^{n} w_{ij} \left(x_i - \bar{x}\right)\left(x_j - \bar{x}\right)}{\sum_{i=1}^{n} \left(x_i - \bar{x}\right)^2},$$

where W_{ij} is the weights matrix that reflects spatial dependence in the formula; n is the number of variables; x_i is the variable of interest.

The weights matrix can be constructed in a number of ways. These are crucial to the analysis and are discussed in the next section.

However, although Moran's I index is adequate for indicating the presence of spatial autocorrelation, it does not show in the form in which it occurs (Anselin, 2005, p. 197).

Spatial models should be applied when theoretical indications related to the research problem suggest that there is spatial influence among observations. In this chapter, for instance, spatial-dependent modeling is considered because the price, quality and availability of services of nearby dwellings influence the values of the dwelling. According to the simple model of spatial process identification proposed by Anselin (2005, p. 199), if there is no detected spatial influence, one can stick to traditional models with more guarantees of their robustness.

One of the most cited econometric spatial models in the literature is the spatial lag, which includes a term of influence of neighbors in the regression process (Lesage, 1999):

Equation 37 – Spatial lag model

$$Y = \rho Wy + X\beta + \varepsilon,$$

where Y is the dependent variable; Wy is a spatial lag vector, derived from the choice of spatial weights matrix[60]; ρ is a spatial auto-regressive coefficient; X is the matrix of independent variables and their coefficients β; and ε is the error vector.

Another very common model in spatial econometrics is the spatial error model[61] which is capable of capturing a "spatial effect that has not been included in the model" (Almeida, 2004, p. 61). When this occurs, the spatial effect implicit in the problem has not been totally included and is thus captured by the error.

Equation 38 – Spatial error model
$Y = X\beta + u$, and

Equation 39
$u = \lambda W u + \varepsilon$, where

λ is the error auto-regressive spatial parameter[62]. A full spatial model includes spatial autocorrelation as a variable and in the errors. Furthermore, it is a combination of the two presented models (Lesage, 1999).

Equation 40
$Y = \rho W y + X\beta + u$, and
$u = \lambda W u + \varepsilon$.

The tests recommended for identifying spatial models are: (a) the Lagrange Multiplier tests for Lags, for lag-spatial models, and (b) the Lagrange Multiplier test for Error specifications to indicate the use of an error auto-regressive model. Finally, there is also the LM-SARMA test that indicates the need for a more complex model in which there is spatial autocorrelation in both the variable(s) and the errors (equation 40).

Ordinary Least Squares (OLS) estimations are not appropriate for treating spatial analysis because the literature indicates that in the presence of spatial autocorrelation, OLS results are inconsistent (Lesage, 1999). The suggested alternative is to use a Maximum Likelihood (ML) estimator to estimate a parameter that would most likely be generated in the observed sample[63] (Anselin, 1988; Anselin and Bera, 1998; Smirnov and Anselin, 2001; Anselin, 2005).

6.2.2 Weight matrices

Weight matrices play an essential role in spatial analysis, represent the influence of neighbors in each observation, and account for how this influence diminishes in space. Many authors insist that different weight matrices may alter results significantly and so it is essential to choose a matrix that is theoretically appropriate. Also, robustness tests for different specifications of weight matrices are recommended (Anselin, 1988; 1999; Lesage, 1999; Arbia, 2006).

The simplest models of weight matrices are models of spatial contiguity that indicate, through a binary relation, whether one is a neighbor. There are two basic types: the *queen* matrix that considers as neighbors all individuals which share a border or have at least a vertex in common; and the *rook* matrix in which only individuals who have a common border are considered to be neighbors. In both cases, it is possible to include higher order matrices so neighbors of neighbors might be considered to influence 'own' issues. A strong underlying theory of the phenomenon being studied should indicate the most appropriate choice.

Alternatively, there are matrices built on spatial distance, in which an arbitrary limit is specified and individuals within that limit are considered to be related and dependent on each other. This specification is naturally preferred, which is indicated in cases where the influence can be readily measured[64]. There is also the matrix of the k-nearest neighbors in which neighbors – as the name suggests – are defined as individuals who are closest to the observation. Once more, this number *k* is often set in a rather arbitrary manner. In this chapter we apply three different matrices, including one that captures a newly proposed neighborhood approach (see section 6.4, below).

6.2.3 Quantile regression analysis

Quantile regression analysis[65] is based on points taken in regular intervals from the accumulated distribution function[66] of a stochastic variable. The intervals denote sub-groups of the sample that can be of any order (a hundred, for centis). In this chapter, the reference is the one proposed by Buchinsky (1997), who reinforces the need to verify whether the patterns observed in a traditional regression (OLS) repeat themselves in the quantile regression. If not, the quantile analysis adds valuable information to interpretation of the phenomenon. Moreover, it helps determine whether the studied behavior at higher/lower levels of the dependent variable differs. The quantile regression is not particularly sensitive to the presence of outliers (like OLS is) and it enables change analysis "at different points of the distribution". Buchinsky (1997, p. 90) summarizes the relevance of quantile analysis, stating that "clearly, it is not enough to investigate changes in the mean when the entire shape of the distribution changes dramatically".

Wooldridge (2002, p. 367) recommends quantile regression when the sample is heterogeneous:

> Median regression is a special case of quantile regression, where we model quantiles in the distribution of y given x. For example, in addition to the median, we can estimate how the first and third quartiles in the distribution of y given x change with x. Except for the median (which leads to Least Absolute Deviations), the objective function that identifies a conditional quantile is asymmetric about zero.

6.2.4 Spatial-quantile regression analysis

A recent theoretical advance to make a spatial and quantile model was proposed by Su and Yang (2007). Starting from a traditional quantile model,

Equation 41

$$Y_n = X_n \beta_{o\tau} + u_n,$$

where β0τ is a quantile regressor that can vary with the value of τ. They propose that

Equation 42

$$Y_n = \rho o \tau W_n Y_n + Xn\beta o\tau + U_n,$$

where $\beta_{0\tau}$ is a scalar spatial-lag parameter that varies according to τ (Su and Yang, 2007, p. 4)[67].

The advantages of using this model is that one can consider different degrees of spatial dependence on the regressor at different points of the distribution.

However, Su and Yang (2007, p. 14) note that given the lack of a distributional assumption, inferences based on the results would demand a method for the calculation of the variance-covariance matrix that would "complicate the matter to a great deal". A log-likelihood is not calculated for the same reason. In addition, calculation of an R-squared would make no sense when estimating using instrumental variables (Wooldridge, 2003, p. 471).

Su and Yang (2007) show through Monte Carlo experiments that the instrumental variabe quantile estimation of spatial autoregressive models perform better than both the quasi maximum likelihood (QML) and the generalized method of moment (GMM). They also show that it is robust to the presence of outliers and heteroskedasticity. However, in terms of comparison with spatial and quantile models, one can only use the coefficient values.

6.3 Dataset

The dataset[68] contains information on actual transaction prices for all 5,512 observations that were negotiated from June to August 2007 in the city of Belo Horizonte. The transactions are linked to a real estate city tax database that includes information on: (a) the property value for tax purposes (which may differ from actual transaction prices); (b) the year of construction of the dwelling; (c) the fraction of parcels ownership[69]; (d) finishing and estate quality levels; (e) the real estate use type (see Table 6-2); (f) a depreciation factor, which considers both quality of estate and years of construction; (g) a typology factor, which considers whether the original building has been subject to usage adaptations; (h) a commercialization factor, which considers whether the estate has been subjected to a change in tax value; (i) the average square meter price of construction given building quality; (j) the built surface; (k) a topography factor; (l) the level of public infrastructure; (m) a pedology factor; (n) a parcel position factor; (o) the land square meter price; (q) the parcel size; (r) the City tax value; and (s) the property's geographic coordinates[70].

Table 6.1 Descriptive basic statistics of estate's sample. Source: GEAVI/PBH, 2007

Variable	Observations	Minimum	Maximum	Average	Standard-deviation
Transaction value (€)	5512	493	1401899	50934	62999
Fraction of parcels	5512	0	1	0.39	0.40
Built surface (sq. m.)	5512	2	1905	145	110
Annual city tax value (€)	5512	0	8126	694	391
Value of sq. m. construction (€)	5512	0	595	159	97
Value of sq. m. of land (€)	5512	8	1061	86	94
Value of sale based city tax (€)	5512	398	703608	34708	38508
Transaction value by surface (€/sq.m.)	5512	44	13202	362	361
Year of construction	5512	1924	2006	1986	14.5
Latitude (UTM)	5512	7785573	7811839	7798245	5015
Longitude (UTM)	5512	598612	617013	609021	3247

Table 6.2 Number of observations by estate type. Source: GEAVI/PBH, 2007

Typology	Observations	Percentage
Apartment	3233	0.587
House	1588	0.288
Shop unit	301	0.055
Low-quality house	182	0.033
Industry/working shed	113	0.021
Second floor shop unit	40	0.007
Comercial garage unit	28	0.005
Residencial garage unit	27	0.005
Total	5512	1.000

Table 6.3 Number of observations by construction quality standards. Source: GEAVI/PBH, 2007

Construction quality	Observations	Percentage
P1 (lowest)	309	0.06
P2	1330	0.24
P3	3160	0.57
P4	634	0.12
P5 (highest)	79	0.01
Total	5512	1.00

Descriptive statistics are presented in Table 6-1. The average price of commercialized estates during the period is around 51,000 Euros[71], with an average built surface of approximately 150 square meters an average age of twenty years. The properties are spatially well distributed within the municipality (Figure 6-1).

The apartment unit typology is the dominant type in sample, accounting for almost 60% of all observations, followed by single family housing units (Table 6-2).

The description of how the City calculates the 'construction quality index' is provided in the annex of "Decreto 10925 of 2001". It proposes five levels of quality, P1 being the worst and P5 being the best, depending on the items present in each unit. The presence of items of quality (such as internet cable availability, window material, or external finishing) sums to certain levels that classify the standard for that estate. Different items have different valuations depending on the type of estate (house, apartment, shop). For low-quality housing, for instance, the maximum standard is P3 (average quality). The building standard dominant in the sample, representing more than half of observations, is the P3 standard (Table 6-3). In cumulative terms, 86.5% of estates are classified at P3 or lower and only 13.5% of the sample is of higher quality (P4 or P5).

Table 6-4 shows the spatial units (as defined in chapter 5) that have the highest number of transactions. The 10 spatial units with the highest concentration of transactions account for nearly one third of the total number of transactions (Table 6-4), and 38 spatial units (out of 161)

Table 6.4 Spatial units with highest number of transaction. Source: author's elaboration based on GEAVI/PBH, 2007 and FJP, 2007.

Top 10 spatial units in number of transactions	Observations	Percentage
BURITIS/ESTORIL/MANSÕES	268	4.81%
PARQUE URSOLINA DE MELO/CASTELO/RECREIO/ITAMARATI	268	4.81%
JD. ATLÂNTICO/STA. AMÉLIA/CONJUNTO STA. MÔNICA	166	2.98%
SÃO PEDRO/SANTO ANTÔNIO	165	2.96%
GRAJAU/GUTIERREZ	153	2.74%
BARROCA/PRADO	153	2.74%
DONA CLARA/JARDIM SANTA BRANCA/ ITAPOÃ	141	2.53%
CIDADE NOVA/SILVEIRA	139	2.49%
CARMO/SION	132	2.37%
FUNCIONÁRIOS	130	2.33%
General total	5512	100.00%
General average of observations per neighborhood	34.86	0.63%
General standard-deviation	49.22	0.88%

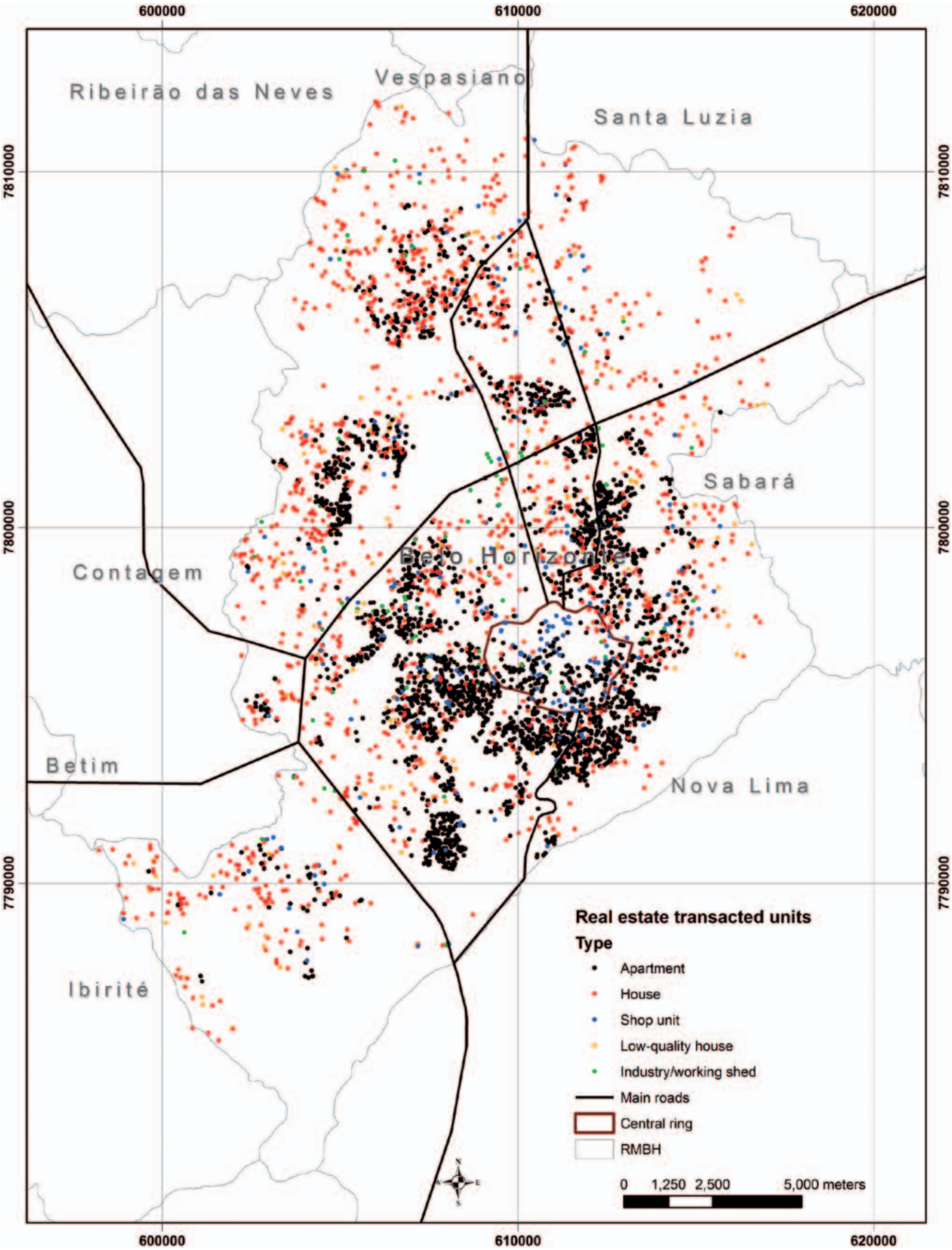


Figure 6.1 Commercialized real estate units in Belo Horizonte, 2007, by type Source: author's elaboration based on data provided by GEAVI/PBH, 2007. These data are linked to the data from chapter 5 to provide the full dataset for the econometrics exercise.

did not include any commercialized property within the analyzed period. The data show that the Castelo and Buritis neighborhoods have the highest rate of transactions. Traditional areas, such as São Pedro, Santo Antônio, Itapoã, Dona Clara, Sion and Funcionários, also have a high number of observations.

6.3.1 Dataset used in the model

The dataset actually implemented in the models is related to each dwelling or to the neighborhood where it is located. In terms of the dwelling, we have: (a) Neperian logarithm of the price of the transaction that actually occurred (*Lnprice*); (b) Neperian logarithm of transaction value per surface (in square meters); (c) Neperian logarithm of estate's surface; (d) construction quality level, as described above; (e) age of the dwelling, measured in years since construction and three *dummy* variables to indicate whether the observation is an apartment, a house, a shop or none of the above.

In terms of neighborhood, we have the following variables derived from chapter 5. The first variable (*domic_n*) represents the neighborhood in terms of the average quality of the housing building and the families who inhabit them, in terms of income and education level (Figure 5-2,). The second variable (*Pca_activ*) reflects quantitative information on the number of economic establishments, specifically the availability of services such as bakeries, gas stations, banks, and private offices. Both a high-class area and a relatively poor neighborhood might have low values for this variable (Figure 5-3). The third variable (*pca_ind_n*) reflects information on the presence of industry, exclusive of other economic activities. That is to say that industry is present at locations with higher values but stands out in relation to other economic activities. The fourth urban attribute of the neighborhood (*pca2_ino_n*) is based on company turnover in sectors that are strong in innovative service provision, specifically commerce, science and education, publicity, and information technology. In order to complement the description of location, the presence of public services – mainly arborization – is depicted by *f_melpub*. The presence of major shopping centers up to a radius of 1,500 meters is represented by *shop_1500*, presence of slums up to a radius of 200 meters is represented by *slum_200*. Finally, the presence of an avenue or arterial roads up to a radius of 200 meters (*aven_200*) (see Table 6-5).

Dataset details related to the neighborhood characteristics are available on request.

6.4 Model, diagnoses, tests, weight matrix alternatives, and results

The basic model to be applied has the log of the price – the actual transaction value – as the dependent variable and the summarizing characteristics of the estate and the neighborhood, according to Table 6-5. The estimating equation is:

$$lnPrice = \alpha + \beta_1 lnArea + \beta_2 Standard + \beta_3 Age + \beta_4 D\_ap + \beta_5 Domic\_n + \beta_6 PCA\_activ + \beta_7 PCA\_innovation + \beta_8 Factor\_melpub + \beta_9 Shop\_1500 + \beta_{10} Slum\_200 + \beta_{11} Aven\_200 + \varepsilon.$$

The first step concerns the test for spatial dependence. According to Anselin (Anselin, 1988, pp. 103-104; Anselin, Syabri *et al.*, 2006), this can be done using the LM error statistic for spatial correlation in residuals and the LM lag statistic for omitted spatial lags. However, to implement the tests, a weight matrix, that is, the exogenously determined spatial dependence of the data

Table 6.5 Basic statistics of variables used in the model

Variable	Variable code	Observations	Minimum	Maximum	Average	Standard-deviation
				Estate's attributes		
Log of transaction value - log price	lnprice	5512	7.155	15.109	11.470	0.747
Log of transaction value per surface	lnpricearea	5512	4.730	10.444	6.689	0.506
Log of estate's surface	lnarea	5512	0.880	7.552	4.781	0.629
Construction quality levels	standard	5512	1	5	2.790	0.771
Age of estate (in years of construction)	age	5512	1	83	21.185	14.525
Dummy apartment	d_ap	5512	0	1	0.587	0.492
Dummy house	d_house	5512	0	1	0.288	0.453
Dummy shop	d_shop	5512	0	1	0.055	0.227
				Location's attributes		
Normalized index of characteristics of the family and domicile by spatial units (see chapter 5) [0 worst;1 best]	domic_n	5512	0	1	0.229	0.190
Normalized index (PCA) of level of activities. Includes services, number of establishments	pca_activ	5512	0	1	0.150	0.179
Principal component that isolates industry activities in relation to services and commerce	pca_ind_n	5512	0	1	0.246	0.246
Second component of PCA analysis that concentrates innovative sectors and education services	pca2_ino_n	5512	0	1	0.540	0.123
Presence of public services. 98% of those without services refer to lack of arborization	f_melpub2	5512	0 (43%)	1 (57%)	0.566	0.496
Presence of major shopping mall within 1500 m	shop_1500	5512	0 (78%)	1 (22%)	0.217	0.412
Presence of slums within 200 m	slum_200	5512	0 (84%)	1 (16%)	0.163	0.369
Presence of an avenue or arterial road within 200 m	aven_200	5512	0 (48%)	1 (52%)	0.519	0.500

has to be constructed. As suggested before, different definitions of the matrix might generate different, even contradictory results. We therefore test for different conceptualizations of matrices to confirm the robustness of the matrices chosen.

6.4.1 Weight matrices

Three matrices are used. The most traditionally defined is the *contiguity matrix*, in which all neighbors are calculated based on Delaunay triangles. In this case, distance is not important and neighbors are all the closest individuals to a certain observation. Another matrix typically applied is the *distance matrix*, which considers the distance to neighbors instead of a binary matrix (neighbor, non-neighbor). In the default alternative (Anselin, 2005) all observations have at least one neighbor, preventing an island (no neighbor) situation from happening. For the dataset, this maximum distance was 1,057 meters. The third weight matrix to be considered is called the *neighborhood matrix* and is defined as observations within the same neighborhood – the spatial unit (UDH) – are considered neighbors. The idea is that this matrix captures the notion of belonging to a certain identity-based neighborhood[72].

All matrices are row-standardized.

Visually, we present the matrices in Figure 6-2, Figure 6-3 and Figure 6-4. Every line indicates a connection among nodes (house transactions). The contiguity matrix, most widely used in the literature, considers all the closest neighbors, independently of distances in the spatial configuration. This implies some deformations in the spatial structure tested. The number of connections is smaller when compared to other matrices. The distance matrix is almost as dense as the neighborhood matrix; however, its links encounter no limits, such as railways or rivers. The neighborhood matrix depicts the idea of belonging, in which observations are clearly separated in different groups, which further conveys the idea that the influence of observations of one group on another is less important.

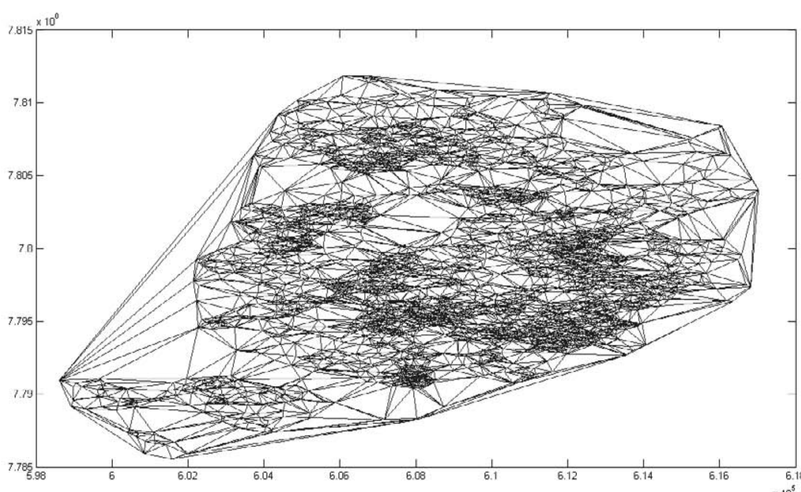


Figure 6.2 Visualization of matrix of contiguity

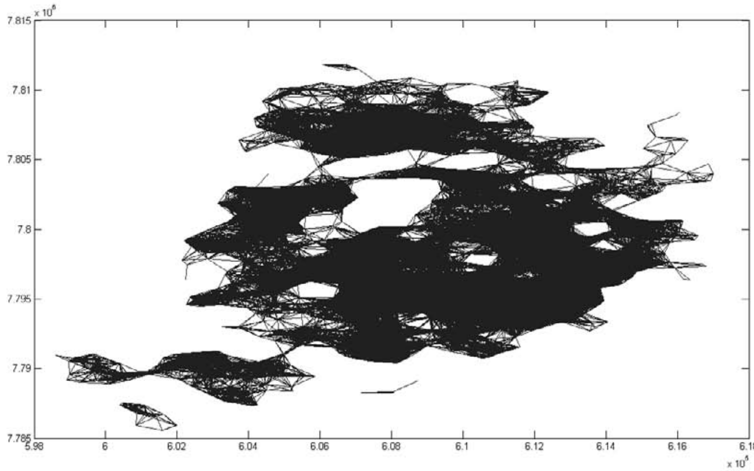


Figure 6.3 Visualization of matrix of distance. Source: elaborated by the author based on the dataset (GEAVI/PBH, 2007)

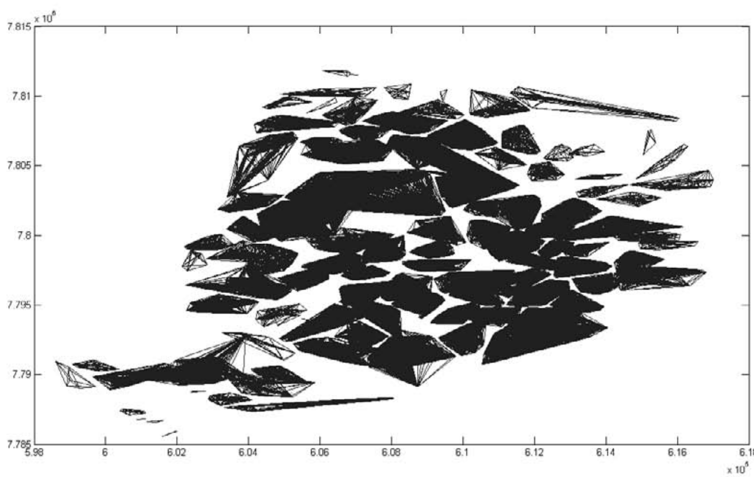


Figure 6.4 Visualization of matrix based on neighborhood. Source: elaborated by the author based on the dataset (GEAVI/PBH, 2007)

Table 6.6 Morans' I statistic as confirmation of spatial dependence

	W contiguity	**W distance**	**W neighborhood**
Moran's I	0.104	0.046	0.072
Moran's I-statistic	14.307	19.658	26.143
Marginal Probability	0.000	0.000	0.000
mean	-0.001	-0.001	-0.001
standard deviation	0.007	0.002	0.003

6.5 Interpretation of results and use of matrices

This section discusses the outcomes of the spatial, quantile and spatial-quantile models.

As expected, all results (Table 6-6, Table 6-7) confirm the strong spatial dependence of the problem, and therefore the need for spatial modeling. The OLS estimation is presented as a comparative parameter.

6.5.1 OLS and spatial

The results presented in Table 6-8 show that the coefficients are statistically significant and that there is a high degree of consistency among the different (weight-matrix) estimates. The observed differences help enhance understanding of the model, by providing insight into the spatial structure. Although the values for the R-squared are also close to one another, the best fit is for the spatial estimates of the neighborhood matrix with a value of 0.689. The log-likelihood is even better, at 2,028 in that model.

The presence of different spatial structures represented by the weight matrices should be kept in mind when interpreting the results. Given the results of the tests in Table 6-7, we consider the OLS results to be only indicative of order and parameters, because it is clear that the lack of spatial information renders this model inferior. The estimation with the contiguity matrix includes a weak spatial structure. The construction of the matrix, through utilization of Delaunay triangulation (Voronoi tessellation), defines only the closest observations as neighbors.

The results estimated with the distance and neighborhood matrices, on the other hand, include neighbors in a somewhat more broadly defined manner. The distance matrix includes every observation within a given radius, and the neighborhood matrix includes all neighbors who are within the same homogeneous spatial unit (see chapter 5). The difference between the two is that, while in the distance matrix, one includes neighbors within an abstract physical distance, in the neighborhood matrix, one considers identity in establishing neighborhood membership. Viewing the problem from an urban perspective in which features such as rivers, rails or roads divide space and in which being on one side or the other clearly matters, the neighborhood matrix seems more plausible, practically speaking. This is confirmed in the results, which indicate that this estimation fits the data best. Furthermore, the estimation with the contiguity matrix seems to perform slightly better than the distance matrix if one considers the value of the R-squared statistic.

In terms of the influence of estate-surface on value, there is little variation between the estimates, and on average, an increase of 1% in individual living area will increase prices about 0.76%. However, there is a slight decrease from the estimations of the OLS to the neighborhood matrix, suggesting that as model explanation power improves, less value is connected to the building-surface coefficient. When it comes to the quality of the estate, represented by the variable *standard*, it is clear that some of its importance is due to location instead of individually defined attributes, when comparing OLS results with other estimations. Therefore, if one considers the OLS estimate, moving upwards one category within the standard levels (from P3 to P4, for instance), there would be an increase in estate prices that is higher than average on spatial estimates[73]. The influence of the age of the estates is largely similar, indicating that aging one year reduces prices around 0.5%, on average.

Table 6.7 Tests for spatial autocorrelation

Anselin's tests for spatial autocorrelation				
	LM ERROR (chi 0.01 17.611)		**Robust (chi 0.01 6.64)**	
	value	**prob.**	**value**	**prob.**
W contiguity	198.767	0.000	203.083	0.000
W distance	351.104	0.000	151.690	0.000
W neighborhood	626.101	0.000	561.710	0.000
	LM LAG (chi 0.01 6.64)		**Robust (chi 0.01 6.64)**	
	value	**prob.**	**value**	**prob.**
W contiguity	3.551	0.060	7.867	0.005
W distance	326.606	0.000	127.192	0.000
W neighborhood	97.383	0.000	32.991	0.000

There are much larger differences over model specifications concerning the estimates of urban attributes. It seems, however, that the results of the estimates with the neighborhood matrix are again most plausible.

We used the information gathered thus far to illustrate the impact on estate prices when considering moving into a better neighborhood or a worse neighborhood. When comparing an estate located in an emerging middle-class neighborhood such as *Buritis* with an estate with exactly the same attributes but in an upper-class traditional neighborhood such as *Lourdes* or with a traditional, stable neighborhood such as *Caiçara*, we find the results presented in Table 6-9.

Note that a change of neighborhoods is not a linear move. For example, even though domiciles and families in Buritis and Caiçara are only slightly different in value (0.67 to 0.63), and Caiçara has an older but more stable availability of services (*pca_activ* from 0.23 to 0.09), Buritis has a lower negative impact from the presence of heavy industry (0.32 to 0.54) and, on the contrary, a higher presence of dynamic services and commerce (0.86 to 0.59). In this case, moving into a somewhat lower-class or less dynamic neighborhood, means a decrease in price an average of 9%.

Comparing the matrices' impact, if one considers the matrix of neighborhood embeddedness, than this difference is 2 percentage points higher than the average. If, on the other hand, the distance matrix is used, the decrease in prices is downplayed to a mere 5%, compared to an average of 9%.

A movement from a middle-class neighborhood to nearly the best neighborhood in terms of presence of economic activities, innovative sectors, domiciles and family characteristics would mean an increase of 51%, on average, of prices of apartments. Once again, the distance matrix tends to diminish differences, while the neighborhood matrix amplifies them.

The other variables considered come from the socioeconomic characterization of the urban fabric (Figure 5-1) and aspects (c) and (e)[74] as defined by Galster (2001). There is not much difference in how they impact estimates. All models show (as expected) positive effects on evaluation of being near major shopping centers, but this is less so when considering the

Table 6.8 OLS and spatial model results

	OLS	Spatial model		
		W contiguity	W distance	W Neighborhood
		Lnprice		
Lnarea	0.772	0.765	0.764	0.758
	(77.93)**	(81.17)**	(79.13)**	(79.02)**
Standard	0.144	0.118	0.113	0.110
	(15.23)**	(13.26)**	(12.9)**	(11.61)**
Age	-0.003	-0.004	-0.005	-0.004
	(-6.95)**	(-7.62)**	(-9.72)**	(-8.94)**
D_ap	-0.083	-0.111	-0.129	-0.149
	(-5.99)**	(-7.75)**	(-9.39)**	(-10.3)**
Domic_n	0.365	0.350	0.239	0.266
	(10.76)**	(8.8)**	(7.16)**	(4.44)**
Pca_activ	0.343	0.403	0.232	0.499
	(8.11)**	(7.67)**	(5.37)**	(5.97)**
Pca_ind_n	-0.278	-0.269	-0.151	-0.248
	(-5.89)**	(-4.79)**	(-3.09)**	(-2.53)*
Pca2_ino_n	0.270	0.288	0.158	0.429
	(5.26)**	(5.37)**	(3.19)**	(4.78)**
Shop_1500	0.170	0.176	0.079	0.113
	(10.14)**	(8.38)**	(4.65)**	(5.47)**
Slum_200	-0.073	-0.054	-0.081	-0.063
	(-4.54)**	(-2.94)**	(-5.1)**	(-3.72)**
Aven_200	0.061	0.056	0.045	0.038
	(5.11)**	(4.15)**	(3.84)**	(3.05)**
F_melpub2	0.108	0.077	0.067	0.061
	(8.23)**	(5.78)**	(5.21)**	(4.61)**
Constant	6.974	7.161	3.271	6.779
	(126.51)**	(561.19)**	(20.24)**	(646.43)**
rho		-0.005	0.355	0.037
		(-3.00)**	(27.2)**	(5.69)**
lambda		0.313	0.070	0.594
		(65.69)**	(8)**	(172.64)**
Observations	5512	5512	5512	5512
Log-likelihood		1968.5	1983.7	2028.1
R-squared	0.666	0.683	0.680	0.689

Absolute value of t statistics in parentheses
* significant at 5%; ** significant at 1%

neighborhood matrix in contrast to OLS models or a contiguity matrix. Again, the negative impact of the proximity of slums and the positive impact of being near main avenues are captured. The presence of public services (mainly arborization) presents the lowest value in the estimates when the neighbourhood matrix is considered. This is probably because having a pleasant, green neighborhood might be valued as an attribute of the neighborhood itself, which is accounted for by using the neighborhood matrix.

Table 6.9 Illustration of results' interpretation. Source: elaborated by the author based on results of chapter 5 and regressions

Neighborhoods/Values	Domic_n	Pca_activ	Pca_ind_n	Pca2_ino_n	
Lourdes (upper-class)	0.95	1.00	0.12	1.00	
Buritis (trendy, emerging, middle-class)	0.67	0.09	0.32	0.86	
Caiçara (traditional low-middle-class)	0.63	0.23	0.54	0.59	
Difference (moving into upper-class from Buritis)	0.28	0.91	-0.20	0.14	
Difference (moving into lower-class from Buritis)	-0.04	0.13	0.22	-0.26	
Coefficients for different estimations	**Domic_n**	**Pca_activ**	**Pca_ind_n**	**Pca2_ino_n**	
OLS	0.37	0.34	-0.28	0.27	
W contiguity	0.35	0.40	-0.27	0.29	
W distance	0.24	0.23	-0.15	0.16	
W Neighborhood	0.27	0.50	-0.25	0.43	
Impact on price for apartments for upward move (%)					**Sum**
OLS	0.10	0.31	0.06	0.04	0.51
W contiguity	0.10	0.37	0.05	0.04	0.56
W distance	0.07	0.21	0.03	0.02	0.33
W Neighborhood	0.07	0.45	0.05	0.06	0.64
Average	0.08	0.34	0.05	0.04	0.51
Impact on price for apartments for downward move (%)					**Sum**
OLS	-0.02	0.05	-0.06	-0.07	-0.10
W contiguity	-0.01	0.05	-0.06	-0.08	-0.10
W distance	-0.01	0.03	-0.03	-0.04	-0.05
W Neighborhood	-0.01	0.07	-0.05	-0.11	-0.11
Average	-0.01	0.05	-0.05	-0.08	-0.09

6.5.2 Quantile Regression Results

The results of the quantile regression are insightful for decomposing the variance within different quantile, but they should be viewed with caution because the tests clearly demonstrate the need for a spatial model.

In general, the analysis in Table 6-10 clearly shows that the impact of the various factors differs from the lowest considered quantile (at.1) to the highest (at.9). All present significance. The results and signs are conform to intuition.

The size of the apartment is a much more important factor for valuation in the cheapest segment of the city *vis-à-vis* more expensive parts of the city. It appears that as apartments become larger and more expensive (as in higher quality city segments), surface area becomes proportionally less important (the marginal utility of more surface area decreases with general improvements in quality). The same interpretation applies to the quality of the apartment. Most likely, more luxurious apartments already have a high-quality standard. It is clear that houses are preferred in the high-priced market, whereas apartments are more valued in the low-priced market.

The quality of neighborhood has a mixed influence on estate prices (see previous section), but appears to have a similar relationship with prices across all quantiles. Still, the neighborhood

Table 6.10 Results for quantile regression

	q .1	q .25	q .5	q .75	q .9
			Lnprice		
Lnarea	0.909	0.895	0.806	0.761	0.667
	(79.15)**	(106.42)**	(101.37)**	(64.15)**	(34.78)**
Standard	0.221	0.192	0.201	0.170	0.181
	(29.96)**	(28.41)**	(26.44)**	(14.29)**	(9.85)**
Age	-0.011	-0.009	-0.005	-0.002	-0.003
	(32.31)**	(27.92)**	(14.07)**	(3.83)**	(3.41)**
D_ap	0.201	0.101	-0.061	-0.181	-0.440
	(16.66)**	(9.54)**	(5.45)**	(11.73)**	(20.77)**
Domic_n	0.240	0.385	0.351	0.250	0.263
	(10.34)**	(17.73)**	(12.91)**	(5.36)**	(3.54)**
Pca_activ	0.149	0.272	0.400	0.517	0.528
	(4.15)**	(8.42)**	(11.77)**	(10.86)**	(7.70)**
Pca_ind_n	0.001	-0.110	-0.230	-0.377	-0.461
	-0.030	(3.23)**	(6.08)**	(6.61)**	(5.44)**
Pca2_ino_n	0.173	0.089	0.181	0.205	0.209
	(4.03)**	(2.30)*	(4.38)**	(3.50)**	(2.54)*
Shop_1500	0.090	0.117	0.154	0.181	0.208
	(6.20)**	(9.06)**	(11.45)**	(9.56)**	(7.79)**
Slum_200	-0.088	-0.048	-0.074	-0.091	-0.096
	(6.70)**	(3.99)**	(5.73)**	(4.95)**	(3.77)**
Aven200	0.015	0.024	0.030	0.056	0.094
	-1.590	(2.74)**	(3.09)**	(4.13)**	(5.00)**
f_melpub2	0.048	0.058	0.079	0.099	0.151
	(4.51)**	(6.02)**	(7.52)**	(6.56)**	(7.05)**
Constant	5.844	6.100	6.713	7.349	8.168
	(116.71)**	(149.09)**	(151.67)**	(107.07)**	(75.35)**
Observations	5512	5512	5512	5512	5512
Pseudo-R2	0.502	0.475	0.455	0.464	0.472

Absolute value of t statistics in parentheses
* significant at 5%; ** significant at 1%

appears to matter less for prices of the cheapest apartments. The preference for economic activities, however, increases steadily over quantiles, and reaches its highest value in relation to pricing of the most expensive apartments. The same pattern can be observed for the relation of the intensity of innovative sectors and education with prices in the sorted quantiles, but in a less steep manner.

Being within the 1,500 meters of large shopping centers appears to be more important for apartments within and above median prices, and less so for apartments at the lowest quantile. The proximity to slums presents negative values for all price ranges. The same is true for the presence of a nearby avenue, with greater importance for pricing of estates in the upper range.

6.5.3 Spatial-quantile (IVQR)

The results of the spatial-quantile model (Table 6-11) confirm the indications of both spatial and quantile estimations and should be considered the best possible estimate, according to the theoretical discussion provided by Su and Yang (2007).

For all weight matrices tested, the estimates vary little over the estate's attributes and show, in general, the same trend. The coefficient for estate-level surface *area*, for instance, consistently decreases in importance from the cheapest to the most expensive estates, and is slightly smaller in the case of estimation using the neighborhood matrix. The same trend can also be observed in relation to the quality of the apartment (*standard*) and the dummy variable that characterizes the estate as an apartment. The age of the estate also has approximately identical coefficients for all matrices as well.

The increase in importance of the quality of the domiciles and neighborhoods from the lowest to the highest quantiles manifests in specifications with different weight matrices. However, its effect is strongest for the neighborhood matrix.

In relation to the level of activities (*pca_activ*), the strength of the coefficients differs substantially between specifications using the matrix of contiguity compared to that of the neighborhood. Clearly, because the matrix of contiguity itself does not account for centrality, the coefficients partly express this aspect as well. For the models that use the matrix of neighborhood on the other hand, part of the effect of centrality is already embedded in the neighbourhood, so that the coefficient expresses, to a larger extent, the isolated advantage of having a higher service level.

The variable that measures the dominant industrial activity (*pca_ind_n*) consistently shows that repulsiveness is an increasing function of the price of the estate, and it is nearly zero for the lowest quantile. This does not hold for the model using the neighborhood matrix. In this case, the negative effect of the industry only lowers prices for the most expensive estates.

In terms of capturing the innovative activities effect, the model using the neighborhood matrix presents the highest significant values, with innovation being much more valued (and used) by individuals seeking more expensive apartments.

The proximity to large shopping centers is more expressed in the valuation of expensive estates in both the contiguity and neighborhood specifications, whereas in the model using the distance matrix, it has a slightly negative effect. This is probably because the variable itself is also based on distance.

All models confirm the negative valuation of proximity to slums (*slum_200*) and the positive valuation of the presence of arborization (*f_melpub2*).

The negative ρ for all models with different matrices and quantiles suggests that there is still heterogeneity in observations within each neighborhood framework. This is further confirmed by the analysis of spatial autocorrelation measured by Moran's I index (for the neighborhood matrix, the 0.5 quantile, for instance, has a statistic of 0.0723, which is statistically significant at the 1% level) and interpolation of the residuals (Figure 6-5). The variation of the residuals is not prominent across the city (left side of the figure), but it is present in the overall, aggregated plot of observations. The figure on the right side indicates that the difference of prices might be spatially attributed to distances of about 100 meters. The framework of the cognitively perceived urban space cannot capture this level of detail because citizens cannot identify, classify and name urban spatial units at this scale. Perceptions would overlap and be too fuzzy to be considered as valid general variables. Furthermore, the idiosyncrasies seem to be attached to the level of each

Table 6.11 Results for instrumental variables spatial-quantile regression

	W. Contiguity					**W. Distance**					**W. Neighborhood**				
	Quantiles					**Quantiles**					**Quantiles**				
Variables	0.1	0.25	0.5	0.75	0.9	0.1	0.25	0.5	0.75	0.9	0.1	0.25	0.5	0.75	0.9
constant	6.193	6.591	7.231	7.231	8.735	6.884	7.077	7.467	7.807	7.879	6.01	6.112	6.911	7.486	8.106
lnarea	0.916	0.877	0.783	0.783	0.629	0.914	0.876	0.778	0.698	0.639	0.910	0.876	0.770	0.703	0.616
standard	0.209	0.176	0.169	0.169	0.127	0.189	0.177	0.157	0.123	0.118	0.195	0.181	0.154	0.119	0.132
age	-0.012	-0.010	-0.007	-0.007	-0.006	-0.013	-0.010	-0.008	-0.006	-0.006	-0.013	-0.010	-0.008	-0.006	-0.006
d_ap	0.147	0.043	-0.115	-0.115	-0.489	0.146	0.043	-0.153	-0.314	-0.525	0.157	0.037	-0.155	-0.304	-0.482
domic_n	0.029	0.093	0.126	0.126	0.199	0.104	0.204	0.130	0.175	0.242	-0.018	0.040	0.138	0.103	0.243
pca_activ	0.141	0.331	0.676	0.676	0.654	0.039	0.048	0.238	0.350	0.225	-0.008	0.109	0.039	0.192	0.257
pca_ind_n	0.051	-0.114	-0.371	-0.371	-0.327	0.002	-0.040	-0.168	-0.275	-0.372	0.143	0.119	0.048	0.148	-0.251
pca2_ino_n	0.002	0.011	0.127	0.127	0.192	0.263	0.068	0.155	0.150	0.323	0.226	0.387	0.314	0.321	0.693
shop_1500	0.005	0.041	0.118	0.118	0.231	0.011	0.012	0.036	-0.007	-0.007	0.041	0.030	0.055	0.083	0.121
slum_200	-0.038	-0.002	-0.012	-0.012	-0.054	-0.099	-0.055	-0.065	-0.112	-0.112	-0.090	-0.049	-0.075	-0.098	-0.083
aven_200	-0.006	-0.002	0.018	0.018	0.033	0.007	0.010	0.023	0.039	0.054	0.005	0.011	0.015	0.035	0.035
f_melpub2	0.025	0.028	0.048	0.048	0.076	0.036	0.025	0.037	0.043	0.047	0.031	0.033	0.045	0.041	0.047
rho	-0.119	-0.126	-0.172	-0.172	-0.217	-0.092	-0.107	-0.078	-0.085	-0.034	-0.094	-0.059	-0.094	-0.152	-0.237

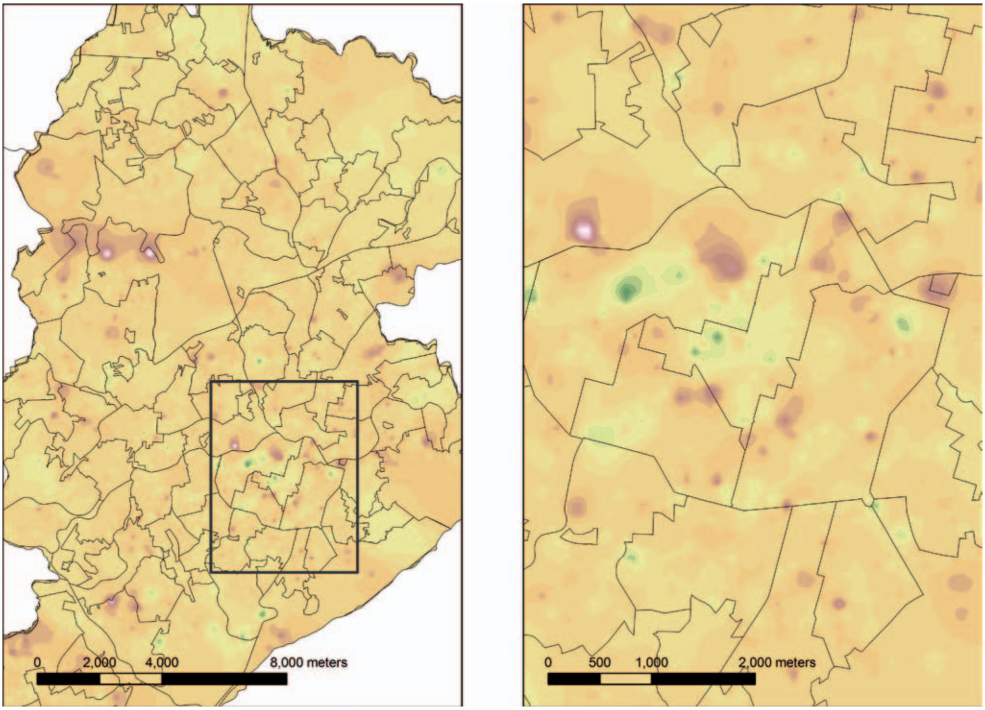


Figure 6.5 Spatial illustration of distribution of residues of IVQR estimation, neighborhood matrix, 0.5 quantile

observation. Consequently, what is not accounted for in the model might be an architectural style, for instance, or other particularities of the dwelling itself. The larger negative values of ρ are observed with the higher quantiles of the neighborhood matrix, which are more likely to have unaccounted, unique characteristics.

6.6 Concluding remarks

Our results confirm that we cannot evaluate real estate markets without taking spatial dependencies into account. The analyses also show that urban complexity embedded within cognitively-perceived neighborhoods can be included in estimation modeling in detail. The methodology applied allowed us to describe economic relations between city space and housing prices better than one could if only measuring distance to CBD.

This interpretation of the urban fabric using neighborhoods, which stems from the indices in chapter 5, allowed for the construction of a number of factors (indexes) that are easy to interpret. These indexes, in turn, made it easier to distinguish among the preferences of socially-heterogeneous actors. Moreover, these factors and the quantile estimation methodology explicitly show that those who can afford the most expensive estates value variables[75] in a considerably different manner from those seeking cheaper estates.

The different matrices also confirmed that the choice of matrix should always be made cautiously. The results show that the neighborhood matrix: (a) is the best fit to the data; (b) has the strongest theoretical support; (c) the best at reproducing a cognitive description of the city (see Figure 6-4); and (d) captures the idiosyncrasies of large urban metropolitan complexities without the need to restrict spatial influence to an arbitrarily-decided central place (i.e., the Central Business District).

Finally, this discussion and the results in the previous section determined the extent to which the various factors considered in the introduction – both at the estate level and at the neighborhood level – influence housing prices.

The insights from this chapter, in turn, are also helpful in calibrating the urban actor preferences that are dynamically modeled in chapter 8.

7 Description of reference data for calibration and validation

This chapter presents the methodology and results of the construction of reference maps of RMBH that serve as a reference and validation tool for the cellular automata model of the next chapter. The maps show the spatial configuration of actors in the RMBH grouped by income for 1991 and 2000. This also allows for the analysis of spatial change over this period.

Apart from providing an actor's configuration, which is used as reference and validation of the cellular automata model, the resulting maps provide a number of insights concerning urban heterogeneity.

First, they enable investigation of the dynamics occurring during 1991-2000. In practice, this means that the development of spatial configuration of inhabitants classified by income can be described.

The spatialization of level-income categories in the city allows for the empirical validation of the classic assumption of how land-use classes are positioned in relation to the CBD. Despite significant developments since Alonso's (1964) seminal work (Capello and Nijkamp, 2004b), urban economics and closely related fields of science retain 'distance to CBD' as a core factor (Lucas and Rossi-Hansberg, 2002). This assumption can be tested on the dataset that is generated.

The map generation took place for the conurbated municipalities of the Greater Area of Belo Horizonte (RMBH) in Brazil, with data from two census years: 1991 and 2000, which allows for the evaluation of the performance of cellular automata models of urban development (Pontius, Boersma *et al.*, 2008).

The chapter also describes the generation of a map of land prices, which is based on the dataset described in the previous chapter. This map is also used to validate the cellular automata model in the next chapter.

The remainder of the chapter presents the methodology and data used to construct income distribution and land price maps, discusses the applied metrics, results, and conclusions.

7.1 Methodology: multivariate analysis – clustering

Clustering can be described as any statistical procedure that classifies its elements in restricted, internally homogeneous groups that enable constructions of aggregated structures and typologies from a finite and multidimensional set of data (Simões, 2003; 2005). Clustering has been used in real estate by Bourassa (1999), and in agent-based modeling by, among others, Fernandez (2005). Clustering has also been used extensively in the ecological and natural sciences (Lillesand, Kiefer *et al.*, 2004).

Clustering is necessary because the available data are divided into ten classes of income and the proposed cellular automata model uses only three classes of income. Thus, there is a need to aggregate classes. For every census tract, the number of domiciles for every income level division is provided.

We propose clustering into three cluster classes that represent income levels thought to capture social heterogeneity: (1) high-income, (2) average-income, and (3) low-income. CA models usually restrict themselves to urban by non-urban analysis with no consideration for income levels (Pontius, Boersma *et al.*, 2008). The proposed division is also closer to the perception of the city by its inhabitants.

The proposed clustering procedure uses a modified iterative optimization, which is also known as a migrating means technique (Ball and Hall, 1965; Richards, 1986). It is performed using the ESRI ArcGIS 9.2 software package. The algorithm is an iterative process (Iterative Self Organizing – ISO/K-mean, ArcGIS 9.2 Spatial Analyst extension) for that compute the minimum Euclidean distance when assigning each candidate cell to a cluster.

A clustering procedure can be either supervised or unsupervised. In unsupervised clustering, the algorithm is applied to the data without any interference from the analyst.

Supervised clustering is applied when a certain class or category is known *ex-ante* by the researcher[76]. These areas of known value are called 'training samples', and they are used as initial subsets for the clustering procedure.

In this study samples are regions of the city that we are sure belong to one of the specific intended classes. The 'training samples' have for reference information provided by a neighborhood's classification scheme performed by the Foundation Institute of Economic, Administrative and Accountancy Research of Minas Gerais (IPEAD)[77] which is affiliated with the Federal University of Minas Gerais (UFMG). Their classification scheme is used to calculate real estate inflation indicators in the city of Belo Horizonte and reflects the inhabitants' perception of neighborhoods and the Neighborhood quality index calculated in chapter 5.

After a cluster class has been statistically defined, all candidate cells are assigned to a specific cluster class using a maximum likelihood classification. According to ESRI, "The maximum likelihood classifier calculates for each class the probability of the cell belonging to that class given its attribute values. The cell is assigned to the class with the highest probability" (2006). When the sample consists of sufficient cells, the classification will be more reliable, in the sense that spatial units are assigned to the cluster the researcher intends.

The clustering technique has two outcomes: (1) a dendrogram that displays the distance between resulting clusters, and (2) the average and co-variance of the cells of the original data in relation to each cluster generated. Together, the results allow for understanding and confirmation of the characteristics of each new cluster class.

To have additional parameters to compare and validate the results of the cellular automata, despite the visual description (Clarke, Hoppen *et al.*, 1997), some metrics typically used in ecological studies and physical geography (Mcgarigal, Cushman *et al.*, 2002; Visser and De Nijs, 2006) are presented. They are also referred to by Brown *et al.* (2005) (see Appendix 11.5, p. 217, for further details).

7.2 Data

The data used in this study consists of polygonal census tracts for the year 2000 (Brazil, 2003), defined by the Brazilian Statistics Foundation (IBGE) and a number of domiciles per income class[78] for both census years 1991 (Brazil, 1991) and 2000 (see Appendix for notes on data availability for the year 1991).

Some general statistics of the data are presented in Table 7-1 and Table 7-2.

To perform the aforementioned supervised clustering procedure, the data that is available in vector format (table data associated with the polygons) is transformed into a raster map (a grid format in which every cell – in this case of 86 by 86 meter grid – receives a value in the table). The study area encompasses the capital of the state of Minas Gerais, Belo Horizonte, and the neighboring municipalities, which, when combined, form the conurbated urban area of the Greater Belo Horizonte Area[79], which includes 90% of the RMBH population.

Table 7.1 Basic statistics of domiciles by census tract and levels of income, RMBH, 1991

Domiciles by census tract/ income level	Average	Standard-deviation	Maximum	Minimum	Total (2788 census tracts)
Above 20 times the minimum-wage salary	11.23	25	209	0	31296
15 and 20 m.w.	6.16	10.2	88	0	17184
Ten and 15 m.w.	13.02	16.47	113	0	36310
Five and ten m.w.	34.04	26.54	194	0	94915
Three and five m.w.	36.96	20.51	171	0	103040
Two and three m.w.	37.1	20.07	168	0	103448
One and two m. w.	61.18	39.1	276	0	170558
One half and one m.w.	46.43	40.35	257	0	129451
Up to half of one m.w.	19.45	24.87	233	0	54217
No income	8.27	9.19	91	0	23056
Total	273.84	92.79	770	2	763475

Table 7.2 Basic statistics of domiciles by census tract and levels of income, RMBH, 2000

Domiciles by census tract/ income level	Average	Standard-deviation	Maximum	Minimum	Total (4121 census tracts)
Above 20 times the minimum-wage salary	14.49	31.35	323	0	59707
15 and 20 m.w.	9.1	12.78	75	0	37502
Ten and 15 m.w.	12.57	13.15	80	0	51819
Five and ten m.w.	41.19	26.96	252	0	169736
Three and five m.w.	39.75	20.71	201	0	163801
Two and three m.w.	31.8	18.57	203	0	131947
One and two m. w.	47.67	32.41	396	0	196465
One half and one m.w.	31.81	22.76	206	0	131081
Up to half of one m.w.	1.23	1.84	16	0	5055
No income	21.71	19.87	290	0	89466
Total	251.32	80.41	1306	11	1035679

7.3 Results and discussion

7.3.1 Results of the clustering procedures

The clustering procedure results in well-defined classes for both 1991 and 2000 (Table 7-3 and Table 7-4). That is, the relative distance between classes is large enough that a distinctive typology is clearly established. Note, however, that the distance is much larger for 2000.

The analyses in Table 7-5 and Table 7-6 indicate that cells in the high-income cluster class mostly contain domiciles for which the head of the household has an income higher than 20 times a minimum-wage salary. The average-income cluster class has its highest frequency in the interval between 5 and 10 times the minimum-wage salary and the low-income cluster class is mostly comprised of households whose head of household has an income of one minimum-wage salary.

7.3.2 The reference year: 1991

This section describes in detail the spatial configuration of RMBH in 1991.

Figure 7-1 shows that the high-income residential area is clustered south of the original planned area with some other locations further to the north. The morphology is convoluted,

Table 7.3 Dendrogram of clustering RMBH, 1991

Distance between the clustered classes		
classes		distance
1	2	7,064
1	3	11,055

```
 0       1,2      2,5      3,7      4,9      6,1      7,4      8,6      9,8     11,1
 |--------|--------|--------|--------|--------|--------|--------|--------|--------|
2 ---------------------------------------------|
                                               |-------------------------|
1 ---------------------------------------------|                         |-
                                                                         |
3 -----------------------------------------------------------------------|
 |--------|--------|--------|--------|--------|--------|--------|--------|--------|
 0       1,2      2,5      3,7      4,9      6,1      7,4      8,6      9,8     11,1
```

Table 7.4 Dendrogram of clustering RMBH, 2000

Distance between the clustered classes		
classes		distance
1	2	8,169
1	3	18,499

```
 0       2.1      4.1      6.2      8.2     10.3     12.3     14.4     16.4     18.5
 |--------|--------|--------|--------|--------|--------|--------|--------|--------|
2 ------------------------------|
                                |---------------------------------------|
1 ------------------------------|                                       |-
                                                                        |
3 ----------------------------------------------------------------------|
 |--------|--------|--------|--------|--------|--------|--------|--------|--------|
 0       2.1      4.1      6.2      8.2     10.3     12.3     14.4     16.4     18.5
```

Table 7.5 Average and co-variance for supervised clustering RMBH, 1991

	High-income class									
Number of cells	**514**									
Layer	**20 times the m.w.**	**15 and 20 m.w.**	**Ten and 15 m.w.**	**Five and ten m.w.**	**Three and five m.w.**	**Two and three m.w.**	**One and two m. w.**	**One half and one m.w.**	**Up to half of one m.w.**	**No income**
Average	107.02	29.17	37.56	37.40	13.08	9.07	12.46	7.59	3.60	2.98
Co-variance										
20 times the minimum-wage salary	1208.8	80.8	-129.1	-257.2	-127.9	-80.9	19.1	71.7	16.6	16.5
15 and 20 m.w.	80.8	99.2	63.9	54.3	1.9	-18.2	-59.9	-48.3	-21.7	8.5
Ten and 15 m.w.	-129.1	63.9	187.0	159.5	39.3	2.7	-74.1	-80.0	-31.0	0.8
Five and ten m.w.	-257.2	54.3	159.5	323.6	90.8	38.5	-45.2	-60.6	-16.8	3.8
Three and five m.w.	-127.9	1.9	39.3	90.8	55.8	35.6	47.0	33.5	17.5	0.6
Two and three m.w.	-80.9	-18.2	2.7	38.5	35.6	50.0	103.5	94.4	42.2	1.4
One and two m. w.	19.1	-59.9	-74.1	-45.2	47.0	103.5	346.0	333.2	142.6	2.7
One half and one m.w.	71.7	-48.3	-80.0	-60.6	33.5	94.4	333.2	351.7	148.9	-0.1
Up to half of one m.w.	16.6	-21.7	-31.0	-16.8	17.5	42.2	142.6	148.9	64.3	-0.2
No income	16.5	8.5	0.8	3.8	0.6	1.4	2.7	-0.1	-0.2	5.6

(Table 7.5 continued)

Medium-income class

Number of cells	**744**									
Layer	**20 times the m.w.**	**15 and 20 m.w.**	**Ten and 15 m.w.**	**Five and ten m.w.**	**Three and five m.w.**	**Two and three m.w.**	**One and two m. w.**	**One half and one m.w.**	**Up to half of one m.w.**	**No income**
Average	17.09	13.60	31.97	70.29	46.06	30.57	32.14	19.28	8.15	5.54
Co-variance										
20 times the minimum-wage salary	95.3	40.1	75.2	63.8	-21.1	-39.3	-59.0	-51.5	-29.0	-3.6
15 and 20 m.w.	40.1	45.6	68.8	87.5	16.1	-7.2	-25.8	-24.4	-10.0	-0.2
Ten and 15 m.w.	75.2	68.8	182.4	221.4	63.1	12.6	-45.5	-33.9	-27.6	0.5
Five and ten m.w.	63.8	87.5	221.4	538.7	216.4	107.4	20.8	13.3	-8.4	9.5
Three and five m.w.	-21.1	16.1	63.1	216.4	202.3	109.1	87.8	74.3	38.4	15.0
Two and three m.w.	-39.3	-7.2	12.6	107.4	109.1	110.1	87.2	69.4	40.3	7.1
One and two m. w.	-59.0	-25.8	-45.5	20.8	87.8	87.2	150.8	102.4	53.2	10.7
One half and one m.w.	-51.5	-24.4	-33.9	13.3	74.3	69.4	102.4	129.5	56.6	8.9
Up to half of one m.w.	-29.0	-10.0	-27.6	-8.4	38.4	40.3	53.2	56.6	104.7	3.0
No income	-3.6	-0.2	0.5	9.5	15.0	7.1	10.7	8.9	3.0	17.7

(Table 7.5 continued)

Low-income class

Number of cells	**1428**									
Layer	**20 times the m.w.**	**15 and 20 m.w.**	**Ten and 15 m.w.**	**Five and ten m.w.**	**Three and five m.w.**	**Two and three m.w.**	**One and two m. w.**	**One half and one m.w.**	**Up to half of one m.w.**	**No income**
Average	1,392	0,647	2,660	11,543	25,392	35,445	106,892	130,961	62,131	19,567
Co-variance										
20 times the minimum-wage salary	1.0	0.1	0.4	0.8	1.1	-2.7	-8.2	-7.6	-0.1	-12.3
15 and 20 m.w.	0.1	1.1	0.8	5.2	8.8	7.2	4.8	-2.1	-0.3	-4.5
Ten and 15 m.w.	0.4	0.8	3.3	12.2	14.1	9.5	-2.3	-5.1	-0.2	-6.7
Five and ten m.w.	0.8	5.2	12.2	127.2	156.5	129.5	104.4	33.2	-0.4	13.7
Three and five m.w.	1.1	8.8	14.1	156.5	288.9	243.4	241.3	73.3	-1.7	51.0
Two and three m.w.	-2.7	7.2	9.5	129.5	243.4	349.8	475.2	241.7	1.0	208.6
One and two m. w.	-8.2	4.8	-2.3	104.4	241.3	475.2	1176.3	744.0	15.9	748.7
One half and one m.w.	-7.6	-2.1	-5.1	33.2	73.3	241.7	744.0	684.4	8.6	656.5
Up to half of one m.w.	-0.1	-0.3	-0.2	-0.4	-1.7	1.0	15.9	8.6	4.2	3.0
No income	-12.3	-4.5	-6.7	13.7	51.0	208.6	748.7	656.5	3.0	1317.2

Table 7.6 Average and co-variance for supervised clustering RMBH, 2000

	High-income class									
Number of cells	**422**									
Layer	**20 times the m.w.**	**15 and 20 m.w.**	**Ten and 15 m.w.**	**Five and ten m.w.**	**Three and five m.w.**	**Two and three m.w.**	**One and two m. w.**	**One half and one m.w.**	**Up to half of one m.w.**	**No income**
Average	148.443	31.457	21.457	22.661	6.787	4.268	7.763	3.865	0.014	4.261
Co-variance										
20 times the minimum-wage salary	2414.4	171.0	14.4	-174.4	-132.7	-70.0	-202.2	-114.6	-0.9	-31.4
15 and 20 m.w.	171.0	148.1	92.0	84.1	13.7	-8.5	-25.1	-6.6	-0.1	-9.4
Ten and 15 m.w.	14.4	92.0	95.2	99.0	26.1	2.5	8.3	10.8	0.0	-1.3
Five and ten m.w.	-174.4	84.1	99.0	179.0	53.8	12.4	24.3	20.4	0.1	0.5
Three and five m.w.	-132.7	13.7	26.1	53.8	29.9	10.9	31.6	19.2	0.1	2.9
Two and three m.w.	-70.0	-8.5	2.5	12.4	10.9	10.1	23.3	13.4	-0.0	6.5
One and two m. w.	-202.2	-25.1	8.3	24.3	31.6	23.3	86.4	48.7	-0.1	17.7
One half and one m.w.	-114.6	-6.6	10.8	20.4	19.2	13.4	48.7	29.3	-0.0	10.0
Up to half of one m.w.	-0.9	-0.1	0.0	0.1	0.1	-0.0	-0.1	-0.0	0.0	-0.0
No income	-31.4	-9.4	-1.3	0.5	2.9	6.5	17.7	10.0	-0.0	11.4

(Table 7.6 continued)

	Medium-income class									
Number of cells	**489**									
Layer	**20 times the m.w.**	**15 and 20 m.w.**	**Ten and 15 m.w.**	**Five and ten m.w.**	**Three and five m.w.**	**Two and three m.w.**	**One and two m. w.**	**One half and one m.w.**	**Up to half of one m.w.**	**No income**
Average	25.550	23.198	28.898	69.266	38.479	19.722	23.110	16.121	0.368	10.826
Co-variance										
20 times the minimum-wage salary	198.3	94.9	70.2	73.2	-25.2	-51.3	-81.6	-52.2	-2.9	-14.6
15 and 20 m.w.	94.9	74.3	50.5	65.0	7.0	-17.3	-31.3	-13.7	-1.4	-4.8
Ten and 15 m.w.	70.2	50.5	91.1	114.3	28.2	-8.6	-22.9	-11.6	-0.6	-4.5
Five and ten m.w.	73.2	65.0	114.3	316.5	95.0	28.5	3.8	4.7	-1.8	-13.5
Three and five m.w.	-25.2	7.0	28.2	95.0	119.2	58.6	79.3	57.2	1.9	29.5
Two and three m.w.	-51.3	-17.3	-8.6	28.5	58.6	70.3	78.1	57.4	1.9	27.0
One and two m. w.	-81.6	-31.3	-22.9	3.8	79.3	78.1	146.1	87.9	3.1	40.5
One half and one m.w.	-52.2	-13.7	-11.6	4.7	57.2	57.4	87.9	72.0	2.3	29.4
Up to half of one m.w.	-2.9	-1.4	-0.6	-1.8	1.9	1.9	3.1	2.3	0.5	2.2
No income	-14.6	-4.8	-4.5	-13.5	29.5	27.0	40.5	29.4	2.2	43.3

(Table 7.6 continued)

Low-income class

Number of cells	**2065**									
Layer	**20 times the m.w.**	**15 and 20 m.w.**	**Ten and 15 m.w.**	**Five and ten m.w.**	**Three and five m.w.**	**Two and three m.w.**	**One and two m. w.**	**One half and one m.w.**	**Up to half of one m.w.**	**No income**
Average	0.925	0.904	2.029	19.311	40.856	47.210	86.892	56.346	2.553	51.461
Co-variance										
20 times the minimum-wage salary	1.0	0.1	0.4	0.8	1.1	-2.7	-8.2	-7.6	-0.1	-12.3
15 and 20 m.w.	0.1	1.1	0.8	5.2	8.8	7.2	4.8	-2.1	-0.3	-4.5
Ten and 15 m.w.	0.4	0.8	3.3	12.2	14.1	9.5	-2.3	-5.1	-0.2	-6.7
Five and ten m.w.	0.8	5.2	12.2	127.2	156.5	129.5	104.4	33.2	-0.4	13.7
Three and five m.w.	1.1	8.8	14.1	156.5	288.9	243.4	241.3	73.3	-1.7	51.0
Two and three m.w.	-2.7	7.2	9.5	129.5	243.4	349.8	475.2	241.7	1.0	208.6
One and two m. w.	-8.2	4.8	-2.3	104.4	241.3	475.2	1176.3	744.0	15.9	748.7
One half and one m.w.	-7.6	-2.1	-5.1	33.2	73.3	241.7	744.0	684.4	8.6	656.5
Up to half of one m.w.	-0.1	-0.3	-0.2	-0.4	-1.7	1.0	15.9	8.6	4.2	3.0
No income	-12.3	-4.5	-6.7	13.7	51.0	208.6	748.7	656.5	3.0	1317.2

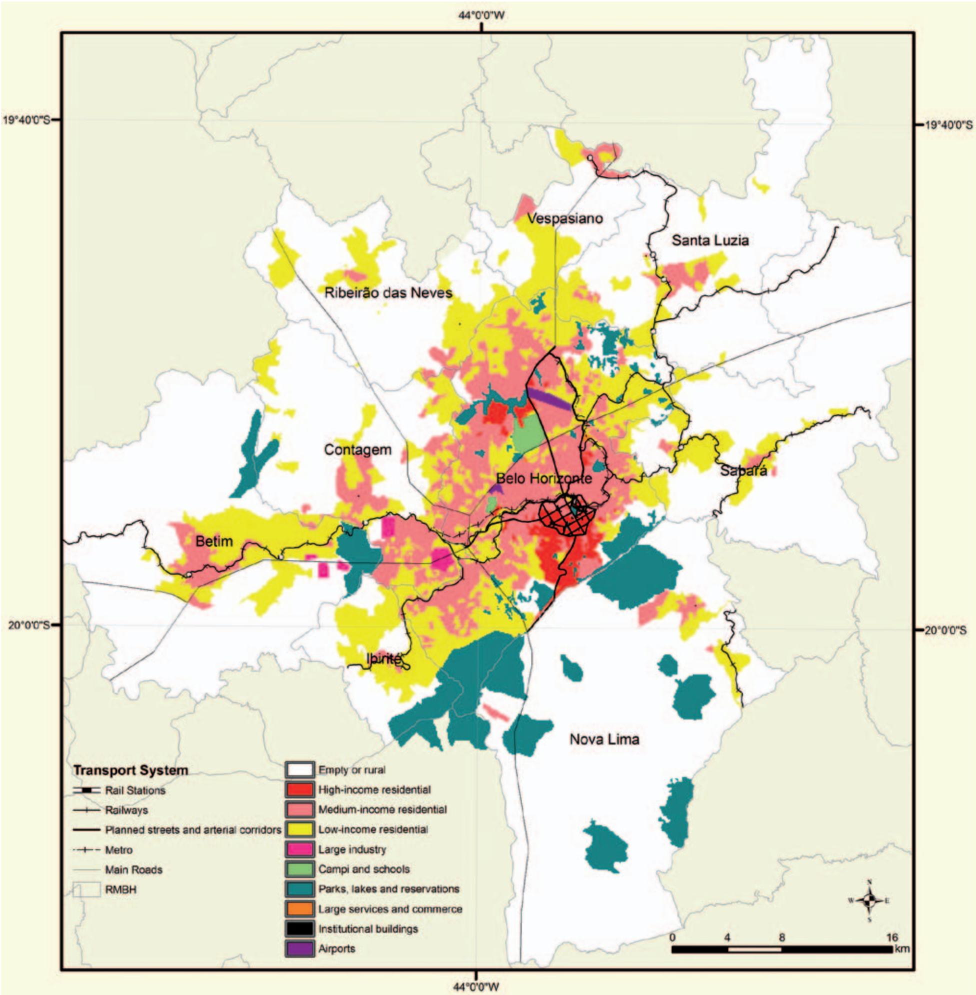


Figure 7.1 Spatial distribution of actors in 1991. Source: elaborated by the author; base map IBGE census tract; data source: IBGE, 2000 and 1991/2000 correlation information

although concentrated, and areas of low incomes are not only adjacent to high-income areas, but they also are found inside the upper-income cluster. Average-income residential areas are heavily clustered in an outer ring around the high-income areas. The spatial pattern is heavily sprawled, fringed, irregular and mixed with low-income residential areas.

These visual configurations can also be expressed quantitatively.

The numbers in Table 7-7 confirm the visual analysis. A higher value for the fractal dimensions of average-income areas compared to other areas indicates that average-income areas have a more convoluted form, along with larger clusters (larger patch size) and a greater shape index.

Table 7.7 Basic measures of reference map – 1991

Basic measures of REFERENCE map 1991			
High-income	**Average-income**	**Low-income**	**Global**
		Fractal dimension	
1.327	1.403	1.343	1.364
		Patch size	
1863	8594	4229	5712
		Perimeter	
641	2900	1168	1775
		Shape index	
3.547	6.941	4.326	5.242

As mentioned in the introduction of this chapter, the spatial distribution of income categories enables validation of the assumption in urban economics that 'distance to the Center Business District (CBD)' and 'cost of transport' determine where actors with different spending power reside. To quantify the presence of different actors in relation to the city center, a radial analysis is conducted. This measures the number of cells of each category in a 20-cell annulus. The initial position is at the city center[81].

The results (Figure 7-2) suggest that, in general, the general assumption in urban economics holds. That is, poorer residents tend to locate farther from the center, average-income households are closer to CBD, and higher-income households locate closest to the center. However, various violations of this general pattern can be observed: (1) in the first 10 km radius, all three categories of income are present; (2) the nearest area to the CBD is not exclusive of high-income occupation; (3) most of the city has overlapping of different income classes; and (4) at 9 km, the number of domiciles for high-income households is the same for the 4 km range, which suggests multinodal configurations.

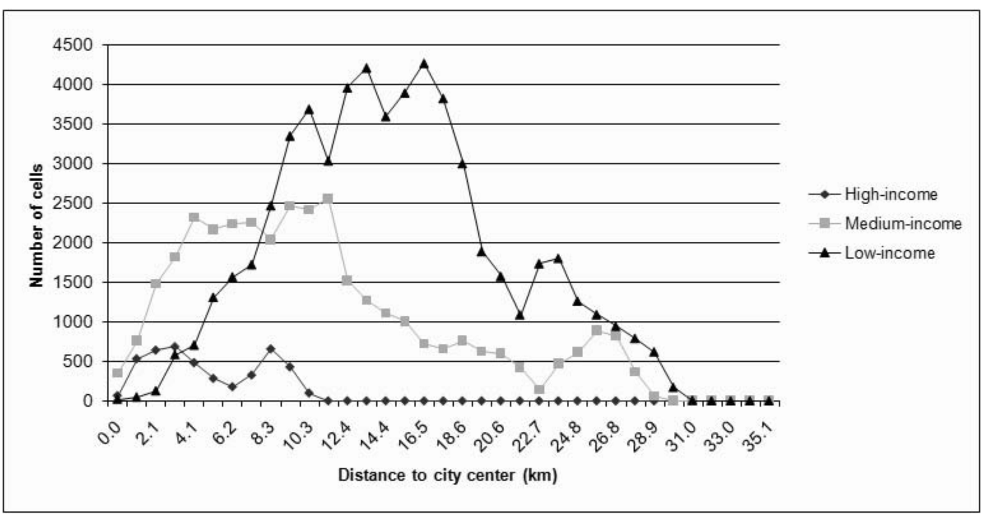


Figure 7.2 Radial distribution of actors from the city center – RMBH – 1991

7.3.3 Validation reference for 2000

The map of the Greater Area of Belo Horizonte for 2000 (Figure 7-3) shows that higher-class residential areas were, like in 1991, located mainly south of the original planned area up to the limits of Belo Horizonte (and the presence of a mountain barrier), with some presence west of the airport, known as Pampulha, and other small scattered areas. There was also a more recent occupation in the southern municipality of Nova Lima. The central areas of Betim and Santa Luzia also show some occupation.

Average-income areas are compact, but seem to be confined to the municipality of Belo Horizonte with some encroachment up on the industrial neighbors of Contagem and Betim. Presence in Ribeirão das Neves, Sabará or Ibirité, is nearly non-existent. Furthermore, note the 'in-between' characteristic of the spatial configuration of the average-income class. They locate as

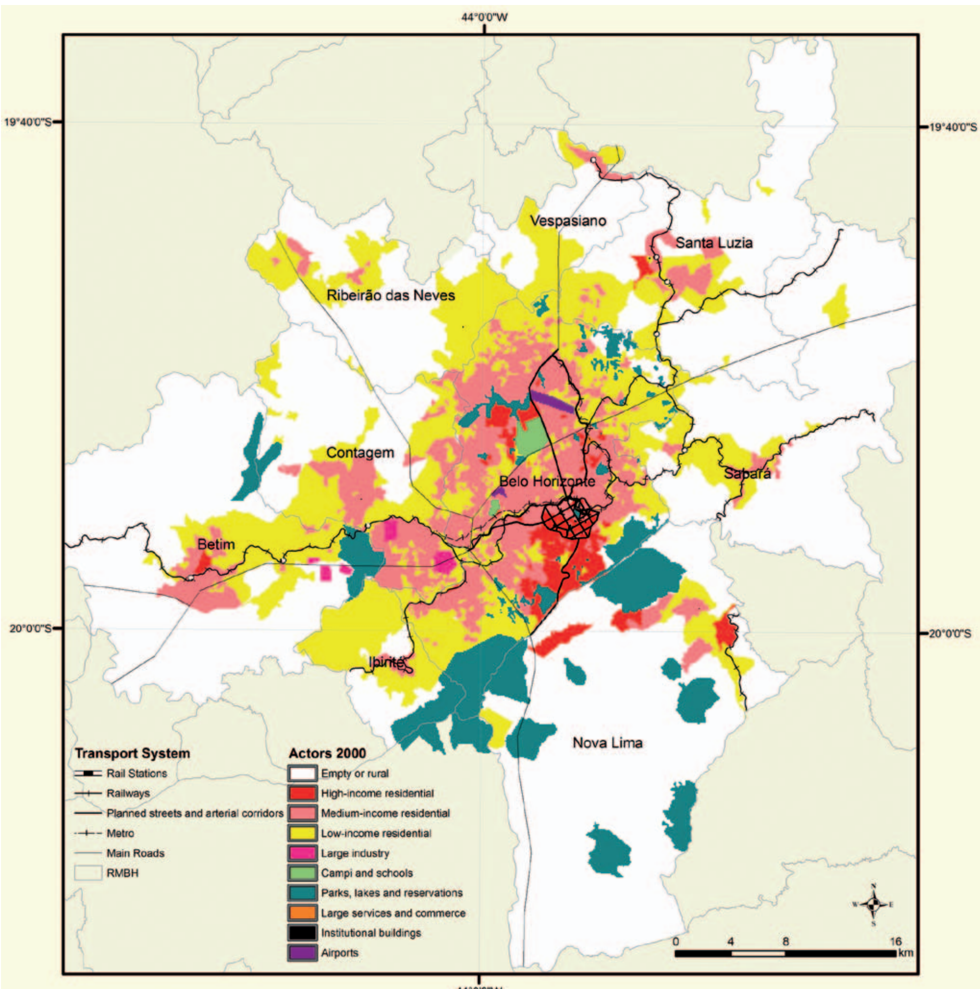


Figure 7.3 Spatial distribution of actors in 2000. Source: elaborated by the author; base map IBGE census tract; data: IBGE, 2000

Table 7.8 Basic measures of validation map – 2000

Basic measures of REFERENCE map 2000			
High-income	**Average-income**	**Low-income**	**Global**
	Fractal dimension		
1.334	1.429	1.386	1.397
	Patch size		
1751	14290	12819	12624
	Perimeter		
669	4985	3204	3651
	Shape index		
3.680	9.246	6.563	7.294

close as possible to high-income areas, but they have to contend with proximity to low-income residential areas.

The lower class is definitely predominant in the outer borders of the conurbated urbanized area of the metropolis. However, it is present in some parts of average-income class areas and even at the heart of some high-income areas.

For 2000, the numbers for RMBH are shown in Table 7-8. For the average-income areas, the patch size is much larger than for high-income areas. Consequently, the perimeter and shape index are also much larger. Furthermore, the patch size for average-income areas confirms its level of aggregation.

The results for the distance to the CDB (Figure 7-4) confirm the suggestions made for 1991. Once again, urban economics theories hold, in general, but are much coarser up to 12 km. Upper-class areas tend to be nearer the center; however, a peak occurs at around 8 km from the center, with another peak at 4 km. Average-income residential areas, in turn, are located in different distance-range areas with a bulk of occupation occurring up to 16 km, with a second peak further

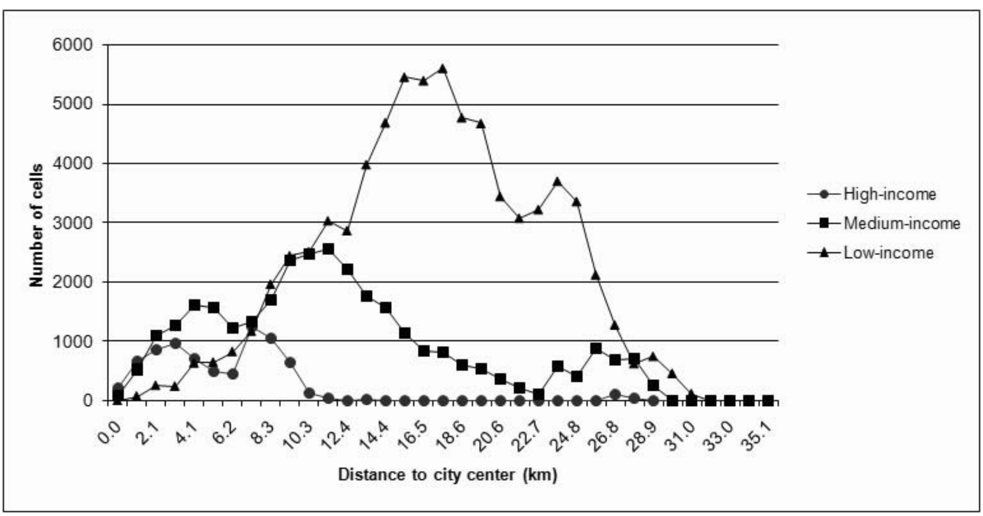


Figure 7.4 Radial distribution of actors from the city center – RMBH – 2000

away. The lower class locates away from the center, but it is also present in the inner regions of the city, and shows numbers similar to those of average-income areas.

7.3.4 Description of evolution from 1991 to 2000

To clarify the evolution of spatial occupation during the period of validation, the maps of differences among different income-levels are presented in this section. The results are produced with a simple superposition of Figure 7-1 and Figure 7-3, using the 'Map Comparison Kit' (Visser and De Nijs, 2006).

A large part of the difference observed is due to an increase in the population. As shown in Table 4-2, the number of inhabitants increases 22%, from 3,212,044 in 1991 to 3,904,172 in 2000.

Figure 7-5 depicts the occupation of high-income residential areas and highlights their decentralization. Although the core upper-income residential areas southern of the central ring

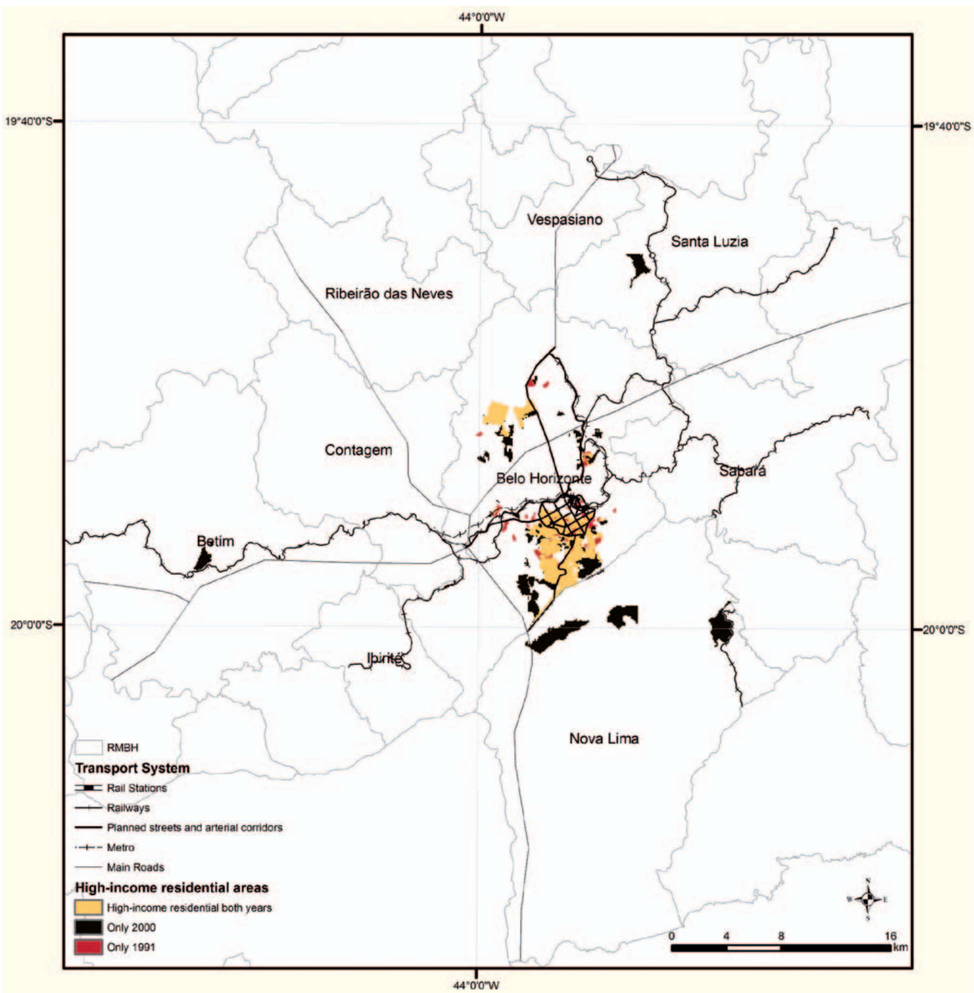


Figure 7.5 Evolution of occupation of high-income residential areas from 1991 to 2000

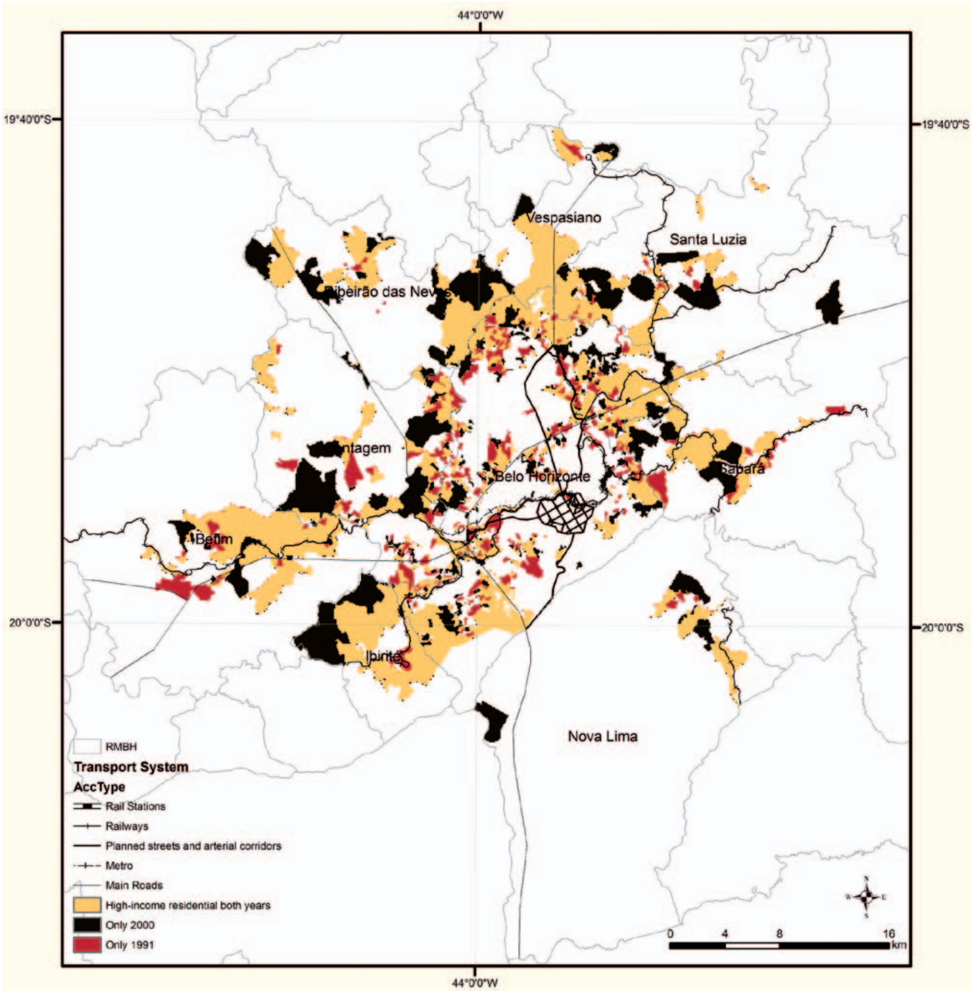


Figure 7.6 Evolution of occupation of low-income residential areas from 1991 to 2000

are maintained, additional occupation is observed towards the South Vector (Costa, H., Costa *et al.*, 2006) in the municipality of Nova Lima, as well as in Belo Horizonte in its eastern area around a large shopping center and its western area near the Federal University campus. The central parts of Betim, Santa Luzia and Nova Lima also present newer areas of occupation, indicating an upward trend of average-income locations becoming higher-income locations.

The change of occupation of low-income areas is shown in Figure 7-6. It confirms the decentralization and physical expansion of the city, which is mostly taken by people with lower incomes. This decentralization is confirmed by radial analysis (Figure 7-2 and Figure 7-4) from which one can say that half of the occupation of low-income residential areas is accumulated 14.4 km from the city center in 1991, compared to 15.5 km in 2000.

Table 7.9 Quantitative occupation of space by different residential actors

	Spatial analysis			
	Year 1991		Year 2000	
	Number of cells (absolute)	Percentage (%)	Number of cells (absolute)	Percentage (%)
High-income	3321	0.04	5850	0.063
Average-income	27164	0.36	31672	0.340
Low-income	44198	0.59	55570	0.597
Total	74683	1.00	93092	1.000
	Percentage analysis (a-spatial)			
	Year 1991		Year 2000	
High-income	0.04		0.06	
Average-income	0.33		0.41	
Low-income	0.63		0.53	

7.3.5 Quantitative occupation of space by different actors

The calculus of how many cells are assigned to each class enables an analysis of how much space is occupied by each income-level. This approach adds to the percentage analysis that considers domiciles in large aggregate areas.

In this chapter, the same database and a similar aggregation of income-level[82] are used to produce the results in Table 5.

The results of Table 7-9 indicate that the numbers for general percentage analysis and spatial analysis are comparable. However, for average-income households – which increase in number proportionally (8 percentage points) – physical occupation is slightly reduced. For low-income domiciles, the result is nearly the opposite. Although the number of families in this segment is proportionally reduced, occupation is practically stable. Together, these effects might suggest that low-income families occupy less dense areas in the outer borders of the metropolitan urban area. This is a similar result to that observed in the visual analysis.

7.4 Spatial configuration of prices in 2007

Because the results of the simulation also include a price map, the actual price map based on observations from the dataset (item 6.3) is presented (Figure 7-7).

The available systematic data on real estate prices refers to legal information provided by the City of Belo Horizonte (PBH). However, the "mapa de valores genericos" (or generic values map) is produced in order to enable the calculation of property taxes, not to accurately reflect market values. The criteria and weights were established by Law 8291 of 2001, which had not been updated as of 2008. A private institution connected to the university (IPEAD[83]) also calculates prices. However, these prices are based on small samples and focus on new supply of units and rental values rather than sale prices.

Older data are available, but only for very aggregated units, especially for 1950-1974 (Plambel, 1987).

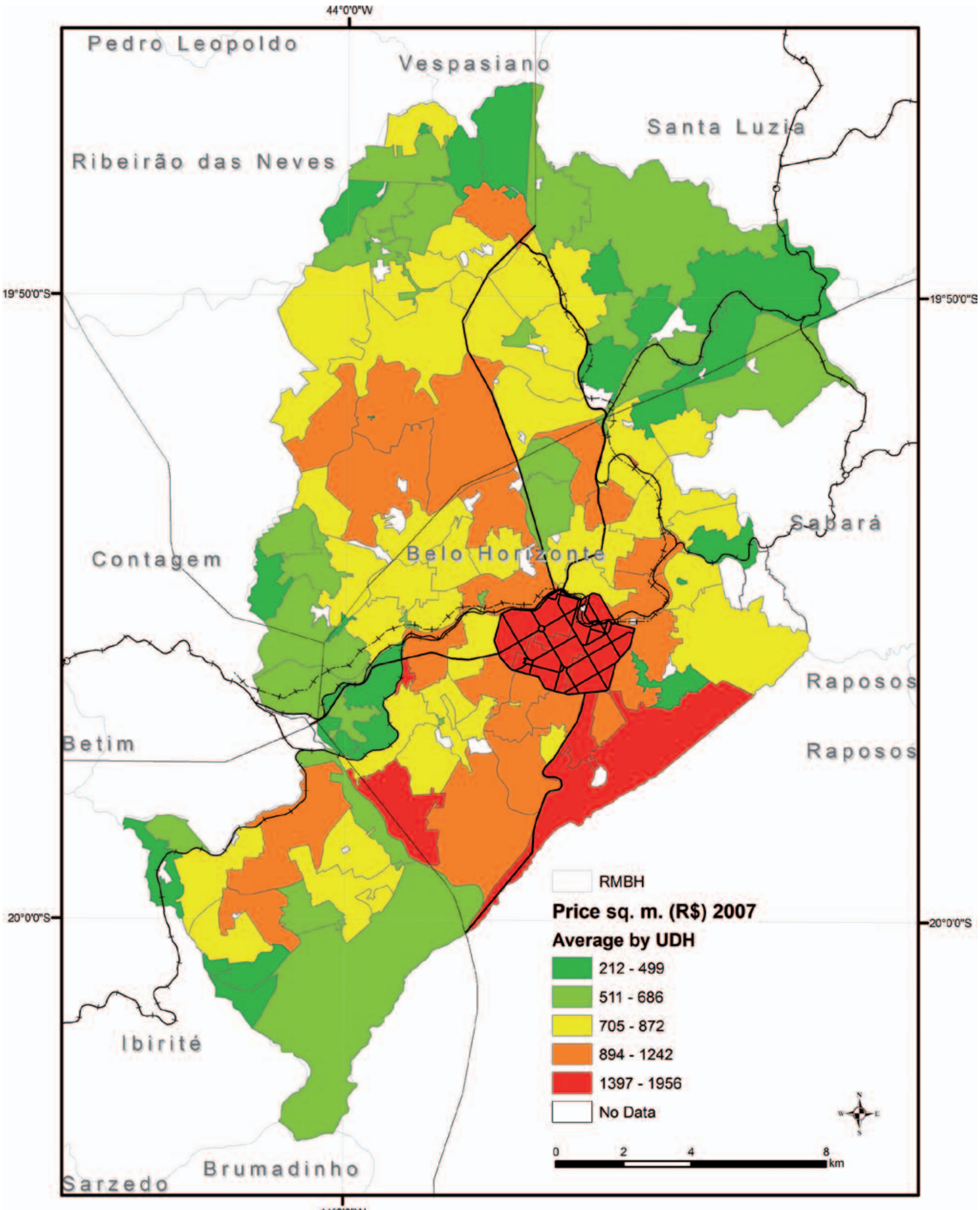


Figure 7.7 Map of price per square meter for Belo Horizonte, 2007. Source: elaboration of the author based on data from GEAVI/SFH/PBH.

Therefore, just the presentation of updated, accurate information is relevant to the city.

The dataset available would enable a construction of maps of rather small spatial units or the use of interpolation and kriging techniques. However, in accordance with the theoretical approach and the rest of the work, the map was constructed using the average per neighborhood (UDH).

A central point to highlight in the analysis of the map is its difference in magnitude among neighborhoods. The average square meter price on the cheapest neighborhood in Belo Horizonte is of R$ 212, which is almost one tenth of the most expensive neighborhood in which the average square meter is R$ 1,956. Despite this huge difference in order, visual analysis of the map shows that neighborhoods of the lowest typology are neighbors of the neighborhoods of the highest typology, with no clear distinction of agglomeration. In a more general perspective, the most expensive areas are closer to the planned original ring and poorer areas tend to concentrate in the outer borders of the municipality. Thus, the map confirms the heterogeneity of the city.

7.5 Concluding remarks

This chapter provides essential data for the reference and validation of the next chapter. Furthermore, it demonstrates that the use of georeferenced, detailed, and socioeconomic data may add considerably to a qualitative understanding of urban areas.

The empirical results contain a spatially detailed analysis of Belo Horizonte and its Metropolitan region. Previous works were not spatially explicit (Monte-Mór, 1994), used different structure (Mendonça, 2002), were restricted to the city of Belo Horizonte, or used less detailed information (aggregated units) (Plambel, 1987; Pinheiro, 2006; Santiago, 2006). The use of information at the scale of the census tract allows for improved results with high degrees of spatial confidence, which are consistent and further enhance more general results published earlier (Mendonça, 2002).

From the results, the urban economics assumption of location within the city in relation to income-level households holds in general for the level of individual cities. When the analysis is applied to a larger scale, however, there is no simple organization of land-uses relative to the city center.

In sum, the exercise provided in this chapter shows that intra-urban analysis has much to gain qualitatively if it seeks to make full use of broadly available quantitative information.

In conclusion, this chapter provides: (1) actor configuration that is useful as reference and validation for the CA model, (2) a demonstration of how quantitative exercises may be enlighten qualitative analysis and descriptions of urban environment, and, finally, (3) an illustration of how physical occupation of domiciles classified by income-level may vary differently from their proportion of the city in general.

8 Application: cellular automata model for the Greater Area of Belo Horizonte (1897-1991)

The theoretical approach described in chapter 3 is empirically tested in this chapter for the Greater Area of Belo Horizonte. As previously noted, the added value is the application of income-differentiated residential actors and the possibility of observing both the dynamics of actors and relative prices throughout the whole period. Also, the application of such models of urban development is new to extremely rapidly developing areas such as the Belo Horizonte area, which differs significantly from the European or North-American contexts.

This chapter begins with the general objective of the model, which is followed by a description of model actors, model data, and scenarios that are defined to test the model's face validity. Also, parameter settings are discussed. Finally, the results of the simulation for both actors and relative prices and their implications for model validity are discussed. A sensitivity analysis is performed to gain further insight into the dynamics of the process. The scenarios proposed are commented on and concluding remarks close the chapter.

8.1. Objective of the application

The objective of the proposed model is to represent the mutual relationship between price formation and actor location decisions, leading to specific spatial configurations. In this respect, both positive and negative feedback mediated by affordability play an important role.

Given the theoretical framework – and the distinction between forecasting and understanding processes made by Brown (2005) – this application's objective is not to predict the exact location of actors, but rather to simulate the general spatial structure, including, in particular: (a) the level of fragmentation or clustering, (b) the relative location of actors, and (c) the structure of the city, interpreted visually.

8.2. Definition of actors

In detailing modeling procedures we distinguish between *functions* and *features*. *Functions are* actors who are actually modeled in the simulation, whereas actors who are provided exogenously are *features*. Features are spatially fixed. Their location is provided top-down, as the result of policy and planning, at the year and place where they actually occupy the land. They influence their neighbors (other types of actors), but, unlike functions, they are not themselves the result of the modeled interaction processes.

In the model, only six actors are endogenously modeled (functions). The study focuses on the intra-urban allocation of residences at different levels of income. Small services, commerce and industry are modeled as supporting functions, whose locational decisions are strongly intertwined with residential development.

For the case study, the following actors[84] are considered:

The spatial resolution is chosen such that cell size allows for the capture of the differences and peculiarities of the intra-urban space at the block level. Cell sizes in applications of urban models usually vary between 50 and 150 meters (Hagoort, 2006; Waddell, Ulfarsson et al., 2007). However, caution should be taken to not impede generalization and the global character of the simulation (Riks, 2007a). In this case study, cells of 86 meters by 86 meters are used (roughly the size of a typical block in Belo Horizonte), which implies a surface area of 7,400 square meters, or about ¾ of one hectare. The data used to initiate and validate the model are census data given at the level of census tracts, which are larger than the cells used.

Additionally, census data do not contain information about non-residential actors, which implies that information on the locations of service, commerce, industry and "in the market" needs to be obtained from other sources. Although other sources (City Hall Tax database/ PBH-SMFA, 2001) provided aggregated information, for example, in the form of the number of establishments by neighborhood, they do not provide their exact location. Based on proportions derived from this database, the demand of these actors is designed so that 1% of the total occupied area is for small commerce and services, and 0.5% of the total occupied area is reserved for small industry. That is, the proportional quantity of cells allocated at each year of the simulation is pre-determined according to the proportion observed in the database. However, spatial location is given by the mechanisms described in the model (see item 8.5). Furthermore, the function in the market should account for the estates available at each period plus new allocations at each year such that even after the newcomers are allocated, 1% still remains unoccupied.

Table 8.1 Actors modeled, RMBH case study

	Actors
FUNCTIONS	Vacant or Rural Areas
	High-income residential area
	Average-income residential area
	Low-income residential area
	Small service and commerce
	Small industry
	In the market
FEATURES	Large industry
	Schools and Campuses (campi)
	Water and Reservation Areas
	Large Service and Commerce
	Institutional buildings
	Airports
	Out of modeling area

In practice, what this means is that only the residential allocation resulting from the simulation is actually compared with real data. More specifically, the results of the simulation for residential actors are compared to data generated through the clustering technique described in chapter 7, which only includes residential actors and exogenous features described in the next section. Given the essential role of the other actors in the dynamics of the urban structure, however, these actors had to be part of the process even if they prevented comparison of the data.

8.2.1 Exogenous historical features implemented

A few *features* were introduced throughout the studied period (1897-2000), including institutional buildings (mainly City Halls), airports, university campuses, industrial districts, and major shopping centers, with their exact location and year. A full list is provided in Table 8-2.

8.3 Application data

8.3.1 Workspace: study area

The workspace is defined as the municipality of Belo Horizonte and its neighbors (Betim, Contagem, Ibirité, Nova Lima, Ribeirão das Neves, Sabará, Santa Luzia and Vespasiano), which form the conurbated urban area of the Greater Area of Belo Horizonte[85]. Using the total area of the Greater Area of Belo Horizonte and its 34 municipalities introduces additional computational burden to the model without adding much in terms of explanatory power. In addition, the population of the selected workspace represented 90% (3,904,172 inhabitants) of the

Table 8.2 Exogenous input of features

Year	Institution	Actors (features)
1907	Law school	Schools and campuses
1914	City Hall Contagem	Institutional building
1933	Pampulha airport	Airports
1939	City Hall Betim	Institutional building
1943	UFMG - university campus	Schools and campuses
1944	Carlos Prates airport	Airports
1948	City Hall Vespasiano	Institutional building
1951	Highways (DER-MG)	Accessibility
1953	City Hall Ribeirao das Neves	Institutional building
1955	Industrial district (energy supply)	Large industry
1958	University campus - PUC-MG	Schools and campuses
1968	Industry (Petrobras refining plant)	Large industry
1970	Industrial district (Cinco)	Large industry
1976	Industry (FIAT plant)	Large industry
1979	BH Shopping Center	Large service and commerce
1991	Minas Shopping	Large service and commerce
1991	Del Rey Shopping	Large service and commerce
1996	Diamond Shopping	Large service and commerce
1992-2002	METRO	Accessibility

total population (4,357,942 inhabitants) in 2000, according to the national statistics foundation of Brazil, IBGE.

The lattice or grid of the workspace contains 795 rows and 786 columns, which contain more than 600,000 cells. However, only 288,592 of these cells are within the modeled area composed of the constituent municipalities (Figure 8-1). In the first year of the simulation, 1897, 33 individual cells, or 0.01% of the total were occupied by "institutional buildings". These represent initially present actors, such as a municipality's central buildings and urban equipments. In the literature, these are usually described as seeds (Clarke, Hoppen *et al.*, 1997).

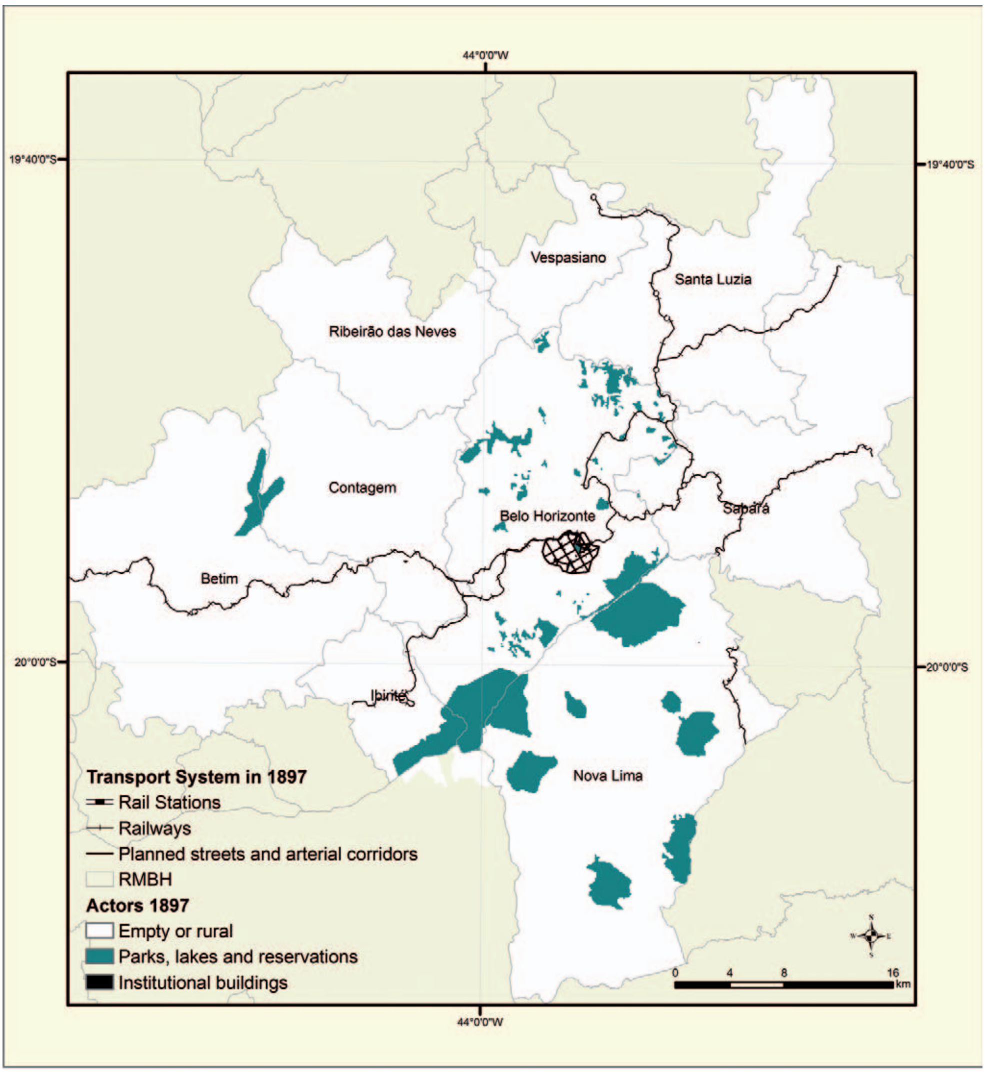


Figure 8.1 Workspace (1897) initial occupation. Source: elaborated by the author.

Around 3%, or 9,892 cells, were already occupied as parks, green areas and conservation units (UCs) in 1897 (see also item 0, p. 162). These cells are called 'excluded areas' by Clarke *et al* (1997).

8.3.2 Exogenous demand

The applied model falls into the "exogenous quantity" classification suggested by Pontius (2008) in which the modeler provides the quantity of each category at each time-step.

The demand of actors (cells) for all other years was estimated based on data from 2000. According to the results described in chapter 7, the spatial allocation for 2000 allocates 5,850 (6.3%) cells of 86 meters by 86 meters to high-income residential areas, 31,672 cells (34%) to average-income areas, and 55,612 cells (59.7%) to low-income areas.

Proportions are maintained and scaled to 1897 data according to the population of each municipality at each time period to obtain the number of cells (demand) per year (population growth is derived from Table 4-2). This implies constant income distributions throughout the previous century. However, we believe that while the structure of the distribution has not changed significantly, the level of income has risen. As discussed in the scenario results (see item 8.9), even a large change in the proportion of each income-level group does not radically change the results.

The reason that the number of cells is calculated at the municipality level is that, in Brazil, municipalities provide a large proportion of local services. The level of organization and available resources is highly differentiated among neighboring municipalities. Thus, the provision of local services, from basic healthcare to education sanitation, and streets and pavement is also rather different. Having, for example, a lower offering of services and consequently cheaper land attracts proportionally more low-income residents (Furtado, B.A., 2006).

8.3.3 Accessibility

In addition to considering neighborhood effects, the model also evaluates accessibility parameters as described in chapter 3 (see item 3.3.2). In the proposed model, accessibility to five different types of features is included: (a) planned arterial streets, (b) roads, (c) railway lines, (d) train stations, and (e) metro lines. Each might influence the neighborhood with a different strength at various distances.

In the proposed simulation, railway lines are only illustrative and have no impact on actors. Train stations, however, influence locational decisions and function as points of concentration for urban functions. They were added to the simulation at the same location and time of actual implementation. Therefore, in 1897, planned arterial streets, railway lines, and train stations are already in place. In 1951, roads that connect the city to São Paulo, Rio de Janeiro, Brasília, Vitória, northern areas of the state (and Pampulha), and the city's Road ring were enhanced or built. In 1965, avenue Cristiano Machado, in the northern part of the city, was added. From 1992 to 2002, the metro line was constructed linking the Eldorado region to Vilarinho, through the central area of the city of Belo Horizonte (see also Table 8-2).

Table 8.3 Projected annual growth rate of population for Minas Gerais state, Brazil. Source: CEDEPLAR/UFMG – 2007

Year intervals	Annual growth rates (%)
2000/05	1.16
2005/10	1.04
2010/15	0.90
2015/20	0.74
2020/25	0.60
2025/30	0.49
2030/35	0.40
2035/40	0.30
2040/45	0.21
2045/50	0.12

Table 8.4 Number of cells in the alternative scenarios

	Year 2050 - Actors - Scenarios (number of cells)	
function	**Scenario 1 - same proportion**	**Scenario 2 - more equal**
High-income res. area	6765	8699
Average-income res. area	35100	53305
Low-income res. area	64668	44529

8.4 Possible scenarios: data

To gain insight into possible future spatial configurations, two scenarios are modeled. The scenarios differ in terms of public policy and urban planning measures, such as new nodes, transport links, or land developments.

Both scenarios refer to the year 2050 and are based on the population projection provided by CEDEPLAR, a demographic research center associated with the Federal University of Minas Gerais (UFMG). The decreasing rates of growth can be seen in Table 8-3.

Both scenarios contain the same number of actors (4,908,002 inhabitants in 2050). In scenario 1, the income distribution across the population is kept constant. In scenario 2, a change in proportions takes place towards a more equal income distribution. This change is implemented so that the share of high-income actors increases from 6.1% to 8.0%, and the share of low-income actors decreases from 58.3% to 40.8%. Consequently, an increase in the proportion of actors in average-income residential areas from 33.2% to 48.8% occurs (Table 8-4; Figure 8-2).

8.5 Model application time scheme and validation

The model application and validation includes comparisons of simulated and real data for various years, and is shown in Table 8-5.

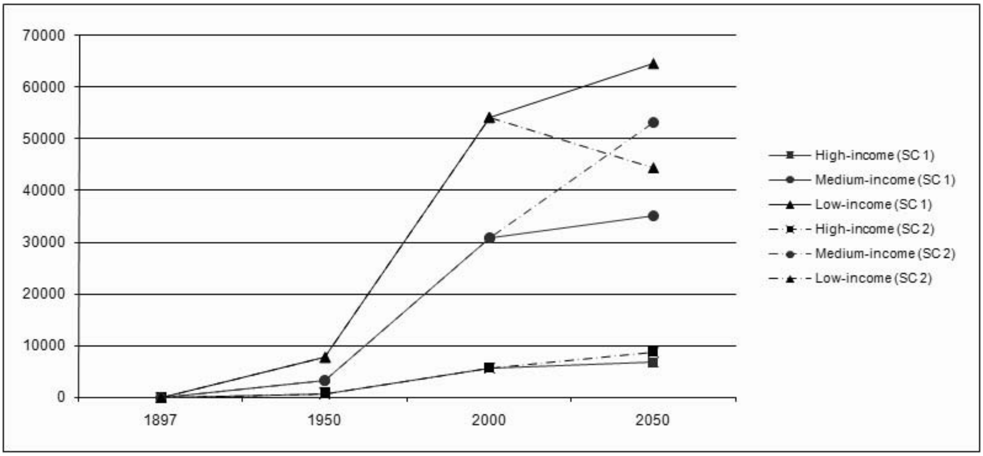


Figure 8.2 Diagram of evolution of actors on both possible scenarios

a) The simulation starts **1897** when planned arterial streets were built. Only 33 known institutional cells are introduced. The aim is to mimic the reference map of **1991**.
b) The parameters are chosen based on theoretical considerations (discussed in the next section). Their magnitude is adjusted via changes in the parameters over manual iterative runs of the application from 1897-1991. The simulation is run once. A parameter change takes place, and the simulation is run again. Results are observed, and more changes are applied. In the literature, this process is called calibration (see item 3.5).
c) **In 1991**, the result of the simulation is compared to actual observations described in chapter 7. The decision of whether to accept model parameters as appropriate is based on visual inspection of the resulting map. The accepted parameters are used for projection for the years after 1991.
d) The simulation (always starting from the initial year of 1897) continues to **2000**.
e) Results are then compared with the 2000 validation map (validation part I).

Table 8.5 Model time scheme and validation

	1897	1991	2000	2007	2050
simulation years					
iterative choice of parameters					
numerous runs aiming at reference year 1991					
comparison year 1991					
and establishment of parameters as definitive					
validation I (simulation x actual obs.)					
actual observation available					
validation II (simulation x price)					
price data available (2007)					
sensitivity analysis					
scenarios					

f) The simulation continues to 2007. The relative prices map (which is also a result of the simulation; see item 3.2, p. 60) is compared with the observed prices map (presented in chapter 7). This is validation II: prices, which is presented in chapter 9.
g) A sensitivity analysis then tests neighborhood rules. The simulation is run again with adjusted parameter settings and results are compared to observed situations for 1991.
h) Finally, the simulation is run to 2050, according to the two proposed scenario alternatives (outlined in the previous section) to provide general insights into future potential spatial configurations of actors.

8.6 Choice of parameters

8.6.1 Neighborhood (N)

The neighborhood effect discussed in this section falls in the six (I to VI) 'general rule shapes' proposed by Hagoort (2006, p. 69). According to Hagoort, neighborhood effects might have:

a) (I and II) a decreasing net positive (or negative) relation that becomes neutral with distance;
b) (III and IV) a net negative (or positive) relationship that switches to a net positive (or negative) relationship and then becomes neutral; and
c) (V and VI) an increasing net positive (or negative) relation that switches to a decreasing relationship, and becomes neutral.

Allocation of each general rule to the modeled actors closely follows the proposal of Allen (1997) and Hagoort (2006). Although general rules and shapes can be observed, Hagoort stresses that there are differences of intensity among different regions.

The parameters ($N_{s(c),s'(c'),d(c,c')}$, see chapter 3) depicted in Table 8-6 are chosen iteratively, manually, and according to the understanding established in the calibration of the model (see item 3.5, p. 69), until the choices that provide the best results (discussed in the next section) are reached. A number of observations can be made regarding the resulting parameters.

In general, a small number of rules (represented by parameters) are sufficient for providing reasonable results (discussed in the subsequent section). Consistently, the values for high-income

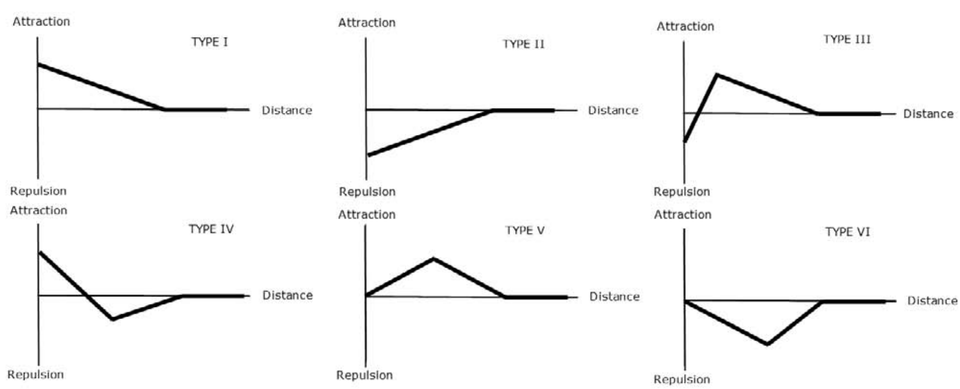


Figure 8.3 Adaptation of Hagoort's six general rules (2006, p. 69)

areas are higher than values for average-income areas, which, in turn, are higher than those for the low-income areas, which implies that preferences are related to income-level, as shown in chapter 6.

The inertia value – at distance 0 – for all rules is much higher than the impact of immediately adjacent cells (distance 1) (Hagoort, 2006).

Rule type I proposed by Hagoort (2006) applies to the impact of residential actors on attractiveness for themselves and the impacts of 'small service and commerce', 'large service and commerce', and 'institutional buildings' on residential use – that is, they all have decreasing net positive relationships.

Rule type II is suitable for describing the attractiveness of large industry – that is, it has a decreasing net negative relationship.

Regarding attractiveness between actors, both high-income and average-income actors attract each other (type I). However, average-income actors are attracted much strongly than high-income actors. Proximity to low-income actors has a negative impact on attractiveness for high-income actors (type II) (see chapter 6).

Another rule, not mentioned as a 'general rule' by Hagoort (2006), implies increasing influence starting at further distances. In this case, an actor would like to locate near another, but, not in its immediate vicinity[86]. This is the case for the impact of 'schools and campuses' on residential use.

For impact on 'small industry', a mix of influences is depicted. The stronger attraction (type I) is for low-income actors; for average-income actors, attraction starts at a farther distance. For high-income actors, attraction is primarily negative but eventually becomes positive (type III).

Airports, water and protected natural areas (reservations) are not key elements in this study. The only airport in the city of Belo Horizonte was built along with the entire 'Pampulha' development project. From the 1940s on, it was intended to be an attractive point. However, this goal was not successfully achieved (see item 4.3 and Villaça (1998). Although open space in developed countries (Yang and Fujita, 1983; Anderson and West, 2006), is restricted[87], water and natural areas in Belo Horizonte do not have a scenic appeal and these factors were not statistically significant in the preliminary exercises of chapter 6. However, their occupancy of space is important because it creates a restriction on other actors.

Institutional buildings impact other actors because they are considered central to urban tissue, by attracting other actors to its surrounding regions.
For the actor 'in the market', which represents vacant, speculative land, there is a lack of information such that only the inertia effect is imposed.

The table with the influences of other actors (small service and commerce, small industry and 'in the market) that have a supporting role in the modeling of residential actors is provided in an annex (Table 11-4).

8.6.2 Accessibility (A)

The parameters of the accessibility function allow for the influence of multiple types of infrastructure (see item 3.3.2). For the proposed application, four types are included. 'Stations' are point locations that refer to actual railway stations. However, more than by facilitating accessibility itself, stations play a role similar to institutional buildings in the sense that they represent a centrality effect. The role of stations as transport nodes is currently very limited

Table 8.6 Parameters of neighborhood influence ($N_{s(c),s'(c'),d(c,c')}$)

	High-income											
Distance (in cells)		1.00	1.41	2.00	2.24	2.83	3.00	3.16	3.61	4.00	4.12	4.24
High-income	95.00	14.00	5.96	1.00	0.96	0.86	0.83	0.81	0.73	0.67	0.65	0.63
Average-income	17.92	5.21										
Low-income	-5.66	-1.65										
Small service and commerce	14.00	4.30	1.84	1.68	1.61	1.45	1.40	1.35	1.23	1.12	1.08	1.05
Small industry	-2.83	-1.08	-0.35	0.68	0.93	0.83	0.81	0.78	0.71	0.65	0.63	0.61
In the market												
Large industry	-1.00	-0.50	-0.29									
Schools and Campuses	0.00	0.00	0.00									
Water and Reservation Areas												
Large Service and Commerce	14.00	12.25	11.53	10.50	10.09	9.05	8.75	8.47	7.69	7.00	6.78	6.58
Institutional buildings	12.00	9.75	8.82	7.50	6.97	5.64	5.25	4.88	3.89	3.00	2.72	2.45
Airports												

	Average-income											
Distance (in cells)		1.00	1.41	2.00	2.24	2.83	3.00	3.16	3.61	4.00	4.12	4.24
High-income	55.66	6.96										
Average-income	85.85	14.00	3.97									
Low-income	0.00	0.00	0.00									
Small service and commerce	5.90	2.20	0.67									
Small industry												
In the market												
Large industry												
Schools and Campuses												
Water and Reservation Areas												
Large Service and Commerce	9.70	4.85	2.84									
Institutional buildings	20.75	7.52	2.04									
Airports												

	Low-income											
Distance (in cells)	0.00	1.00	1.41	2.00	2.24	2.83	3.00	3.16	3.61	4.00	4.12	4.24
High-income												
Average-income												
Low-income	67.00	0.55	0.32									
Small service and commerce	1.20	0.60	0.35									
Small industry	3.77	2.36	1.77	0.94	0.61							
In the market												
Large industry	-0.08	-0.04	-0.02									
Schools and Campuses												
Water and Reservation Areas												
Large Service and Commerce	3.00	1.50	0.88									
Institutional buildings	20.75	16.86	15.25	12.97	12.05	9.75	9.08	8.45	6.72	5.19	4.71	4.24
Airports												

																		Hagoort Rule Type
4.47	5.00	5.10	5.39	5.66	5.83	6.00	6.80	6.32	6.40	6.71	7.00	7.70	7.21	7.28	7.62	7.81	8.00	
0.59	0.50	0.48	0.44	0.39	0.36	0.33	0.32	0.28	0.27	0.22	0.17	0.15	0.13	0.12	0.06	0.03	0.00	I
																		I
																		II
0.99	0.84	0.81	0.73	0.65	0.61	0.56	0.54	0.47	0.45	0.36	0.28	0.26	0.22	0.20	0.11	0.05		I
0.57	0.48	0.47	0.42	0.38	0.35	0.32	0.31	0.27	0.26	0.21	0.16	0.15	0.13	0.12	0.06	0.03		III
																		II
															0.07	0.19	0.30	*
6.17	5.25	5.08	4.58	4.10	3.80	3.50	3.36	2.93	2.79	2.26	1.75	1.63	1.38	1.26	0.67	0.33		I
1.94	0.75	0.53																I

4.47	5.00	5.10	5.39	5.66	5.83	6.00	6.80	6.32	6.40	6.71	7.00	7.70	7.21	7.28	7.62	7.81	8.00	Hagoort Rule Type
																		I
																		I
																		I
													0.20	0.43	1.55	2.20	2.83	*
																		II
															0.01	0.03	0.05	*
																		I
																		I

4.47	5.00	5.10	5.39	5.66	5.83	6.00	6.80	6.32	6.40	6.71	7.00	7.70	7.21	7.28	7.62	7.81	8.00	Hagoort Rule Type
																		I
																		I
																		I
																		II
															0.01	0.03	0.05	*
																		I
3.35	1.30	0.91																I

Table 8.7 Parameters of accessibility: distance decay ($a_{y,s}$) and relative importance ($w_{y,s}$)

	High-income residential area		Average-income res. area		Low-income residential area	
	Distance decay	**Relative importance**	**Distance decay**	**Relative importance**	**Distance decay**	**Relative importance**
Stations	10	0.4	10	0.32	5	0.25
Railway	0	0	0	0	0	0
Planned streets and arterial corridors	10	0.9	50	0.29	15	0.15
Metro	10	0.05	10	0.1	20	0.2
Roads	20	0.01	20	0.05	10	0.01
	Small service and commerce		**Small industry**		**In the market**	
	Distance decay	**Relative importance**	**Distance decay**	**Relative importance**	**Distance decay**	**Relative importance**
Stations	10	0.4	8	0.7	30	0.35
Railway	0	0	3	0	0	0
Planned streets and arterial corridors	10	0.8	2	0.2	100	0.4
Metro	10	0.2	6	0.1	30	0.2
Roads	10	0.1	6	0.4	100	0.2

because no passenger or freight services are present, except for one 12-hour day trip to Vitoria, a coastal city in the state of Espírito Santo, which has limited importance in Brazil88.

The other three types of infrastructure are: (a) streets originating from the original design that was implemented in 1897 (see chapter 4), (b) metro lines that were gradually created from 1985 onwards (power of this line is limited because it is a single line that runs through already densely occupied areas), and (c) roads. These types of infrastructure represent general features of the city. Although they are more specific than a simple variable like 'distance to CBD', they are not sufficient in providing detailed access information at the level of the neighborhood. Further detailing this accessibility information with a historical study of their implementation would provide better results, but it would also hamper generalization of the results to other case studies. The parameters of Table 8-7 are chosen such that:

(a) For high-income actors, the most important factor is the 'formally', planned city represented by the original central ring that had acquired the status of capital or political center (outer localities were deemed rural, suburban areas). Stations are also important for municipalities other than Belo Horizonte. Metro lines are less important, given the ownership of cars by most households in 'high-income' areas. Roads, although important, are responsible for inter-city connections, rather than intra-urban connections that are the focus of this study.
(b) Average-income actors value both planned streets and stations that give them access to the city.
(c) Low-income actors value accessibility of stations the most, followed by the metro line, and then the original planned streets.

Table 8.8 Parameters used for prices

Function (s)	µs	τ
High-income	8	0.1
Average-income	3.2	0.24
Low-income	0.8	0.95
Small service and commerce	10	0.05
Small industry	2	0.3
In the market	8	0.24

8.6.3 Price (P)

The prices are implemented in the model according to the description in item 3.2 (p. 60).

As explained, according to equation 32 ($^{t}Pc = \Sigma_{c' D(c)\ s} f(.)$)), the first stage is to calculate prices of land based on the actual configuration of actors. The impact of each actor on prices is given by parameters (μ_s), as depicted in Table 8-8. When calculating the transition potential for the next time-step, according to equation 30, ($^{t}Tp_{s,c} = [(1+e)^{t}N_{s,c}.^{t}TA_{s,c}] - [\beta\ e' + (1 - \beta)\ ^{t}P.\ \tau_s]$)), the affordability of different actors concerning land prices is weighted by τs (Table 8-8).

Parameters are chosen iteratively in a trial-and-error procedure as described in the calibration section (item 3.5). They refer to the relative differences in empirically obtained prices for the city of Belo Horizonte (see Figure 7-7).

The radius used for the calculation of ($D_{(c)}$) is 4 cells.

8.6.4 Comparison with the reference map to confirm parameter choice

The results of the simulation, using the above parameters, provided results for every 10 years, starting at 1900[89] (Figure 8-4). Figure 8-5 presents the results for 1991 in detail, which are compared to the reference map produced in chapter 7 (Figure 7-1).

As expected, given the chosen parameters, the allocation of actors is scattered throughout the study area and is polycentric in the initial years. However, the model predicts clustering near the planned area. The comparison with actual initial development (see Figure 4-2 to Figure 4-7) shows that in the simulation the central ring is occupied earlier. As discussed in chapter 3, it is rather difficult to simultaneously simulate attraction and repulsion. As time unfolds, however, the pattern of occupation observed in the simulation is similar to the actual pattern of occupation.

The change starting in the 1950s captures the expansion of the number of inhabitants[90]. These years also show the expansion to the west of mainly average-income residential areas, which the model realistically predicts.

The clustering of high-income actors in the center, surrounded by a large patch of average-income actor occupation, along with external, more dispersed, conurbated low-income residential areas, is similar to the actual, observed pattern.

As described in chapter 7, metrics of the results are used with visual analysis to facilitate comparison of the results and reference maps. Because the simulated and reference maps do not contain the exact same actors (the simulation contains small service and commerce, small industry and 'in the market', which are not present in the reference map), the comparison is only indicative and merely support visual inspection.

Table 8-9[91] shows that the general structure depicted in the reference map of 1991 (on the left) is present in the map of results (on the right). The high-income fractal dimension in

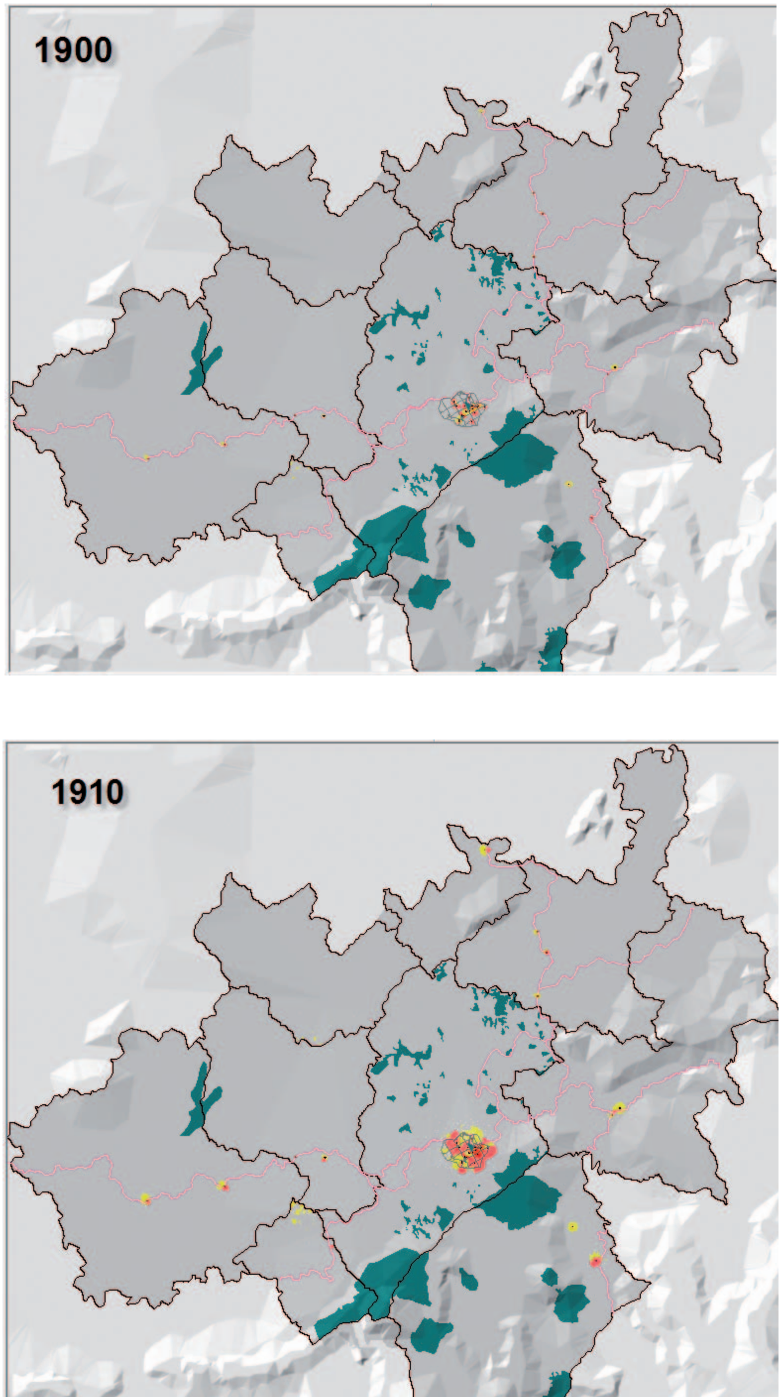
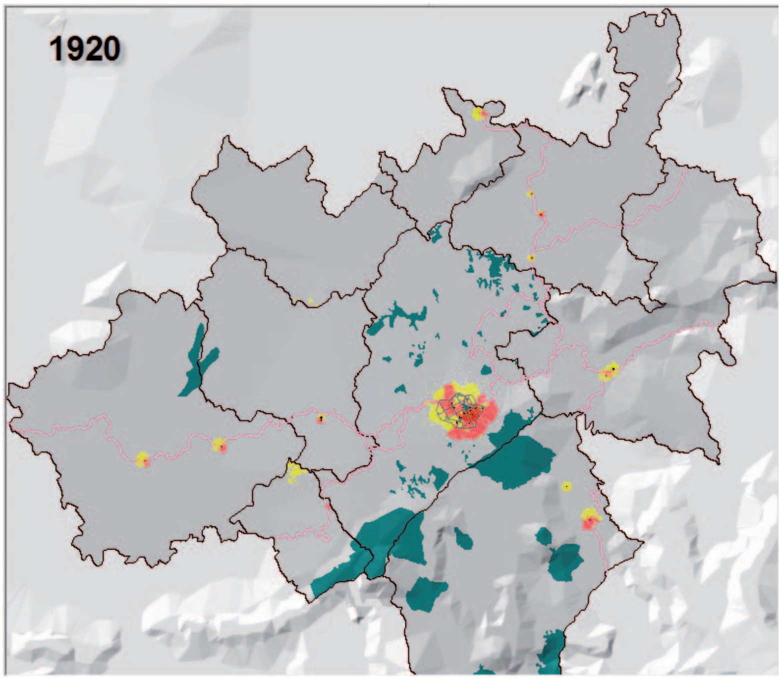
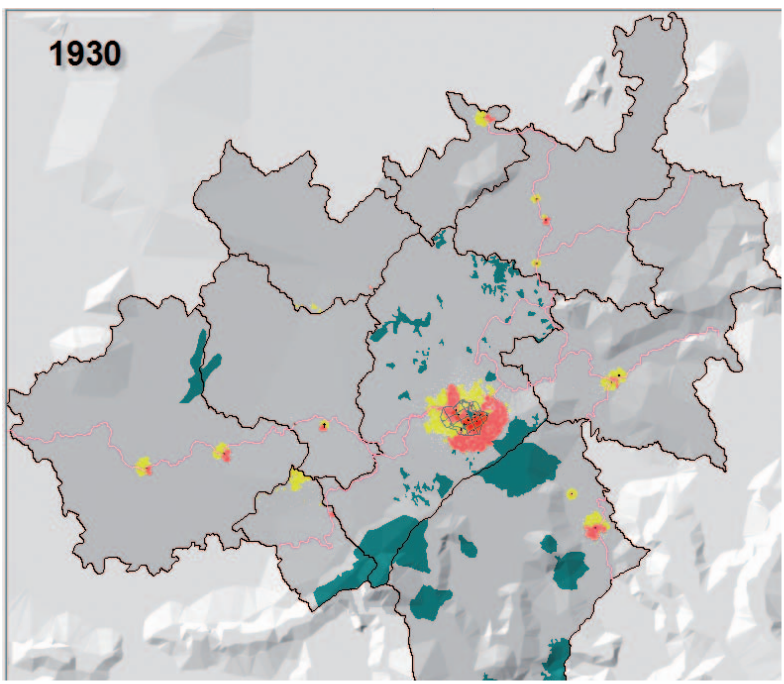
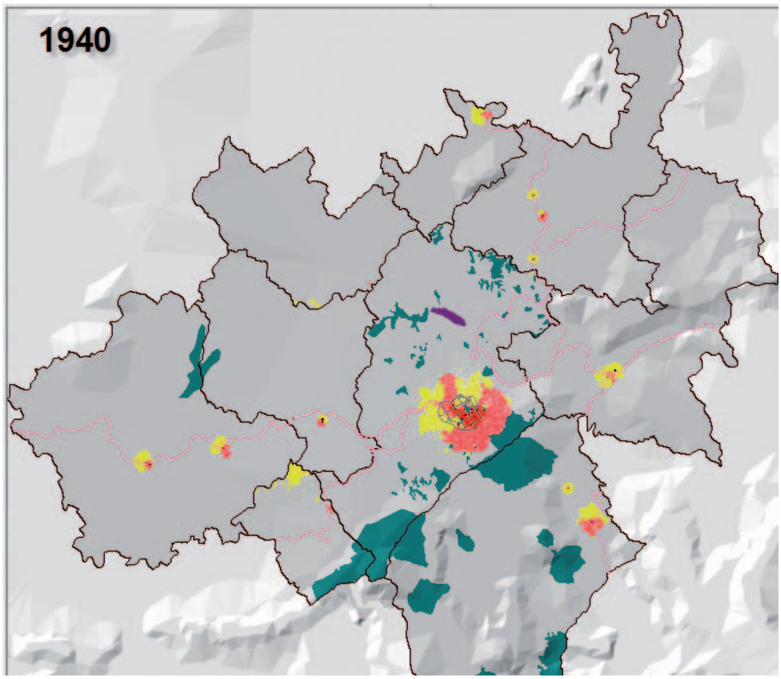


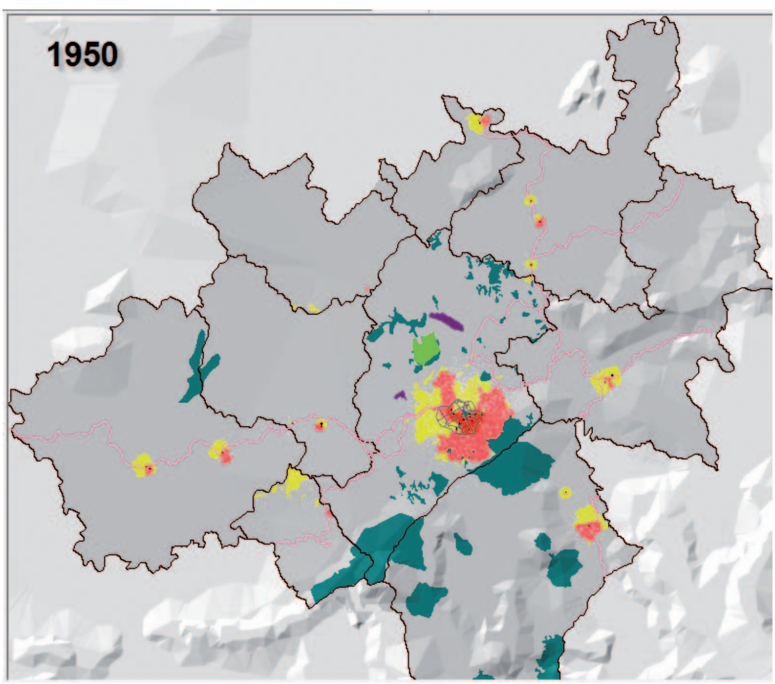
Figure 8.4 Evolution of actors RMBH, every 10 years (1900-1990)



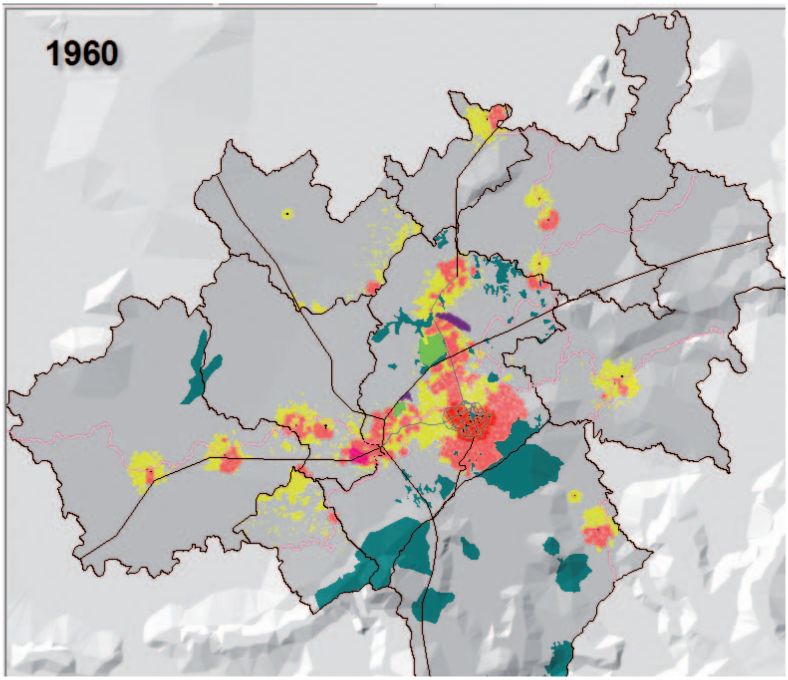


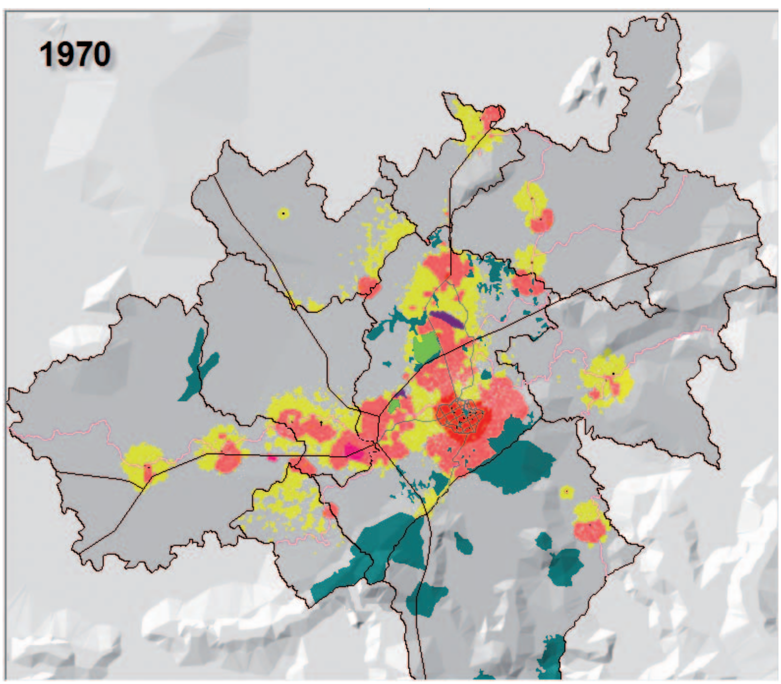
(figure 8.4 Continued)



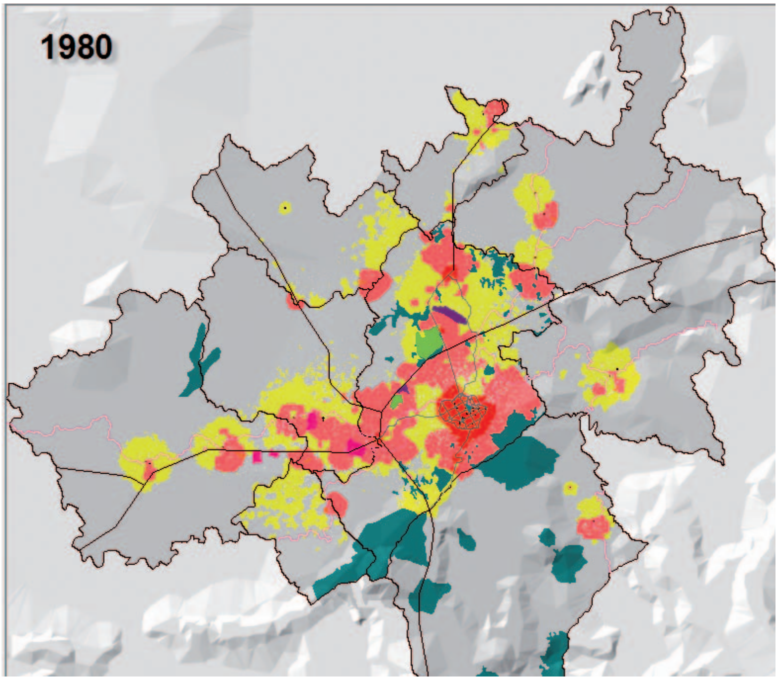


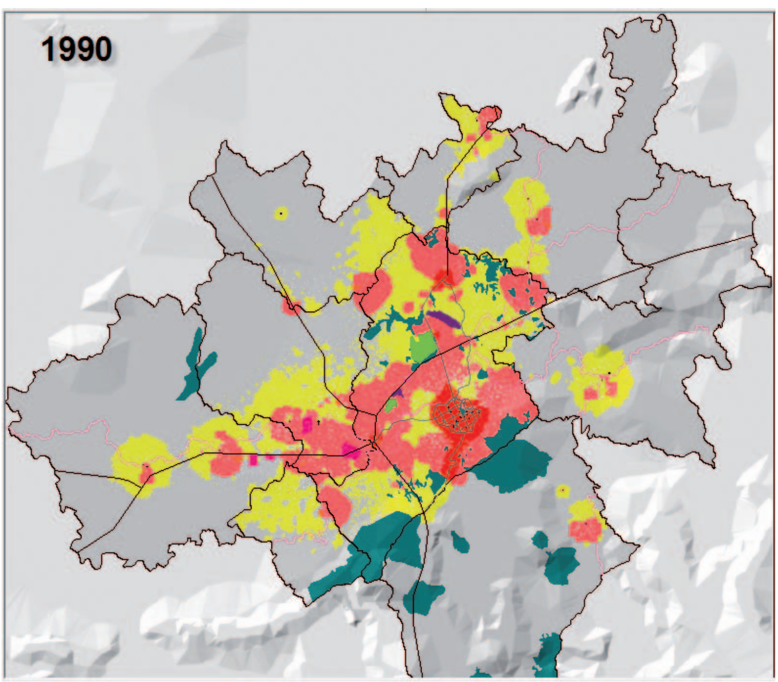
(figure 8.4 Continued)





(figure 8.4 Continued)





(figure 8.4 Continued)

the simulation is close to observed data, but with a slightly higher patch size, which indicates that the simulation map is more clustered than the actual results. For average-income actors, patch size numbers of the simulation are close to reality, but the perimeter is a bit lower in the simulation results, which suggests a lower degree of clustering than in the real data. The low-income configuration seems to differ more from reality than the others because the patch size of the simulated results is much larger than the one observed. This suggests that low-income actors are more clustered in the simulation than they should have been.

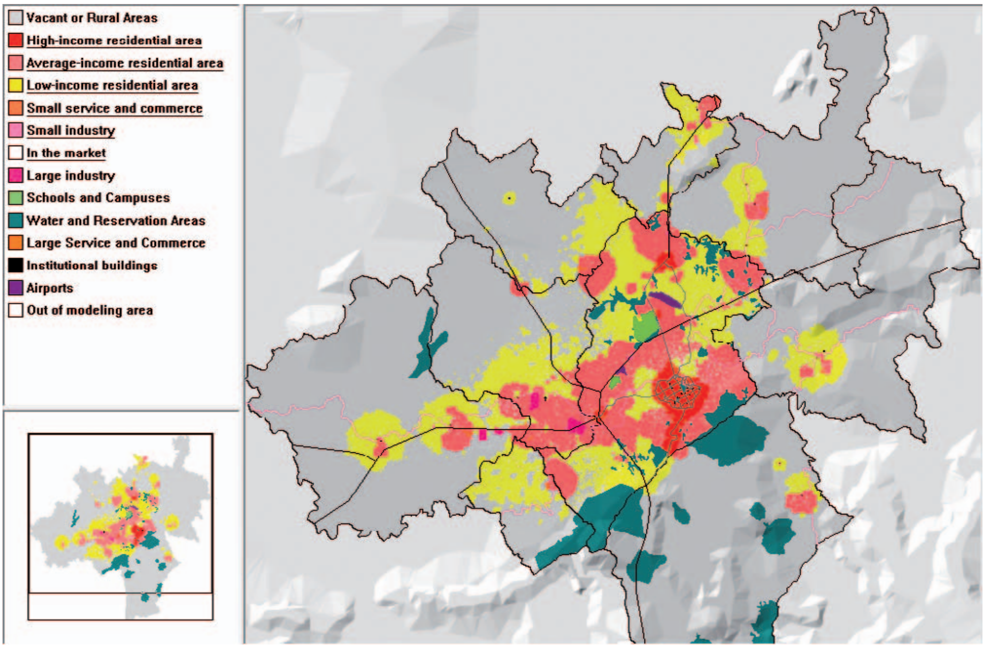


Figure 8.5 Results for comparison, 1991

Table 8.9 Comparison of basic measures of reference with simulation 1991 with defined parameters

Basic measures of actual reference map 1991				**Basic measures of simulation map 1991**			
Fractal dimension				**Fractal dimension**			
High-income	Average-income	Low-income	Global	High-income	Average-income	Low-income	Global
1.327	1.403	1.343	1.364	1.313	1.372	1.413	1.394
Patch size				Patch size			
1863	8594	4229	5712	2338	8377	6868	7215
Perimeter				Perimeter			
641	2900	1168	1775	654	2442	2212	2227
Shape index				Shape index			
3.547	6.941	4.326	5.242	3.342	6.021	6.300	6.067

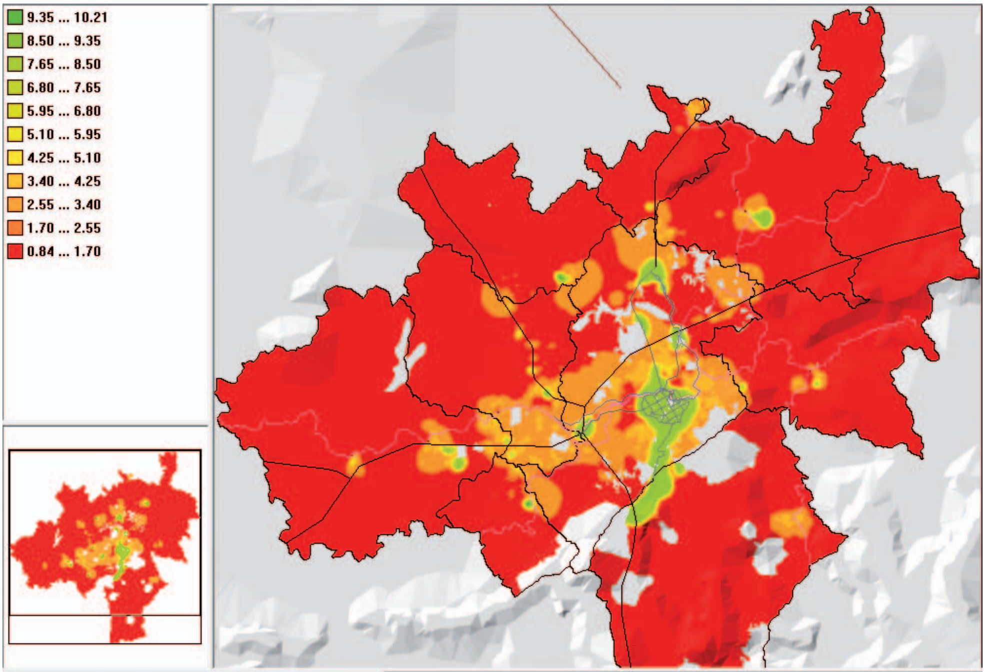


Figure 8.6 Results of relative prices based on simulation, 2007

8.6.5 Relative price configuration and comparison with reference map

This section discusses the extent to which the model accurately predicts the spatial configuration of prices (Figure 8-6).

A visual comparison suggests that the general structure is captured (and validated in item 9, p. 184) with a more expensive area around the original central area and towards the south[92]. Also, less expensive areas are located close to the center in the simulation. However, they do not match the exact locations of observed central, less expensive areas. In the simulation results, they are in the western and northern parts around the central ring, the actual slum conglomerates of Belo Horizonte are three patches south of the central area (Figure 7-1). This is consistent with the character of the model, which aims to describe general patterns rather than cell-to-cell prediction. Deeper idiosyncrasies not captured in the general model would be responsible for their exact locations[93].

A second expensive region is predicted in the northern part of the municipality. This also occurs in the actual price map.

8.7 Validation part I: actors

This section presents the validation of the simulation, and further discusses its results, according to the time scheme discussed above.

The transition between 1991 and 2000 presents the conurbation of low-income areas to the north and a strong expansion of high-income actors towards the southern municipality of Nova

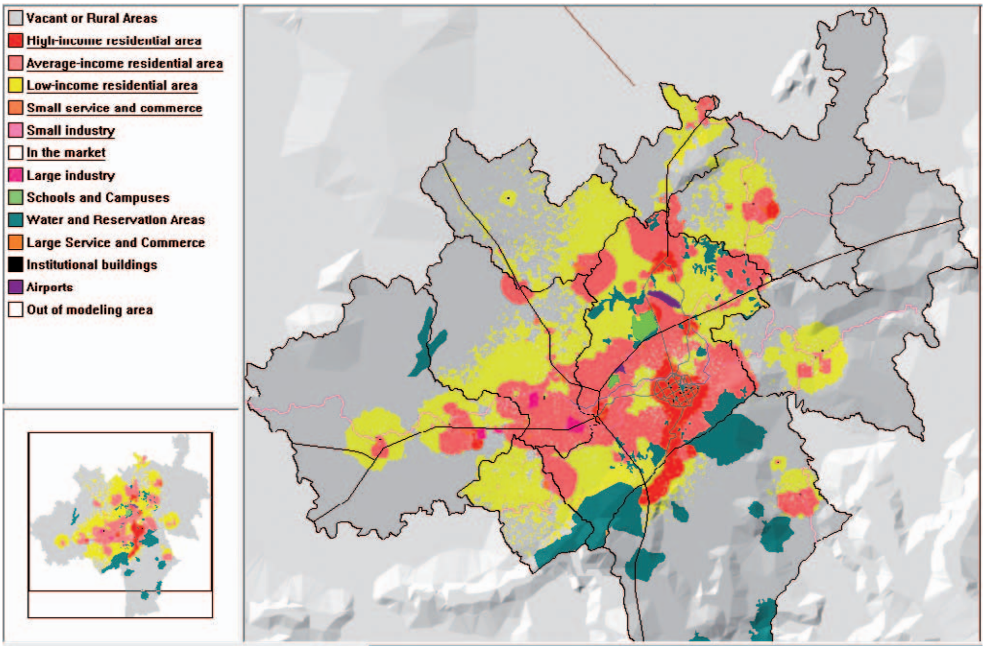


Figure 8.7 Results with price included in the formula, 2000

Table 8.10 Comparison of basic measures of reference with RESULT simulation 2000

Basic measures of actual reference map 2000				**Basic measures of simulation result map 2000**			
Fractal dimension				**Fractal dimension**			
High-income	Average-income	Low-income	Global	High-income	Average-income	Low-income	Global
1.334	1.429	1.386	1.397	1.348	1.370	1.439	1.410
Patch size				Patch size			
1751	14290	12819	12624	3079	8695	12554	10646
Perimeter				Perimeter			
669	4985	3204	3651	906	2412	3958	3240
Shape index				Shape index			
3.680	9.246	6.563	7.294	3.972	5.857	8.033	7.037

Lima. This reflects the construction of the first major shopping center in the late 1970s (see item 4.3) and a continuous preference of residential actors for the southern region (Villaça, 1998; Costa, H., Costa *et al.*, 2006).

In relation to the validation map of 2000, the metrics (Table 8-10) indicate that the simulation performs well and that the similarity between simulated and observed land-use patterns is larger than for 1991. Apart from the patch size and perimeter of average-income areas, all indicators remain close to the actual data of 2000. This suggests that the model that is calibrated on the 1897-1991 period with the chosen parameters is valid – i.e., the model is able to reproduce observed spatial patterns beyond the calibration period (Batty, M. and Torrens, 2005).

8.8 Sensitivity analysis

To gain further insight in how different inputs may influence model outcomes, a sensitivity analysis is performed in this section.

8.8.1 Simulation results without prices

The results of the simulation in which changes in price proposed in chapter 3 are not implemented are presented in Figure 8-8, so that one can compare results that included the change with those that did not (Figure 8-5). Thus the effect of ignoring price developments can be investigated.

The simulation assumes that prices P, and parameters β and τs in equation 30 $^{t}Tp_{s,c} = [(1+e)\ {}^{t}N_{s,c} \cdot {}^{t}TA_{s,c}]\ [\ .e' + (1\ \ )\ {}^{t}P.\ _{s}]$ are 0, so that the second squared box has no effect and the model reduces to the models developed at RIKS (Riks, 2007b).

The main point that can be concluded from visual comparison is that the effect of the neighborhood influence is very strong and generates a structure that is relatively similar in both models with and without prices. The edges and borders of the simulation without prices are more compact and regular.

However, the indicators in Table 8-11 show that the model without price and with the exact same set of neighborhood rules and accessibility parameters perform worse. High-income actors are much more clustered (as indicated by patch size and fractal dimension). Average-income actors have similar numbers. However, the average-income perimeter is much smaller, which indicates less convoluted patterns that are easily recognizable by visual comparison. Low-income areas have a similar fractal dimension and perimeter but a much larger patch size.

This comparison is made with the model calibrated given the existence of a price mechanism. Ideally, the present model should be compared with a model that does not include a price mechanism. However, what happens when there is no price mechanism is that the calibration is biased because it attributes an attraction or repulsion based on implicit price influence. It does

Table 8.11 Basic measures of result simulation without price, 1991

Basic measures of result simulation without prices 1991			
Fractal dimension			
High-income	**Average-income**	**Low-income**	**Global**
1.406	1.342	1.284	1.311
Patch size			
3277	8223	7593	7630
Perimeter			
1184	1865	1498	1618
Shape index			
5.163	4.773	3.957	4.308

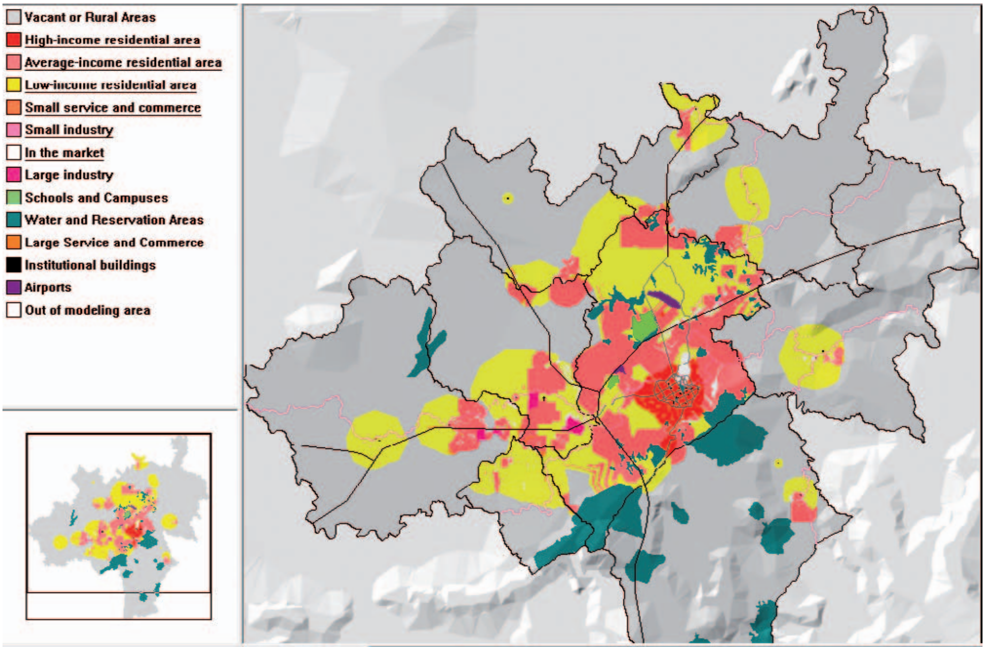


Figure 8.8 Results of the simulation without price changes, 1991

not differentiate between repulsion for price and repulsion for preference. It is not that a specific actor would not be attracted to a certain area, but just that the actor cannot afford it. This is a central argument in the simulation: parameters are more likely to conform to reality when they include both facets of influence (see chapter 3 for a discussion, see specifically Figure 3-1).

8.8.2 Simulation results with double influence of prices

To better understand the influence of prices in the process, the parameters of Table 8-8 (μ_s) are doubled, resulting in the configuration presented on Figure 8-9.

In this case, equation 30 becomes:

Equation 43

$${}^tTP_{s,c} = [(1+e)\ {}^tN_{s,c}.{}^tTA_{s,c}] - [\beta.e'+(1-\beta)\ {}^tP.2\tau_s].$$

As expected, the stronger the impact of price, the stronger the effect of succession and a larger part of the low-income residential area is expelled to the outskirts of the conurbation patch. This happens because the value of τ is smaller for high-income actors and higher for low-income actors (see Table 8-8). However, even with these strong impacts of price, the effect of some low-income residential area inertia enables them to remain in central areas, and a smaller portion of them do so. Apart from that and a more convoluted edge, there are no other significant differences.

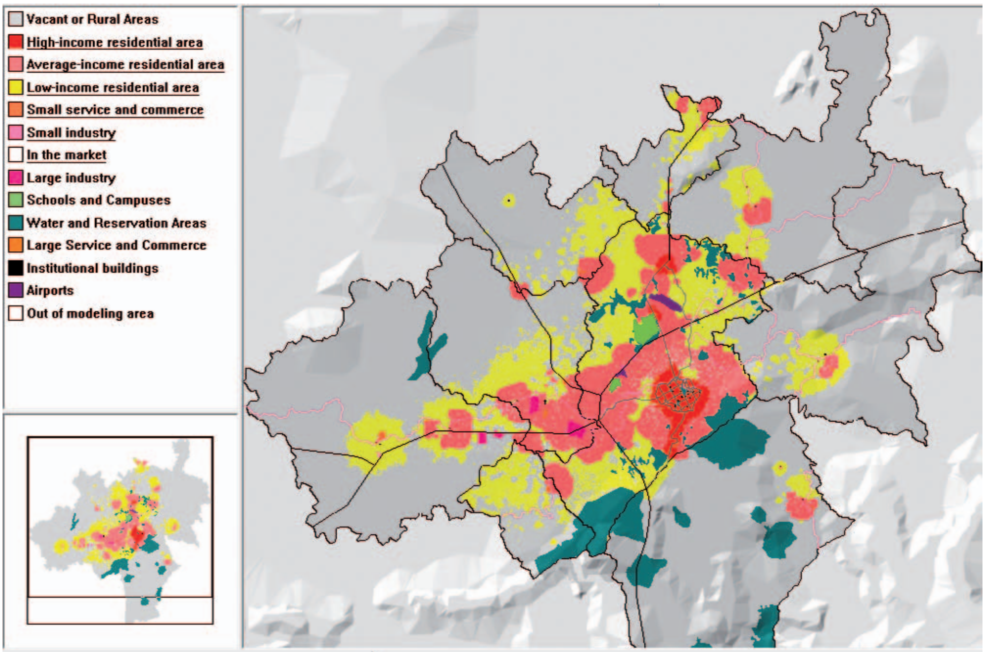


Figure 8.9 Simulation results with double influence of prices, 1991

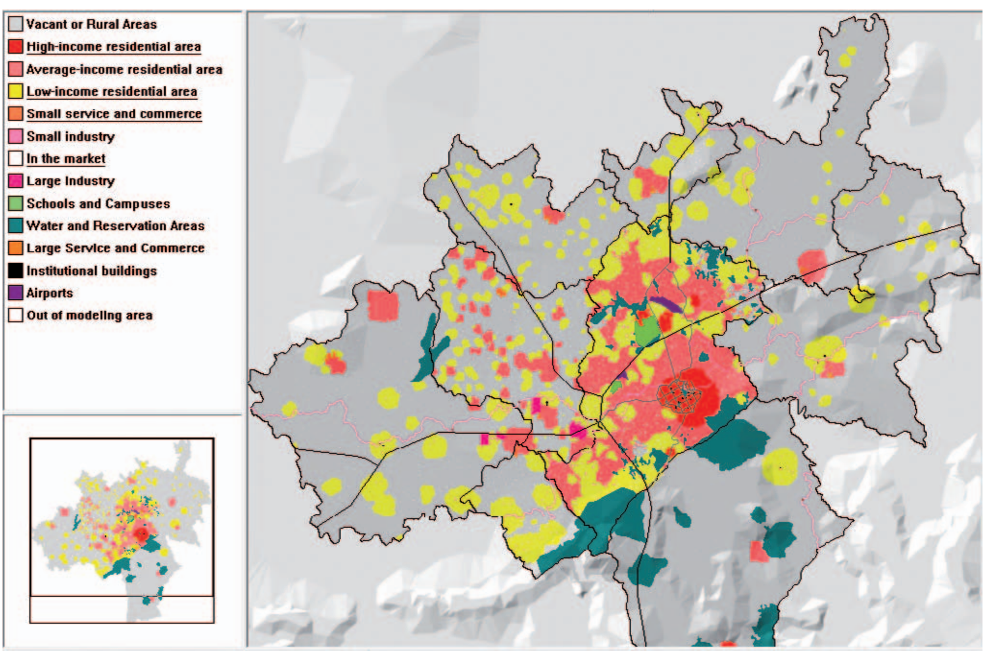


Figure 8.10 Simulation results without accessibility parameters, 1991

8.8.3 Simulation results without accessibility parameters

Simulation results without accessibility parameters (Figure 8-10) show the importance of the centrality represented by these parameters.

From equations 30 and 35,

$^{t}Tp_{s,c} = [(1+e)^{t}N_{s,c} . {}^{t}TA_{s,c}] - [\beta . e' + (1 - \beta)^{t} P_{c} . \tau_{s}]$, and.

$$^{t}TA_{s,c} = \frac{1 - \left[\prod_{s \in S_y} \left(1 - w_{y,s} . {}^{t}A_{c,y,s}\right) \right]}{1 - \left[\prod_{s \in S_y} \left(1 - w_{y,s}\right) \right]}$$

and are set to 0 and, therefore, total accessibility ($^{t}TA_{s,c}$) is 1, which give accessibility no influence on transition potential.

This modification of the model gives results that are very different from the observed pattern (Figure 8-5). Instead of presenting a large patch of low-income residential areas attached to central high-income and average-income areas, it is presented as a number of smaller patches in an archipelago-like configuration. This is also true for the average-income area that is much more dispersed. For the high-income residential areas, the configuration is similar, but its expansion is more irregular and does not start to unfold towards the southern vector as was observed in reality. The impact of institutional buildings and their over-powering of both neighborhood parameters and price affordability seem to restrict allocation to the central, planned area.

8.8.4 Simulation results with similar N parameters for all residential income levels

Another alternative to evaluate the adequacy of the chosen parameters is to test similar parameters for all residential actors. Implicit in this test is the homogeneous treatment of actors like in other CA models, only with one residential actor.

To apply this change, the influence parameter $N_{s(c),s'(c'),d(c,c')}$, from equation 33, $^{t}Ns,c = {}_{c'D(c)} N_{s(c),s'(c'),d(c,c')}$, is set to have similar values for different residential actors (s)[94].

The results in Figure 8-11 demonstrate the importance of explicitly accounting for social heterogeneity.

The pattern observed in this exercise is rather different from patterns in Figure 8-5. The high-income residential area becomes too attached to the accessibility network, especially, to 'roads'. Average-income areas do not seem to conform to any regular pattern, and are subdivided into a larger number of patches. The low-income areas are more clustered than they should be, although they do not form one large external patch.

8.8.5 Simulation without accessibility parameters, with similar N parameters, and without price parameters

As another test, a model in which there are no accessibility parameters (as item 8.8.3 above) and in which the neighborhood influence parameters are similar (as item 8.8.4 above), simultaneously shows a result (Figure 8-12) that resembles a random pattern.

In the example, not even the high-income residential area is clustered around the planned area, nor is there any distinguishable resemblance with observed results.

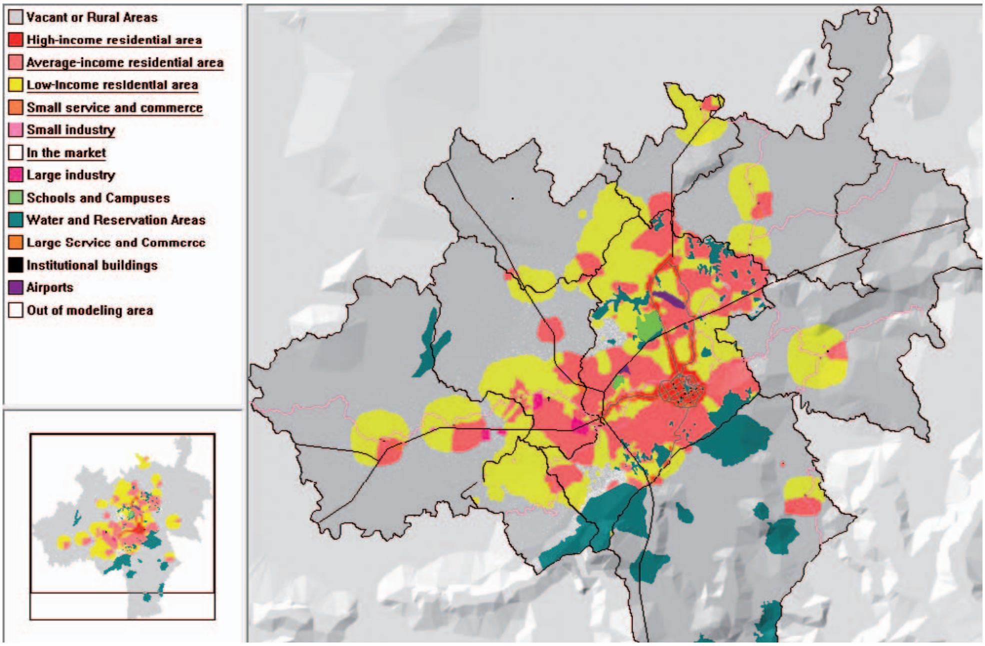


Figure 8.11 Simulation results with similar N parameters for all residential income levels, 1991

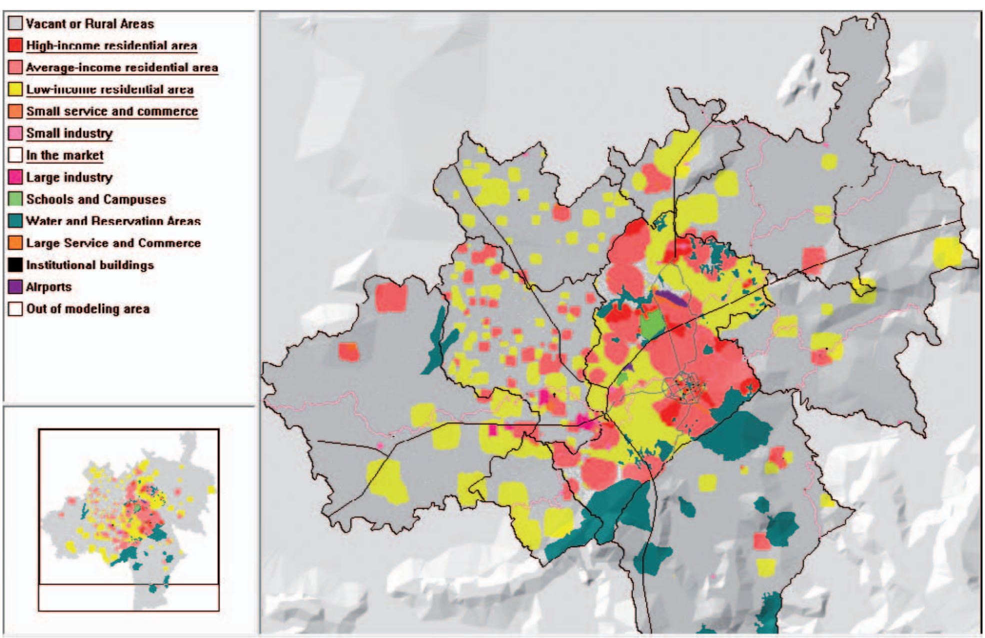


Figure 8.12 Simulation without accessibility parameters and similar N parameters, 1991

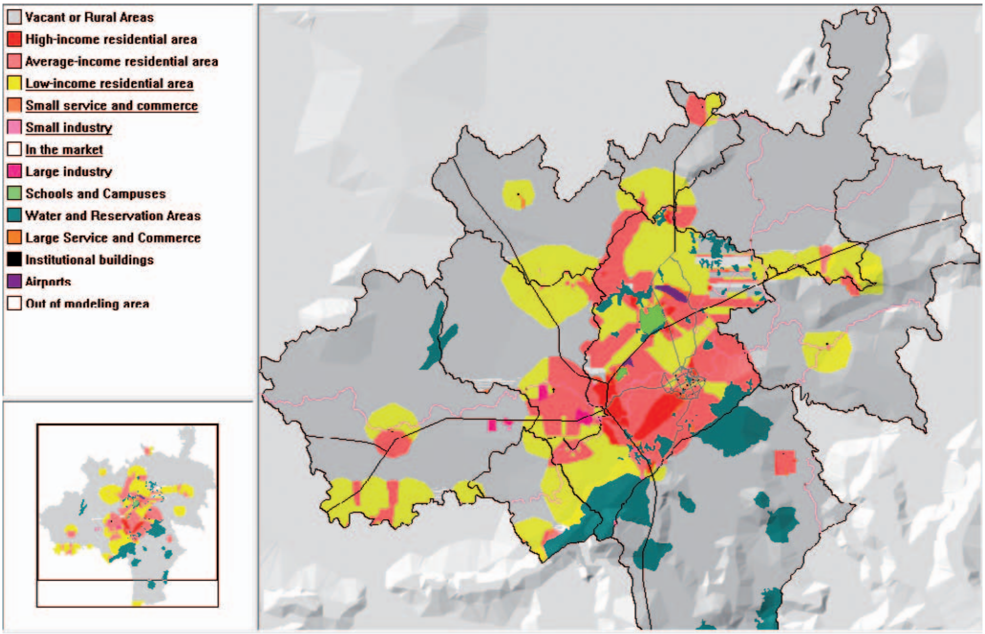


Figure 8.13 Simulation with similar N parameters, no accessibility and no price parameters, 1991

In a final test presented in Figure 8-13, there is no influence of accessibility, the neighborhood parameters have similar influences on each other, and the price mechanism of negative feedback is disabled (see item 8.8.1, above).

The result is even worse than the previous one. It is easy to see that a random stochastic effect is the major effect driving the spatial configuration. Some areas on the borders of the study area present some occupation that has not appeared in any of the other models.

The sensitivity analysis suggests that neighborhood influence is central to the spatial configuration of the city, and prices play an important role in the dynamics of this influence. That is, the higher the influence of prices, the faster is the segregation and succession among different income-level neighbors. Inertia – the tendency of earlier occupants to remain for long periods of time in a given area – is the counter-effect of succession. Accessibility is also an essential item because it organizes the concentration of actors in the city. Finally, results show that an intra-urban analysis cannot fail in differentiating actors by income-level. Exercises that include only urban x non-urban land-use are reasonable when studying urban morphology as a whole, but not when studying its internal configuration.

8.9 Scenario analysis

As explained in item 8.4, two scenario analyses are performed here to provide insights into probable future configurations of urban space and prices in the Greater Area of Belo Horizonte.

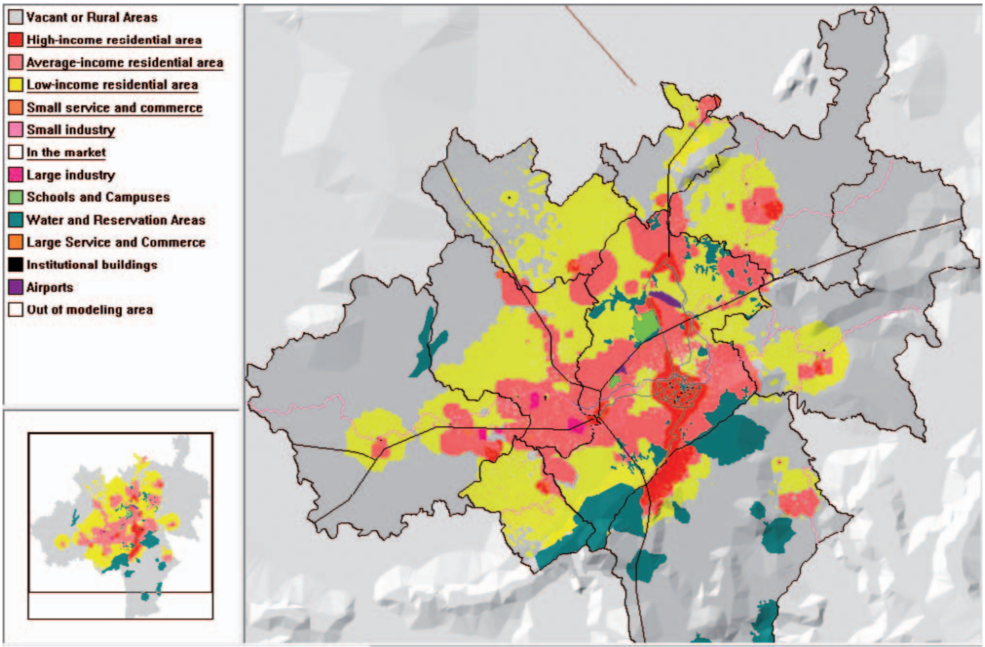


Figure 8.14 Results for scenario 1, 2050

Both scenarios confirm that the general configuration is not likely to change radically (which reinforces the presence of historical factors as key-elements of the market). Moreover, both show the permanence of low-income residential areas near the original central planned areas.

In scenario 1, in which the proportion of each income-level is the same, low-income occupation occurs particularly in external areas of the metropolitan conurbation, making it more compact and dense than what is presently observed.

In scenario 2, there is a relative decrease in low-income residential areas. The results show that the occupation of average-income areas overlaps low-income areas, with the exception of some internal patches that remain unchanged. The allocation of new cells in neighboring municipalities also tends to occur nearer the border of the central city of Belo Horizonte. High-income areas develop mostly next to existent areas, mainly towards the south, which reflects the same pattern observed through the twentieth century and which was described by Villaça (1998).

That the number of actors is different (see Figure 8-2) highlights the succession of higher-income actors overtaking lower-income actors. This results in higher clustering of high-income and average-income actors and allocation closer to the center and the transportation network.

8.10 Concluding remarks

In answer to the general questioning of importance of parameters posed in chapter 3 (item 3.5, p. 69) the dynamics of the observed processes demonstrates that succession does occur and inertia effects are important. This means that higher-income occupation takes over lower-income

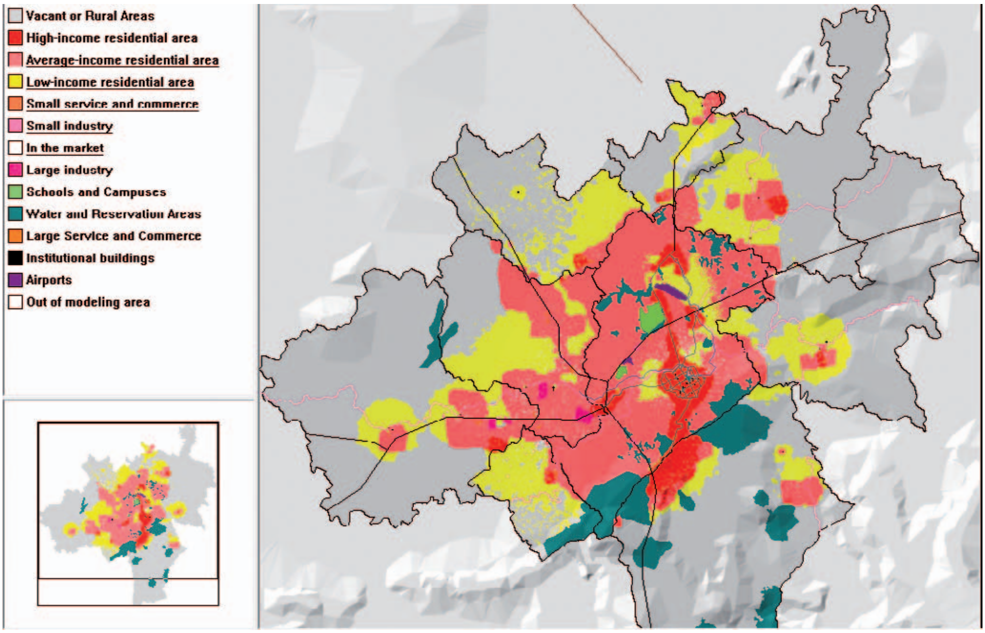


Figure 8.15 Results for scenario 2, 2050

occupation as their numbers expand. Furthermore, inertia of the actors plays an important role in keeping them in places where early occupation occurred. In other words, although most lower-income actors are subject to succession, a small number succeed in retaining early-acquired territory. As the test of doubling price parameters indicates, the higher the economic factor impact represented by relative prices, the smaller the number of urban actors who are able to retain their land.

In relation to influence of accessibility parameters, accessibility has proven to be essential because it brings to the model influences other than a unit's immediate neighborhood. Although a distance decay factor is used, influences on rather distant places, such as at a 100 cell (8.6 km) distance, can still have an effect due to accessibility transportation. Furthermore, accessibility is an influence that is much more precise than distance to CBD because it accounts simultaneously for four different types of accessibility aids (stations, roads, planned arterial streets and metro lines) that vary in the spatial plane according to the actual shape of the influencing object. In the case at hand, for example, stations are considered attraction centers and there are a number of stations that constitute a multicentric continuous influence. This method of modeling is closer (in the theoretical perspective) to the methods of Fujita and Ogawa (1982) and Wheaton (2004), mentioned in chapter 1. Furthermore, the analysis of parameters (neighborhood, accessibility, price) shows (item 8.6) that they have different influences for the various actors. Tests with different parameters show that they are important in establishing the structure of the city.

9 Cross-validation: part II, prices

The aim of this chapter is to further validate both the econometric model in chapter 6 and the CA model presented in chapters 3 and 8, by cross-validating them. This is achieved using input from one model in the other and analyzing the results. In doing so, we are increasing the explanatory power and reliability of the models and gaining further insight into what the specific contributions of each model are.

9.1 Models

The input for the econometric analysis stems from applications of previous models. Both the model and the data (location and independent variables) are drawn from the econometric analysis in chapter 6. From the CA model in chapter 8, the land price map for 2007 is used as a dependent variable.

Two analyses are proposed. One analysis considers the same model as in chapter 6, where independent variables comprise dwellings and city characteristics. The other model is on in which dwelling attributes are excluded, thus generating a regression of city characteristics on price per square meter as the dependent variable.

(a) The first model involves regressing the independent variables on the simulated generated land price data, rather than on the original empirically-collected price data as in chapter 6. The Simulation value is drawn from the land price map (Figure 8-6) at specific locations of the observations (Figure 6-1)[95].
The model of item 6.4 then becomes:

LnPrice = α + β1lnArea + β2Standard + β3Age + β4D_ap + β5Domic_n + β6PCA_activ + β7PCA_innovation + β8Factor_melpub + β9Shop_1500 + β10 Slum_200 + β11 Aven_200+ ε.

Simulation (lnPriceMap) = α + β1lnArea + β2Standard + β3Age + β4D_ap + β5Domic_n + β6PCA_activ + β7PCA_innovation + β8Factor_melpub + β9Shop_1500 + β10 Slum_200 + β11 Aven_200+ ε.

(b) The second analysis is one in which the same database is used. However, the regression is on the natural logarithm of the price per square meter (lnpricearea). This model is proposed because the CA model does not include specificities of the household. In this case, only the neighborhood and city characteristics are used as independent variables. Hence, the variables that control for the dwelling (lnarea, age, standard, d_ap) are excluded. Thus, the original model becomes:

Analysis 1	
Original model chapter 6 and Original model with dependent variable stemming from CA model	
Dependend variable	Independent variables
Lnprice (empirically collected) and LnpriceMap (from CA simulation)	Area
	Standard
	D_ap
	Domic_n
	PCA_activ
	PCA_innovation
	Factor_melpub
	Shop_1500
	Slum_200
	Aven_200
Analysis 2	
Modified model chapter 6 and Modified model with dependent variable stemming from CA model and exclusion of dwellings' characteristics	
Dependend variable	Independent variable
Lnprice/area (empirically collected) and LnpriceMap (from CA simulation)	Domic_n
	PCA_activ
	PCA_innovation
	Factor_melpub
	Shop_1500
	Slum_200
	Aven_200

Figure 9.1 Models' description

$$Lnpricearea = \alpha + \beta_1 Domic\_n + \beta_2 PCA\_activ + \beta_3 PCA\_innovation + \beta_4 Factor\_melpub + \beta_5 Shop\_1500 + \beta_6\ Slum\_200 + \beta_7\ Aven\_200 + \varepsilon.$$

And, using the CA generated data:

$$Simulation\ (lnPriceMap) = \alpha + \beta_1 Domic\_n + \beta_2 PCA\_activ + \beta_3 PCA\_innovation + \beta_4 Factor\_melpub + \beta_5 Shop\_1500 + \beta_6\ Slum\_200 + \beta_7\ Aven\_200 + \varepsilon.$$

9.2 Results, analyses and interpretation

To make the comparison easier, the original results for the spatial model (see Table 6-8) and the Neighborhood matrix (see 6.4.1) are replicated in the first column of Table 9-1 (*Lnprice*). They are followed by results with the dependent variable as the simulated land price (*Simulation*). The results for a regression in which the dependent variable is a randomly generated number are presented in the third column (Random #). Finally, the results for the Simulation land price (*Simulation*) with the other two weight matrices (contiguity and distance) are shown in the columns on the left.

Table 9.1 Comparison of the model from chapter 6 and simulation results as the dependent variable

	Spatial model (0.5)				
		W Neighborhood		W Contiguity	W Distance
	Lnprice	**Simulation**	**Random #**	**Simulation**	**Simulation**
Lnarea	0.758	0.049	-0.001	0.077	0.023
	(79.02)**	(1.01)	(-0.11)	(1.52)	(0.46)
Standard	0.110	0.036	0.012	0.091	-0.011
	(11.61)**	(0.78)	(1.94)	(1.87)***	(-0.25)
Age	-0.004	0.004	0.000	0.007	0.000
	(-8.94)**	(1.65)	0.769	(2.70)**	(0.13)
D_ap	-0.149	-0.014	0.003	0.090	-0.033
	(-10.3)**	(-0.22)	0.359	(1.21)	(-0.49)
Domic_n	0.266	0.220	0.002	0.531	0.254
	(4.44)**	(1.73)	0.109	(2.47)**	(1.63)
Pca_activ	0.499	0.497	0.021	2.295	0.464
	(5.97)**	(3.09)**	0.710	(8.26)**	(2.32)*
Pca_ind_n	-0.248	-0.228	0.033	-1.238	-0.181
	(-2.53)*	(-1.33)	1.601	(-3.95)**	(-0.87)
Pca2_ino_n	0.429	-0.027	-0.001	0.044	-0.101
	(4.78)**	(-0.15)	-0.090	(0.13)	(-0.43)
Shop_1500	0.113	0.613	-0.006	1.138	0.470
	(5.47)**	(5.25)**	-0.549	(10.13)**	(5.87)**
Slum_200	-0.063	0.267	0.023	0.381	0.120
	(-3.72)**	(3.63)**	(2.17)*	(3.91)**	(1.50)
Aven_200	0.038	-0.046	-0.021	-0.056	-0.033
	(3.05)**	(-0.83)	(-2.75)**	(-0.78)	(-0.56)
F_melpub2	0.061	0.121	0.016	0.159	0.057
	(4.61)**	(1.93)*	1.824	(2.30)*	(0.90)
Constant	6.779	0.379	0.498	2.037	0.356
	(646.43)**	(1.58)	(6.82)**	(6.47)**	(1.34)
rho	0.037	0.675	0.001	-0.023	0.802
	(5.69)**	(35.79)**	(-3.21)**	(-1.48)	(27.11)**
lambda	0.594	-0.454	0.464	0.336	-0.244
	(172.64)**	(-7.96)**	0.732	(26.89)**	(-3.17)**
Observations	5512	5512	5512	5512	5512
Log-likelihood	2028.1	-7078.3	4100.0	-7104.3	-6978.7
R-squared	0.689	0.218	0.005	0.208	0.242

Absolute value of assymptotic t-statistics in parentheses
***significant at 10; * significant at 5%; ** significant at 1%

The results in Table 9-1 show that the model with the land price map as the dependent variable (*simulation*) performs better than the regression with the random value as the dependent variable.

The results also demonstrate that although the model with all three matrices performs similarly well, there are differences. The model with the best fit (measured by R-squared and log-likelihood) is the model with the distance matrix (see chapter 6). Not surprisingly, this matrix is constructed using a standard distance pattern (see item 6.4.1).

The model with the contiguity matrix, however, has a greater number of significant variables. This happens perhaps because of the weak spatial link represented by this matrix. The freedom of strong spatial links allows the model to perform better at the individual level of the variables. The models with other matrices, on the other hand, – distance and neighborhood – perform worse with these variables, probably because the structure of the CA model is able to capture strong spatial influences, but not nuances and specificities of hedonic regression. In summary,

Table 9.2 Further comparison of model chapter 6 and simulated price result

	Spatial model (0.5) W Neighbourhood	
	Lnpricearea	**Simulation**
Domic_n	5.864	0.244
	(4.95)**	(2.03)*
Pca_activ	0.362	0.550
	(4.57)**	(3.42)**
Pca_ind_n	-0.342	-0.218
	(-3.64)**	(0.2)
Pca2_ino_n	0.339	-0.035
	(3.52)**	(-0.19)
Shop_1500	0.104	0.635
	(4.73)**	(8.67)**
Slum_200	-0.053	0.267
	(-2.94)**	(3.58)**
Aven_200	0.045	-0.040
	(3.37)**	(-0.73)
F_melpub2	0.069	0.141
	(4.89)**	(2.31)*
Constant	5.864	0.737
	(2018.53)**	(5.01)**
rho	0.055	0.677
	(4.59)**	(31.88)**
lambda	0.522	-0.436
	(70.67)**	(-5.02)**
Observations	5512	5512
Log-likelihood	1627.8	-7081.9
R-squared	0.212	0.217

Absolute value of assymptotic t-statistics in parentheses
* significant at 5%; ** significant at 1%

they have a better general fit, which reflects that the general structure has indeed been generated by the CA model, but misses on the specifics. These indications seem to be reinforced by the high values and significance of *rho* for the distance and neighborhood matrices and small values and non significance for the contiguity matrix.

The variables related to the characteristics of the apartment, i.e., *standard* and *age*, are significant only for the contiguity matrix, probably due to the general attributes of the region where they are located (suggesting that there is a correlation of independent variables with the error, which in turn reinforces the need for an instrumental variables estimation method). On the other hand, the variable that represents the attributes of the domicile and the family in general (*domic_n*) is significant at the 10% level for all matrices. This is expected because neighborhood rules are the main drives of CA models. However, the level of differentiation among neighborhoods captured by the proposed index, in which every neighborhood has its own value (Table 11-3), is far more detailed than the division among high-income, average and low-income areas proposed by the CA model. It is likely that a more detailed design of actors in the simulation would result in higher significance for the variable.

The general attractiveness of activities (*pca_activ*) is captured by the simulation in all matrices. However, the repelling effect of industry (although present in all coefficients) is only significant for the contiguity matrix. The specific dynamics of the city represented by the variable that captures predominance of innovative sectors and education (*pca2_ino_n*) are not present in the results of the simulation, which suggests that, although the model does captures general trends and structures in the long run, it fails to address fast-paced changes. This means that a reliable model for drawing inferences should be constantly updated to account for innovative changes that eventually transform neighborhood characteristics and citizen perception. Alternatively, the inclusion of other actors who could account for such changes in the model could be introduced to ameliorate the results.

Not surprisingly, the effect of large shopping centers is present in the simulation results. This is because the effect has explicitly been included in the CA model as an exogenous *feature*. The negative influence of slums is not captured in the results, even though it is present in the neighborhood effect influence table in the form of a mutually negative effect between high and low-income residential areas. Nonetheless, as mentioned below, this simultaneous effect of attractiveness and repulsion proved to be the most difficult part of the modeling.

In terms of accessibility, the econometric model configuration, in which accessibility is expressed by variable *Aven_200*, is much more detailed than the CA model. Whereas the CA model contains only major axis and planned original streets, the econometric model accounts for all the arterial structure presently implemented. One could go down to all the details of avenue implementation and idiosyncrasies of Belo Horizonte Greater Area's evolution in the CA model. This would definitely improve the accuracy of the results. However it would also limit the generalization of the model and therefore its scientific strength. This applies for cellular automata models in general, according to Brown *et al.* (2005). The central dilemma, for Brown, is that the more constrained and detailed the model, the better the results, but the less capable the model is to explain generalities and trends. Therefore, the modeler might over-fit a model in attempting to explain a specific city or case, rather than attempting to construct a model that can explain a more of a general phenomenon.

The results for the second model, in which the dwelling characteristics are excluded (Table 9-2), show that, except for the index of presence of industry (pca_ind_n), the same variables of the first model are statistically significant.

In terms of goodness of fit, the R-squared for both regressions is at the same level, with the simulation performing slightly better (0.005). However, the log-likelihood is much worse when the simulation results are used.

10 Concluding remarks

This concluding chapter starts with an evaluation balance of the approaches. We then focused on the added value of the research in this thesis. The following section provides answers to the research questions and discusses limitations of the research. Suggestions for further research are provided at the end of this chapter.

In this thesis, two general approaches to the problem of understanding real estate prices are presented. In chapter 6, we presented an econometric model that emphasized local social heterogeneity, given by income-levels and the importance of the neighborhood. In chapter 8, a CA model was applied that studied real estate development within an intra-urban income-differentiated framework.

We concluded that the econometric model performs well and provides insights into what factors contribute (and to what extent) to the formation of real estate prices. It is precise, consistent, and it makes broader inferences possible. It is descriptive in the sense that, given a *status quo*, the model can be readily applied. It does not, however, account for the factors that dynamically evolve into a specific configuration.

The CA model, on the other hand, provides exactly this. It shows that neighborhood characteristics, when complemented with general accessibility information, generate an empirically comparable configuration of actors and prices in the city at the end of a substantial period of time. Comparatively, however, its prediction is less precise because the CA format aims to create a general structure rather than specific cell evolution. Thus, the CA model misses some of the factors captured by the econometric model, such as fast-paced innovative changes and detailed network influences (see chapter 9). As suggested below, some of these factors may be included in a more detailed case study, over a shorter time-span.

10.1 Critical concluding remarks

Real estate markets have been extensively researched in the literature, mainly from an economic perspective. In the context of this burgeoning field, this thesis presented a central focus on the neighborhood-scale as perceived by inhabitants of the city. The neighborhood plays an essential role in understanding the city and its driving forces, like land prices. The choice of neighborhoods as cognitively self-contained units of analysis can be witnessed in the elaboration of the indices in chapter 5, and in the construction of the weight matrices in chapter 6. The CA model, in turn, demonstrates that basically, the sense of attraction for the same functions or the willingness to locate near individuals who are similar accounts for a large part of the urban structure's evolution, and its associated land prices.

A second point that distinguishes the approach of this study is the investigation of rules and relationships explicitly determined by income. Again, this approach is observed in the construction of indices, in the quantile econometric analysis and in the elaboration of the cellular automata model. Not only do income-differentiated research results vary strongly from median results, they also are a key element in explaining the dynamics of urban structure and land prices. This is discussed in chapter 6 (regular and quantile regressions) and in the sensitivity analysis of chapter 8 (see item 8.8.4).

This research potentially provides policy-makers with knowledge about how real estate prices interact with urban tissue. This adds to the policy scope that is currently available for authorities, institutions (Teixeira, 1999), and academia (Macedo, 1996). These models can aid public bodies regarding insights to changes caused by new major infrastructure allocations, large-scale commerce and service facilities, or industrial conglomerates. This goes beyond the results usually available when only cross-sectional econometric analysis is performed. Inferences, however, should not be made on a cell-to-cell basis but rather with regard to the general change of structure.

We believe that this work highlights the importance of considering local influences and the neighborhood as central elements in urban real estate markets in relation to complex urban environments. We also show that spatial and social heterogeneity are inherently embedded in urban structure and its determinative processes. Negative feedback also contributes to processes of urban development, and should be further developed and included in future cellular automata modeling. Social and spatial heterogeneity should not be left out of urban real estate studies, because they appear to be crucial structural elements. We also hope to have further bridged neighboring fields of science (both theoretically and methodologically). And, finally, the investigation of the case study of Belo Horizonte and its metropolitan area provided us with data and a toolkit filled with analytic information that could potentially contribute to city planning and development.

10.2 Research questions revisited

Although the previous section addressed the research questions in a general manner, this section provides more specific answers.

The first question was "do urban real estate price models improve when considering neighborhood identity and social heterogeneity?". We can answer affirmatively. It appears that crucial information is left out of models when social heterogeneity, expressed by the quantile regression in chapter 6, is not incorporated. Furthermore, the neighborhood level, by incorporating the perception space of urban citizens, attaches specific values to real estate prices. The description of a neighborhood as a group of indices that captures both similarities and idiosyncrasies of urban fabric improves our understanding of the urbanization process.

Spatial heterogeneity influences real estate prices on an even larger scale than the neighborhood level. This conclusion suggests, contrary to what is seen in the literature, that detailed local models of real estate price determination are required, rather than general models that theoretically assume a homogeneous city.

In terms of social heterogeneity, this study confirms that preferences and choices of consumers of real estate differ substantially over different levels of prices. Any analysis that does not take this into consideration provides only median results, which might give misleading signals if considered at different levels of prices. This is relevant for policy-makers, businesses, and developers when weighing the impact of changes in urban configuration on housing prices.

The second question was "do CA models on urban development increase understanding of urban dynamic processes when including social heterogeneity and negative economic feedback mechanisms?". This question is answered using the CA model in chapter 3. The proposed model simulates the dynamics of urban development by explicitly distinguishing between positive and negative forces of agglomeration. By doing so, it goes beyond the modeling framework proposed by White and Engelen (1993), in which all effects of local influence on urban development are modeled in an aggregated manner. Modeling the effects separately (agglomeration and disagglomeration) adds value because the spatial configuration of prices enables a second round of validation with an independent second set of data, as recommended by Batty and Torrens (2005).

Additionally, the empirical model in chapter 8 combined with the comparison and historical discussion in chapters 4 and 7, shows that a reasonable, current and developing urban area can be simulated by using long-term feedback processes in a rather large area, even if this means that the number of actors involved increases considerably. The results indicate that the dynamic influence of relative prices and interactions among actors is important in reaching a comparable result.

Specifically, sensitivity analysis shows that neighborhood influences, or the attraction forces observed among similar and different actors in a neighborhood, are an important issue in explaining urban dynamics for the case study. Spatial heterogeneity also plays an essential role, because actors entering the system weigh possible location alternatives in a non-linear fashion. This results in a multimodal, complex configuration that is in harmonious with the empirically observed distribution. Social heterogeneity is a crucial element as well. As the sensitivity analysis demonstrates, neglecting the socially diverse dimension of actors yields unsatisfying result for the case study. We therefore recommend that future land-use cellular automata modeling aimed at the intra-urban level should include social heterogeneity, unless the case study is highly homogeneous or top-down regulated (like the case in The Netherlands, for instance).

10.3 Limitations

The main limitation of this research is that it aims to understand economic factors as organizers of urban dynamics and structure, in addition to other factors. Heterodox approaches have advantages of linking realistic concepts together, but they also pose considerable strains on the process of integrating concepts – in our case, concepts from architecture and urban morphology, tools from geoprocessing, statistics and econometrics, and theoretical backgrounds from both economics and self-organizing systems. Despite this, we believe that the results offer a cohesive understanding of the determination of real estate prices in cities. Furthermore, the focus on the intra-urban scale provides insights that go beyond generic descriptions.

Specific to the CA model, the difficulty of simultaneously calibrating mechanisms to represent attractiveness and repulsion became apparent in this study. Even though the proposed modification to the original model of White and Engelen (1993) intended to accomplish just that, in practice, the choice of parameters necessary to dynamically generating these effects was complex.

10.4 Suggestions for further research

A more detailed forecast could be achieved given that a more detailed description of the city's idiosyncrasies at a larger scale was provided. This would imply more information on the implementation of historical exogenous inputs and accessibility features. The 1991 reference map can serve as starting year if it is improved with density information and actual satellite-imagery data on land-use. Such an emphasis could enable us to focus on specific areas in the city, rather than on general understanding of the processes. An immediate use of such a detailed model would be to study the effects on land price and spatial configuration as results of a newly constructed highway infrastructure in the North Vector (2008) and a Administrative Center for the state of Minas Gerais on its margins of an area of 80 hectares (also being constructed in 2008).

In relation to other metropolitan areas in Brazil, it would be useful for public policy-makers if a public or private-public consortium could be assembled to apply the same methodology to other metropolitan areas in Brazil (which already have spatial units elaborated by the same institution (FJP) with the same methodology).

Additionally, although this research (chapter 6) provides new indications of real estate price behaviors, given the intra-urban scale of analysis, urban descriptions seem to indicate that differentiation goes beyond the level of municipalities that is generally analyzed currently. As explained, the spatial unit is the smallest possible unit that follows an established methodology (Pinheiro, 2006). If neighborhoods were widely accepted, recognizable by the population, and statistically large enough to provide reliable variables, they would be a better choice. Another alternative to solving this problem would be to define weight matrices in econometric analysis that account simultaneously for neighborhoods and distance. Such a matrix should first indicate neighborhood membership, and second restrict relations to contiguity, defining only to real estate objects near each other as neighbors. A rule should also be conceived to include the "islands", or areas that remain without neighbors. A configuration like this might just be able to capture both aspects of membership and local city idiosyncrasies.

Finally, two further enhancements in the model proposed in chapter 3 might be subject to further research. An effort can be made to better define the mechanisms of attraction of each municipality for each actor. In practice, this would imply altering the exogenous demand factor to make it endogenous through a local fiscal framework. In other words, the attractiveness of each municipality could be included in the model based on these characteristics. This sheds light on the presence and role of positive feedback mechanisms and the attraction towards specific municipalities. Once again, this will be offset by subsequent increases in prices[96].

Additionally, competition among actors can be improved by introducing elements of bargain. Instead of competing only through higher transition potential, rules can be designed to promote specific, local interactions among potential proponents. Such a model extension would benefit from theoretical and empirical conceptualizations that evolved from the original trade mechanisms proposed by Epstein and Axtell (1996)[97]. In doing so, the analysis of the speculative feature typical of the housing market would be captured.

References

Abnt. Avaliação de bens. Parte 2: Imóveis urbanos Associação Brasileira de Normas Técnicas (Brazilian Association of Technical Norms) 14653-2: 34 p. 2004.

Abraham, J. M., Goetzmann, W. N., et al. Homogeneous Groupings of Metropolitan Housing Markets. Journal of Housing Economics, v.3, n.3, p.186-206. 1994.

Abramo, P. Mercado e ordem urbana: do caos à teoria residencial urbana. (Market and urban order: from chaos to urban residential theory, in Portuguese). Rio de Janeiro: Bertrand Brasil. 2001. 224 p.

Allen, P. M. Cities and regions as self-organizing systems: models of complexity. Amsterdam, Netherlands: Taylor & Francis, v.1. 1997. 296 p. (Environmental problems & social dynamics series)

Almeida, E. S. Curso de econometria espacial aplicada (Course on applied spatial econometrics, in Portuguese). Piracicaba: ESALQP-USP. 2004

Alonso, W. Location and Land Use: toward a general theory of land rent. Cambridge, MA: Harvard University Press. 1964. 205 p.

Anas, A., Arnott, R., et al. Urban spatial structure. Journal of Economic Literature, v.36, n.3, September, p.1426-1464. 1998.

Anderson, S. T. and West, S. E. Open space, residential property values, and spatial context. Regional Science and Urban Economics, v.36, p.773-789. 2006.

Anselin, L. Spatial econometrics: methods and models. Dordrecht: Kluwer Academic. 1988. 284 p.

Anselin, L. Spatial dependence and spatial structural instability in applied regression analysis. Journal of Regional Science, v.30, p.185-207. 1990.

Anselin, L. Some robust approaches to testing and estimation in spatial econometrics. Regional Science and Urban Economics, v.22, n.special issue. 1992a.

Anselin, L. SpaceStat tutorial: a workbook for using SpaceStat in the analysis of spatial data. University of Illinois. Urbana-Champaign, p.263. 1992b

Anselin, L. Spatial econometrics. University of Texas at Dallas. 1999

Anselin, L. Under the hood. Issues in the specification and interpretation of spatial regression models. Urbana, IL: Regional Economics Applications Laboratory, University of Illinois at Urbana-Champaign 2002.

Anselin, L. Spatial externalities, spatial multipliers and spatial econometrics. International Regional Science Review, v.26, n.2, April, p.153-166. 2003.

Anselin, L. Exploring spatial data with GeoDaTM: a workbook. Urbana-Champaign, IL: Spatial Analysis Laboratory, Department of Geography, University of Illinois. 2005. 244 p.

Anselin, L. and Bera, A. Spatial dependence in linear regression models with an introduction to spatial econometrics. In: A. Ullah e D. E. A. Giles (Ed.). Handbook of Applied Economic Statistics. New York: Marcel Dekker, 1998, p.237-289

Anselin, L., Syabri, I., et al. GeoDa: An Introduction to Spatial Data Analysis. Geographical Analysis, v.38, n.1, p.5-22. 2006.

Arbia, G. Spatial econometrics: statistical foundations and applications to regional convergence Berlin: Springer. 2006. 224 p. (Advances in Spatial Science)

Archer, W. R., Gatzlaff, D. H., et al. Measuring the Importance of Location in House Price Appreciation. Journal of Urban Economics, v.40, n.3, p.334-353. 1996.

Arnott, R. Economic theory and housing. In: E. S. Mills (Ed.). Handbook of Regional and Urban Economics. Amsterdam: Elsevier Science Publishers, v.2, 1987, p.959-988

Arthur, W. B. Urban systems and historical path dependence. In: J. H. Ausubel e R. Herman (Ed.). Cities and their vital systems. Washington, D.C.: National Academy Press, 1988, p.85-97

Arthur, W. B. Positive feedbacks in the economy. Scientific American, v.262, n.2, p.92-99. 1990.

Arthur, W. B. Inductive reasoning and bounded rationality. The American Economic Review, v.84, n.2, p.406-411. 1994.

Arthur, W. B. Complexity and the economy. Science, v.284, 2 April 1999, p.107-109. 1999.

Arthur, W. B. Out-of-equilibrium. In: K. Judd e L. Tesfatsion (Ed.). Handbook of computational economics: Elsevier/North-Holland, v.2, 2005, p.1551-1563. (Handbook in Economics)

Ball, G. H. and Hall, D. J. A Novel Method of Data Analysis and Pattern Classification. Menlo Park, CA: Stanford Research Institute 1965

Barredo, J. I. and Demicheli, L. Urban sustainability in developing countries' megacities: modelling and predicting future urban growth in Lagos. Cities, v.20, n.5, p.297-310. 2003.

Barredo, J. I., Kasanko, M., et al. Modelling dynamic spatial processes: simulation of urban future scenarios through cellular automata. Landscape and Urban Planning, v.64, p.145-160. 2003.

Barros, R. P. D., Foguel, M. N., et al., Eds. Desigualdade de Renda no Brasil: uma análise da queda recente (Income inequalities in Brazil: an analysis of recent decline, in Portuguese). Brasilia: Instituto de Pesquisa Economica Aplicada/IPEA, p.446ed. 2007.

Batten, D. Complex landscapes of spatial interaction. The Annals of Regional Science, v.35, p.81-111. 2001.

Batty, M. Urban evolution on the desktop: simulation with the use of extended cellular automata. Environment and Planning A, v.30, p.1943-1967. 1998.

Batty, M. Agents, cells and cities; new representational models for simulating multiscale urban dynamics. Environment and Planning A, v.37, n.8, p.1373-1394. 2005a.

Batty, M. Cities and complexity: understanding cities with cellular automata, agent-based models and fractals. Cambridge, MA: The MIT Press. 2005b. 648 p.

Batty, M. and Longley, P. Fractal cities: a geometry of form and function. Academic Press. 1994. 394 p.

Batty, M. and Torrens, P. M. Modelling and prediction in a complex world. Futures, v.37, n.7, p.745-766. 2005.

Behrens, K. Market size and industry location: Traded vs non-traded goods. Journal of Urban Economics, v.58, n.1, p.24-44. 2005.

Bell, M., Dean, C., et al. Forecasting the pattern of urban growth with PUP: a web-based model interfaced with GIS and 3D animation. Computers, Environment and Urban Systems, v.24, n.6, p.559-581. 2000.

Benenson, I. and Torrens, P. M. Geosimulation: automata-based modeling of urban phenomena. London: John Wiley. 2004. 287 p.

Benevolo, L. The history of the city. London: Scholar Press. 1980. 1011 p.

Berlekamp, E. R., Conway, J. H., et al. Winning ways for your mathematical plays. Wellesley, MA: A K Peters Ltd.. 2004

Biderman, C. Forças de Atração e Expulsão na Grande São Paulo. (Forces of attraction and repulsion in the Greater Area of Sao Paulo, in Portuguese). (Ph.D. dissertation). Escola de Administração de Empresas de São Paulo, Fundação Getúlio Vargas, São Paulo, 2001. 226 p.

Blakely, E. J. Fortress America: gated communities in the United States. Brookings Institution Press 1997. 209 p.

Bolaffi, G. Habitação e urbanismo: o problema e o falso problema. (Housing and urbanism: the problem and the fake problem). In: E. Maricato (Ed.). A Produção Capitalista da Casa (e da Cidade) no Brasil Industrial. São Paulo: Ed. Alfa-Omega, 1979, p.37-70

Bourassa, S. C., Cantoni, E., et al. Spatial dependence, housing submarkets and house price prediction. Journal of Real Estate Finance and Economics, v.35, p.143-160. 2007.

Bourassa, S. C., Hamelink, F., et al. Defining housing submarkets. Journal of Housing Economics, v.8, n.2, p.160-183. 1999.

Brasil. Censo demográfico 2000: agregado por setores censitários dos resultados do universo. Documentação do arquivo. M. D. Planejamento e I. B. D. G. E. Estatística: IBGE: 157 p. 2003.

Brasil. Síntese de Indicadores Sociais. O. E. G. Ministério Do Planejamento e I. B. D. G. E. Estatística: IBGE. 19: 317 p. 2006.

Brazil. Censo demográfico 1991: agregado por setores censitários dos resultados do universo. [Demographic census 1991: census tract aggregation of the results of the universe]. Brasilia: Instituto Brasileiro de Geografia e Estatística (IBGE). 1991

Brazil. Censo demográfico 2000: agregado por setores censitários dos resultados do universo. [Demographic census 2000: census tract aggregation of the results of the universe, in Portuguese]. Instituto Brasileiro de Geografia e Estatística (IBGE). 2003. 157 p.

Brown, D. G., Page, S., et al. Path dependence and the validation of agent-based spatial models of land use. International Journal of Geographical Information System, v.19, n.2, p.153-174. 2005.

Brueckner, J. The economics of urban sprawl: theory and evidence on the spatial sizes of cities. Review of Economics and Statistics, v.55, p.479-482. 1983.

Brueckner, J. The structure of urban equilibria: a unified treatment of the Muth-Mills model. In: E. Mills (Ed.). Handbook of Regional and Urban Economics: Elsevier Science Publishers B.V., v.II, 1987, p.821-845

Brueckner, J. Analysing Third-world urbanization: a model with empirical evidence. Economic Development and Cultural Change, v.38, n.3. 1990.

Brueckner, J. K., Thisse, J.-F., et al. Why is central Paris rich and downtown Detroit poor?: An amenity-based theory. European Economic Review, v.43, n.1, p.91-107. 1999.

Buchinsky, M. Recent advances in quantile regression models. Journal of Human Resources, v.33, n.1, p.88-126. 1997.

Burguess, E. W. The growth of the city. In: R. Park e E. Burguess (Ed.). The city. London: The University of Chicago Press, 1925, p.47-62. (The heritage of sociology)

Caldeira, T. P. D. R. City of walls: crime, segregation and citizenship in São Paulo. Berkeley: University of California Press. 2000

Can, A. and Megbolugbe, I. Spatial Dependence and House Price Index Construction. Journal of Real Estate Finance and Economics, v.14, p.203-222. 1997.

Canuto, F. "Pasta 00731 – volumes 01 e 02 end.: quarteirões fechados da praça 7 de setembro apenso ao processo n. – 01.059220.95.10": sobre a produção de uma nova sensibilidade urbana pela arquitetura. ("File 00731 – volumes 01 and 02 address: closed blocks of 7th of September square": about the production of a new urban sensibility for architecture, in Portuguese). (Master's thesis). Escola de Arquitetura/NPGAU, UFMG, Belo Horizonte, 2005. 179 p.

Capello, R. Urban rent and urban dynamics: the determinants of urban development in Italy. The Annals of Regional Science, v.36, p.593-611. 2002.

Capello, R. and Nijkamp, P. The theoretical and methodological toolbox of urban economics: from and towards where? In: R. Capello e P. Nijkamp (Ed.). Urban dynamics and growth: advances in urban economics. Amsterdam: Elsevier, 2004a, p.1-27. (Contributions to economic analysis)

Capello, R. and Nijkamp, P., Eds. Urban dynamics and growth: advances in urban economics. Amsterdam, Netherlands: Elsevier, p.880ed. 2004b.
Castells, M. A sociedade em rede. São Paulo: Paz e Terra, v.1. 1999. 617 p. (A era da informação: economia, sociedade e cultura)
Castro, Í. E. D., Correa, R. L., et al., Eds. Geografia: conceitos e temas. (Geography: concepts and themes, in Portuguese). Rio de Janeiro: Bertrand Brasil, p.352ed. 2001.
Chamberlain, E. The theory of monopolistic competition. Cambridge, Massachusetts: Harvard University Press. 1950
Cheshire, P. and Sheppard, S. On the Price of Land and the Value of Amenities. Economica, v.62, n.246, p.247-267. 1995.
Choay, F. L'urbanisme: utopies et réalités – une anthologie Paris. 1965
Clarke, K. C., Hoppen, S., et al. A self-modifying cellular automaton model of historical urbanization in the San Francisco Bay area. Environment and Planning B: Planning and Design v.24, n.2, p.247-261. 1997.
Cliff, A. and Ord, J. K. Spatial Autocorrelation. London: Pion. 1973
Colby, C. C. Centripetal and centrifugal forces in urban geography. Annals of the Association of American Geographers, v.23, n.1, p.1-20. 1933.
Corbusier, L. Towards a new architecture. London: Architectural Press. 1946. 269 p.
Costa, H. Habitação e produção do espaço em Belo Horizonte. (Housing and space production in Belo Horizonte, in Portuguese). In: R. L. D. M. Monte-Mór (Ed.). Belo Horizonte: espaços e tempos em construção. Belo Horizonte: CEDEPLAR & PBH, 1994, p.51-77. (Coleção BH 100 anos)
Costa, H., Costa, G., et al. Novas periferias metropolitanas: a expansão metropolitana em Belo Horizonte. (New metropolitan periferies: the metropolitan expansion in Belo Horizonte: dynamics and specificities of the South Axis). Belo Horizonte, Brazil: C/Arte. 2006. 467 p. (Coleção Estado da Arte)
Costa, J. Toponímia de Minas Gerais – com estudo histórico da divisão administrativa. (Toponymy of Minas Gerais – with a historical study of administrative divisions, in Portuguese). Belo Horizonte: Imprensa Oficial. 1970. 429 p.
Coulton, C. J., Korbin, J., et al. Mapping residents' perceptions of neighborhood boundaries: a methodological note. American Journal of Community Psychology, v.29, n.2. 2001.
Cullen, G. Townscape. London. 1961. 315 p.
Dale-Johnson, D. An alternative approach to housing market segmentation using hedonic price data. Journal of Urban Economics, v.11, n.3, p.311-332. 1982.
De Nijs, T. C. M., De Niet, R., et al. Constructing land-use maps of the Netherlands in 2030. Journal of Environmental Management, v.72, p.35-42. 2004.
Delden, H. V. and Engelen, G. Combining participatory approaches and modelling: lessons from two practical cases of policy support. Maastricht, The Netherlands, p.7. 2006. (mimeo)
Delden, H. V., Luja, P., et al. Integration of multi-scale dynamic spatial models of socio-economic and physical processes for river basin management. Environmental Modelling & Software, v.22, p.223-238. 2007.
Diamond, D. B. and Tolley, G. S. The economics of urban amenities. New York: Academic Press. 1982
Dickens, C. Oliver Twist. London: J. M. Dent, v.233. 1949. 410 p. (Everyman's library)
Dietzel, C. and Clarke, K. The effect of disaggregating land use categories in cellular automata during model calibration and forecasting. Computers, Environment and Urban Systems, v.30, n.1, p.78-101. 2006.
Dipasquale, D. and Wheaton, W. C. Urban economics and real estate markets. Englewood Cliffs, NJ: Prentice Hall. 1996. 378 p.
Dowall, D. The Land Market Assessment: a new tool for urban management. Washington, D.C.: The World Bank. 1995. 92 p.

Durlauf, S. N. Neighbourhood effects. In: J. V. Henderson e J.-F. Thisse (Ed.). Handbook of regional and urban economics. Amsterdam, Netherlands: Elsevier, v.4, 2004, p.2173-2242. (Hanbooks in Economics)

Engelen, G., Hagen-Zanker, A., et al. Report 2550016006/2005. Environmental Planning Agency. Bilthoven, Netherlands. 2005

Engelen, G., White, R., et al. Using cellular automata for integrated modelling of socio-environmental systems. Environmental Monitoring and Assessment, v.30, p.203-214. 1995.

Engelen, G., White, R., et al. The Moland Model for urban and regional growth. Maastrich, Holanda: Research Institute for Knowledge Systems. 2004. 238 p.

Epstein, J. M. and Axtell, R. Growing artificial societies: social science from the bottom up. Cambridge, MA: Brookings/MIT Press. 1996

Esri. ArcMAP Desktop Help (software, version 9.2). ESRI Inc. 2006

Fávero, L. P. L. O mercado imobiliário residencial de São Paulo: uma aplicação de modelos de comercialização hedônica e correlação canônica. (Real estate market of Sao Pualo: an application of models of hedonic comercialization and canonical correlation). São Paulo: Universidade de São Paulo, v.Doutoramento. 2005. 350 p. (Departamento de Administração da FEA)

Fávero, L. P. L., Belfiore, P. P., et al. Modelos de precificação hedônica de imóveis residenciais na Região Metropolitana de São Paulo: uma abordagem sob as perspectivas da demanda e da oferta. (Models of hedonic precification of residential real estate in the Greater Area of Sao Paulo: an approach under the perspectives of supply and demand). Estudos Econômicos, v.38, n.1, p.73-96. 2008.

Fernandez, L. E., Brown, D. G., et al. Characterizing location preferences in an exurban population: implications for agent-based modeling. Environment and Planning B, v.32, n.6, p.799-820. 2005.

Fujita, M. Urban economic theory: land use and city size. Cambridge: Cambridge University Press. 1989. 366 p.

Fujita, M., Krugman, P., et al. The spatial economy: cities, regions and international trade. Cambridge, Mass.: MIT Press. 1999. 367 p.

Fujita, M. and Ogawa, H. Multiple equilibria and structural transition of non-monocentric urban configurations. Regional Science and Urban Economics, v.12, p.161-196. 1982.

Furtado, B. A. Favelas inseridas no meio urbano em Belo Horizonte. (Slums within urban tissue in Belo Horizonte). Belo Horizonte: UFMG. 1995. 52 p. (Escola de Arquitetura)

Furtado, B. A. Minas Gerais – evolução dos limites municipais: uma análise exploratória. (Minas Gerais – evolution of municipalities' boundaries: an exploratory analysis, in Portuguese). (Master's thesis). Programa de Tratamento da Informação Espacial, Pontíficia Universidade Católica de Minas Gerais, Belo Horizonte, 2003. 117 p.

Furtado, B. A. Crítica à lógica da urbanização. [Criticism to the logics of urbanization]. Geosul, v.21, n.41, p.127-139. 2006.

Furtado, C. Formação econômica do Brasil. (Economic formation of Brazil, in Portuguese). Rio de Janeiro: Companhia das Letras. 1956. 352 p.

Furtado, C. Pequena introdução ao desenvolvimento. (Brief introduction to development, in Portuguese). São Paulo: Companhia Editora Nacional. 1980

Galster, G. William Grigsby and the analysis of housing sub-markets and filtering. Urban Studies, v.33, n.10, p.1797-1805. 1996.

Galster, G. On the Nature of Neighbourhood. Urban Studies, v.38, n.12, p.2111-2124. 2001.

Garreau, J. Edge City: Life on the New Frontier Random House Inc. 1992

Glaeser, E. L. Are cities dying? Journal of Economic Perspectives, v.12, n.2, p.139-160. 1998.

Glaeser, E. L. The economics approach to cities. Working Paper NBER. Cambridge, MA: National Bureau of Economic Research 13696 2007.

Glaeser, E. L., Gyourko, J., et al. Why is Manhattan so expensive? Regulation and the rise in house prices. NBER Working Paper Series. Cambridge, MA: National Bureau of Economic Research 2003.
Glaeser, E. L., Gyourko, J., et al. Urban growth and housing supply. Journal of Economic Geography, v.6, p.71-89. 2005.
Golab, C. The geography of the neighbourhood. In: R. Bayor (Ed.). Neighbourhoods in urban America. Washington: Kennikat Press 1982, p.70-85
Goodman, A. C. and Thibodeau, T. G. Housing market segmentation. Journal of Housing Economics, v.7, p.121-143. 1998.
Gouvêa, R. A questão metropolitana no Brasil. (The metropolitan issue in Brazil). Sao Paulo, Brazil: FGV. 2005. 324 p.
Grigsby, W., Baratz, G., et al. The dynamics of neighbourhood change and decline. London: Pergamon. 1987
Griliches, Z. Hedonic price indexes for automobiles: an econometric analysis of quality change. The price statistics of the Federal Government, v.73. 1961.
Guest, A. M. and Lee, B. A. Consensus on locality names in the metropolis. Sociology and Social Research v.67, p.374-391. 1983.
Guo, J. Y. and Bhat, C. R. Operationalizing the concept of neighborhood: Application to residential location choice analysis. Journal of Transport Geography, v.15, n.1, p.31-45. 2007.
Hagen-Zanker, A. Map comparison methods that simultaneously address overlap and structure. Journal of Geographical Systems, v.8, p.165-185. 2006.
Hagoort, M. J. The neighbourhood rules: land-use interactions, urban dynamics and cellular automata modelling. Utrecht: Faculteit Geowetenschappen Universiteit Utrecht, v.334. 2006. 298 p. (Nederlands Geographical Studies)
Haig, R. M. Toward an understanding of the metropolis. Quarterly Journal of Economics, v.40, May, p.421-423. 1926.
Harvey, D. Social justice and city. Baltimore: Johns Hopkins University. 1973. 336 p.
Harvey, D. A condição pós-moderna: uma pesquisa sobre as origens da mudança cultural. São Paulo: Loyola. 1992. 349 p.
Henderson, J. V. The size and types of cities. American Economic Review, v.64, n.4, September, p.640-656. 1974.
Hermann, B. M. and Haddad, E. A. Mercado Imobiliário e Amenidades Urbanas: A View Through the Window. Estudos Econômicos v.35, n.2, Abril-junho, p.237-269. 2005.
Hobsbawm, E. The age of the capital. London: Weidenfeld and Nicolson. 1975. 354 p.
Hobsbawm, E. The age of empire, 1875-1914. New York: Pantheon Books. 1987. 404 p. (History of civilisation)
Holland, J. Complex adaptive systems. Daedalus, v.121, p.17-30. 1992.
Holland, J. and Miller, J. H. Artificial Adaptive Agents in Economic Theory. The American Economic Review, v.81, n.2, May, p.65-71. 1991.
Ibbotson, R. G., Siegel, L. B., et al. World wealth: Market values and returns. Journal of Portfolio Management, v.12, n.1, Fall, p.4-23. 1985.
Isard, W. Location and space-economy: a general theory relating to industrial location, market areas, land use, trade, and urban structure. New York: John Wiley & Sons. 1956. 350 p.
Isard, W. Methods of interregional and regional analysis. Aldershot: Ashgate. 1998. 490 p. (Regional Science Studies Series)
Jacobs, J. The economy of cities. New York: Vintage Books. 1970. 268 p.
Johnson, R. and Wichern, D. Applied multivariate statistical analysis. Prentice Hall. 1998. 799 p.
Kauko, T. J. Modelling the locational determinants of house prices: neural network and value tree approaches. Faculty of Geosciences, Utrecht University, Utrecht, 2002. 251 p.

Kelejian, H. H. and Robinson, D. P. Spatial correlation: a suggested alternative to the autoregressive model. In: L. Anselin e R. Florax (Ed.). New Directions in Spatial Econometrics. Berlin: Springer-Verlag, 1995, p.75-95

Kelejian, H. H. and Robinson, D. P. A suggested test for spatial autocorrelation and/or heteroskedasticity and corresponding Monte Carlo results. Regional Science and Urban Economics, v.28, p.389-417. 1998.

Kelejian, H. H. and Robinson, D. P. The influence of spatially correlated heteroskedasticity on tests for spatial correlation. In: L. Anselin e R. Florax (Ed.). Advances in Spatial Econometrics. Heidelberg: Springer-Verlag, 1999

Koenker, R. and Bassett, G. Regression Quantiles. Econometrica, v.46, n.1, January, p.33-50. 1978.

Krugman, P. The Self-organizing Economy. Cambridge, MA: Blackwell Publishers Inc. 1996. 122 p.

Lau, K. H. and Kam, B. H. A cellular automata model for urban land-use simulation. Environment and Planning B, v.32, p.247-263. 2005.

Lee, B. A. and Campbell, K. E. Common Ground? Urban Neighborhoods as Survey Respondents See them. Social Science Quarterly (University of Texas Press), v.78, n.4, p.922-936. 1997.

Lee, B. A., Oropesa, R. S., et al. Neighborhood Context and Residential Mobility. Demography, v.31, n.2, p.249-270. 1994.

Lefebvre, H. Le droit à la ville. Paris: Anthropos. 1965

Lefebvre, H. De lo rural a lo urbano. Buenos Aires: Lotus Maro. 1976

Lefebvre, H. The production of space. Oxford: Blackwell. 1991

Lemos, C. B., Ferreira, G. H. D. M., et al. A questão intra-urbana. CEDEPLAR/UFMG. Belo Horizonte. 2004

Leroy, S. F. and Sonstelie, J. Paradise lost and regained: Transportation innovation, income, and residential location. Journal of Urban Economics, v.13, p.67-89. 1983.

Lesage, J. The theory and practice of spatial econometrics. Department of Economics. University of Toledo. 1999. 309 p.

Leven, C., Little, J., et al. Neighborhood change. New York: Praeger Publishers. 1977

Lillesand, T. M., Kiefer, R. W., et al. Remote sensing and image interpretation. New York: Wiley. 2004. 763 p.

Liporoni, A. S., Neto, D. N., et al. Instrumentos para gestão tributária de cidades. (Tools for fiscal management of cities, in Portuguese). São Paulo: Liv. e Ed. Universitária de Direito. 2003. 161 p.

Liu, Y. and Phinn, S. R. Modelling urban development with cellular automata incorporating fuzzy-set approaches. Computers, Environment and Urban Systems, v.27, n.6, p.637-658. 2003.

Longley, P. and Batty, M. Advanced Spatial Analysis: The CASA Book of GIS. London: ESRI Press 2003. 275 p.

Lucas, R. and Rossi-Hansberg, E. On the internal structure of cities. Econometrica, v.70, n.4, July, p.1445-1476. 2002.

Lynch, K. and Camargo, J. L. The image of the city. São Paulo: Martins Fontes. 1960. 227 p.

Macedo, P. B. R. Hedonic price model with spatial effects: an application to the housing market of Belo Horizonte, Brazil. Textos para discussão. Belo Horizonte: CEDEPLAR/FACE/UFMG: 22 p. 1996.

Maricato, E. T. M. A produção capitalista da casa (e da cidade) no Brasil industrial. (The capitalist production of housing (and the city) in industrial Brazil). São Paulo: Alfa-Ômega. 1979. 166 p.

Marshall, A. Principles of economics. New York: Cosimo Classics/Prometheus Books. 1890. 348 p.

Mcgarigal, K., Cushman, S. A., et al. FRAGSTATS: spatial pattern analysis program for categorical maps. Amherst: University of Massachusetts 2002.

Megbolugbe, I. F., Hoek-Smit, M. C., et al. Understanding neighbourhood dynamics: a review of the contributions of William G. Grigsby. Urban Studies, v.33, n.10, p.1779-1795. 1996.

Mendonça, J. G. Segregação e mobilidade residencial na RMBH. (Segregation and residential mobility in RMBH, in Portuguese). IPPUR, UFRJ, Rio de Janeiro, 2002.

Mendonça, J. G. and Godinho, M. H. D. L. População espaço e gestão na metrópole: novas configurações, velhas desigualdades (People, space and management in the metropolis: new configurations, old inequalities). Belo Horizonte: PUC Minas. 2003. 336 p.
Mills, E. S. An aggregative model of resource allocation in a metropolitan area. American Economic Review, v.57, p.197-210. 1967.
Mills, E. S. and Nijkamp, P. Advances in Urban Economics. In: E. S. Mills (Ed.). Handbook of Regional and Urban Economics. Amsterdam: North-Holland, v.2, 1987
Mingoti, S. A. Análise de dados através de métodos de estatística multivariada. (Data analysis using multivariate statistics methods, in Portuguese). Belo Horizonte: Editora UFMG. 2005. 297 p.
Miyao, T. Dynamic urban models. In: P. Cheshire e E. S. Mills (Ed.). Handbook of regional and urban economics. Amsterdam North-Holland, v.2, 1987
Monte-Mór, R. L. D. M. Belo Horizonte: espaços e tempos em construção. (Belo Horizonte: spaces and times in construction, in Portuguese). Belo Horizonte: CEDEPLAR/PBH. 1994. 93 p. (Coleção BH 100 anos.)
Monte-Mór, R. L. D. M. O que é o urbano no mundo contemporâneo. (What is urban in the contemporary world?, in Portuguese). Textos para discussão, v.281, 01 april 2007, p.14. 2006.
Muhammad, S. Future urbanization patterns: in the Netherlands, under the influence of information and communication technologies (Doctorate). Faculty of geosciences, Utrecht University, Utrecht, 2007. 187 p.
Muth, R. F. Cities and Housing. Chicago: University of Chicago Press. 1969
Naik, D. N. and Khattree, R. Revisiting olympic track records: some practical considerations in the principal component analysis. The American Statistician, v.50, n.2, p.140-144. 1996.
Paelinck, J. and Klaassen, L. Spatial Econometrics. Farnborough: Saxon House. 1979
Page, S. E. On the emergence of cities. Journal of Urban Economics, v.45, p.184-208. 1999.
Palm, R. Spatial Segmentation of the Urban Housing Market. Economic Geography, v.54, n.3, p.210-221. 1978.
Park, R. E., Burgess, E. W., et al., Eds. The city. Chicago: University of Chicago Pressed. 1925.
Parker, D. C., Manson, S. M., et al. Multi-Agent Systems for the Simulation of Land-Use and Land-Cover Change: A Review. Annals of the Association of American Geographers, v.93, n.2, p.314-337. 2003.
Pines, D. and Thisse, J. F. Urban systems: market and efficiency. Journal of Public Economic Theory, v.3, n.1, p.7-14. 2001.
Pinheiro, F. J. Desenvolvimento Humano na Região Metropolitana de Belo Horizonte: Atlas Metropolitano. (Human development in the Greater Area of Belo Horizonte: metropolitan atlas, in Portuguese). Belo Horizonte: Centro de Estudos Econômicos e Sociais, FJP. 2006
Plambel. O mercado da terra na Região Metropolitana de Belo Horizonte. (The land market in the Greater Area of Belo Horizonte). Belo Horizonte: Plambel. 1987. 207 p.
Pontius, R., Boersma, W., et al. Comparing the input, output, and validation maps for several models of land change. The Annals of Regional Science, v.42, n.1, p.11-37. 2008.
Portugali, J. Self-organization and the city. Berlin: Springer-Verlag. 2000. 352 p. (Springer series in synergetics)
Quigley, J. What have we learned about urban housing markets? In: P. Mieszkowski e M. Straszheim (Ed.). Current issues in Urban Economics. Baltimore: John's Hopkins, 1979, p.391-429
Rauch, J. Seeing around corners. The Atlantic Monthly, v.289, n.4, April, p.35-48. 2002.
Richards, J. A. Remote Sensing Digital Image Analysis: An Introduction. Berlin: Springer-Verlag. 1986
Riks. METRONAMICA: a dynamic spatial land use model. Maastrich, The Netherlands: RIKS BV. 2005. 17 p.
Riks. Map Comparison Kit 3: user manual. Maastricht, NL: RIKS BV. 2006. 73 p.
Riks. METRONAMICA data requirements. Maastricht, The Netherlands: RIKS BV. 2007a. 4 p.
Riks. METRONAMICA model description. Maastricht, The Netherlands: RIKS BV. 2007b. 70 p.
Rose-Ackerman, S. Rascism and urban structure. Journal of Urban Economics, v.2, p.85-103. 1975.

Rosen, S. Hedonic Prices and Implicit Markets: product differentiation in pure competition. Journal of Political Economy. 1974.

Rossi, A. Architecture of the city (L'architettura della città). Padova/MIT Press. 1966. 201 p.

Santiago, R. N. Obsolescência Programada no Mercado Imobiliário – o espaço como forma de entesouramento. (Planned obsolescence in the real estate market – space as a way of hoarding, in Portuguese). Escola de Arquitetura, Universidade Federal de Minas Gerais, Belo Horizonte, 2006. 136 p.

Santos, C. Políticas Federais de Habitação no Brasil: 1964/1998. (Federal housing policies in Brazil). Textos para discussão. 1999.

Santos, M. A urbanização brasileira. (The Brazilian urbanization). São Paulo: Editora Hucitec. 1994. 157 p.

Santos, M. Estradas Reais: introdução ao estudo dos caminhos do ouro e do diamante no Brasil. (Royal roads: introduction to the study of gold and diamonds' ways in Brazil, in Portuguese). Belo Horizonte: Editora Estrada Real. 2001. 179 p.

Schnare, A. B. and Struyk, R. J. Segmentation in urban housing markets. Journal of Urban Economics, v.3, n.2, p.146-166. 1976.

Serra, G. O espaço natural e a forma urbana. (The natural space and the urban form, in Portuguese). São Paulo: Nobel. 1987. 211 p. (Coleção espaços)

Sheppard, S. Hedonic analysis of housing markets. In: P. Cheshire e E. S. Mills (Ed.). Handbook of regional and urban economics. Amsterdam: North-Holland, v.3, 1999, p.1595-1636

Silva, E. A. and Clarke, K. C. Calibration of the SLEUTH urban growth model for Lisbon and Porto, Portugal. Computers, Environment and Urban Systems, v.26, p.525-552. 2002.

Simões, R. F. Localização industrial e relações intersetoriais: uma análise de fuzzy cluster para Minas Gerais. (Industrial location and inter-sector relations: a fuzzy cluster analysis for Minas Gerais, in Portuguese). (Ph.D. dissertation). Instituto de Economia, Unicamp, Campinas, 2003. 198 p.

Simões, R. F. Métodos de análise regional e urbana: diagnóstico aplicado ao planejamento. (Methods of urban and regional analysis: diagnosis applied to planning, in Portuguese). Belo Horizonte: CEDEPLAR/UFMG. 2005 (Working Paper 259)

Singer, P. Desenvolvimento econômico e evolução urbana. (Economic development and urban evolution, in Portuguese). São Paulo: Companhia Editora Nacional, v.22. 1977. 377 p. (Ciências Sociais)

Sitte, C. City planning according to artistic principles. London, v.2. 1889. 205 p. (Columbia University Studies in Art History and Archaeology)

Smirnov, O. and Anselin, L. Fast maximum likelihood estimation of very large spatial autoregressive models: A characteristic polynomial approach. Computational Statistics and Data Analysis, v.35, p.301-319. 2001.

Soares, S. S. D. O ritmo de queda na desigualdade no Brasil é adequado? Evidências do contexto histórico e internacional. (The rhythm of reduction of Brazilian inequalities is adequate? Evidences from the historic and international contexts). Textos para discussão do IPEA, v.1339, p.21. 2008.

Soja, E. W. Postmodern geographies: the reassertion of space in critical social theory. London: Verso. 1990. 17-30 p.

Soja, E. W. Third space: expanding the geographical imagination. Oxford: Blackwell Publishers. 1996. 354 p.

Soja, E. W. Postmetropolis: critical studies of cities and regions. Malden, MA: Blackwell Publishers. 2000. 464 p.

Somarriba, M. D. M. G., Valadares, M. G., et al. Lutas urbanas em Belo Horizonte. (Urban fight in Belo Horizonte). Belo Horizonte: Fundação João Pinheiro. 1984. 130 p.

Steinnes, D. N. Do people follow jobs' or do jobs follow people? A causality issue in urban economics. Urban Studies, v.19, p.187-192. 1982.

Stevenson, S. New empirical evidence in heteroscedasticity in hedonic housing models. Journal of Housing Economics, v.13, p.136-156. 2004.

Su, L. and Yang, Z. Instrumental Variable Quantile Estimation of Spatial Autoregressive Models. I World Conference on Spatial Econometrics. Cambridge, UK: Pace Projects, 2007. 31 p.

Suttles, G. D. Social construction of communities. University of Chicago Press. 1972. 278 p. (Studies of Urban Society)

Teixeira, L. P. Metodologia das pesquisas do mercado imobiliário. (Methodology of research on real estate, in Portuguese). Belo Horizonte: Câmara Brasileira da Indústria da Construção (Brazilian Chamber of Construction Industry). 1999. 36 p.

Thünem, J. H. V. Der isolierte staat. Hamburg (New York): Pergamon Press, v.I. 1826

Tiebout, C. M. A pure theory of local expenditures. Journal of Political Economy, v.64, p.416-424. 1956.

Tobler, W. R. A computer movie simulating urban growth in the Detroit region. Economic Geography, v.46, p.234-240. 1970.

Torrens, P. M. New advances in urban simulation: cellular automata and multi-agent systems as planning support tools. Working Paper Series. London, UK: Center for Advanced Spatial Analysis 2001.

Tuan, Y. Topophilia: a study of environmental perception, attitudes, and values Englewood Cliffs, N.J.: Prentice-Hall 1980. 260 p.

Tuan, Y. Space and place: the perspective of experience. São Paulo: DIFEL. 1983. 250 p.

Tversky, B. Structures of mental spaces: how people think about space. Environment and Behavior, v.35, n.1, p.66-80. 2003.

Verburg, P. H. and Veldkamp, T. A. Projecting land use transitions at forest fringes in the Philippines at two spatial scales. Landscape Ecology, v.19, n.77-98. 2004.

Villaça, F. Espaço intra-urbano no Brasil. (Intra-urban space in Brazil). São Paulo, Brazil: Studio Nobel. 1998. 373 p.

Visser, H. and De Nijs, T. The Map Comparison Kit. Environmental Modeling & Software, v.21, p.346-358. 2006.

Waddell, P. UrbanSim: modeling urban development for land use, transportation, and environmental planning. Journal of the American Planning Association, v.68, n.3, p.297-314. 2002.

Waddell, P. and Ulfarsson, G. F. Dynamic Simulation of real estate development and land prices within an integrated land use and transportation model system. Transportation Research Board 82nd Annual Meeting. Washington, D.C., 2003. 21 p.

Waddell, P., Ulfarsson, G. F., et al. Incorporating land use in metropolitan transportation planning. Transportation Research, v.41, p.382-410. 2007.

Wendt, P. F. Economic growth and urban land values. The Appraisal Journal, v.26, July, p.427-443. 1958.

Wheaton, W. C. On the optimal distribution of income among cities. Journal of Urban Economics, v.3, p.31-44. 1976.

Wheaton, W. C. Income and urban residence: An analysis of consumer demand for location. American Economic Review, v.67, p.620-631. 1977.

Wheaton, W. C. Commuting, congestion, and employment dispersal in cities with mixed land use. Journal of Urban Economics, v.55, n.3, p.417-438. 2004.

Wheaton, W. C. and Nechayev, G. Does Location Matter? Journal of Portfolio Management, v.32, Sep2005 Special Real Estate Issue, p.100-108. 2005.

White, R. and Engelen, G. Cellular automata and fractal urban form: a cellular modelling approach to the evolution of urban land-use patterns. Environment and Planning A, v.25, p.1175-1199. 1993.

White, R., Engelen, G., et al. The use of constrained cellular automata for high-resolution modelling of urban land-use dynamics. Environment and Planning B, v.24, p.323-343. 1997.

Whitehead, C. M. E. Urban housing markets: theory and policy. In: P. Cheshire e E. S. Mills (Ed.). Handbook of Regional and Urban Economics. Amsterdam: North-Holland, v.3, 1999, p.1559-1594

Whittle, P. On stationary processes in the plane. Biometrika, v.41, p.434-449. 1954.
Wickramasuriya, R. C. Application and assessment of usability of the land use model Metronamica. Centre for Geo-Information, Wageningen University, Wageningen, The Netherlands, 2007. 79 p.
Wingo, L. Transportation and urban land. Washington, D.C.: Resources for the future. 1961. 132 p.
Wirth, L. Urbanism as a way of life. The American Journal of Sociology, v.44, n.1, p.1-24. 1938.
Wooldridge, J. M. Econometric analysis of cross section and panel data. Cambridge, Mass.: MIT Press. 2002. 752 p.
Wooldridge, J. M. Introductory econometrics: a modern approach. Mason: Thomson South-Western. 2003. 819 p.
Yang, C. and Fujita, M. Urban spatial structure with open space. Environment and Planning A, v.15, p.67-84. 1983.
Zietz, J., Zietz, E. N., et al. Determinants of house prices: a quantile regression approach. Journal of Real Estate Finance and Economics, v.37, n.3, p.(forthcoming). 2008.

11 Appendix

11.1 Initial occupation of Belo Horizonte

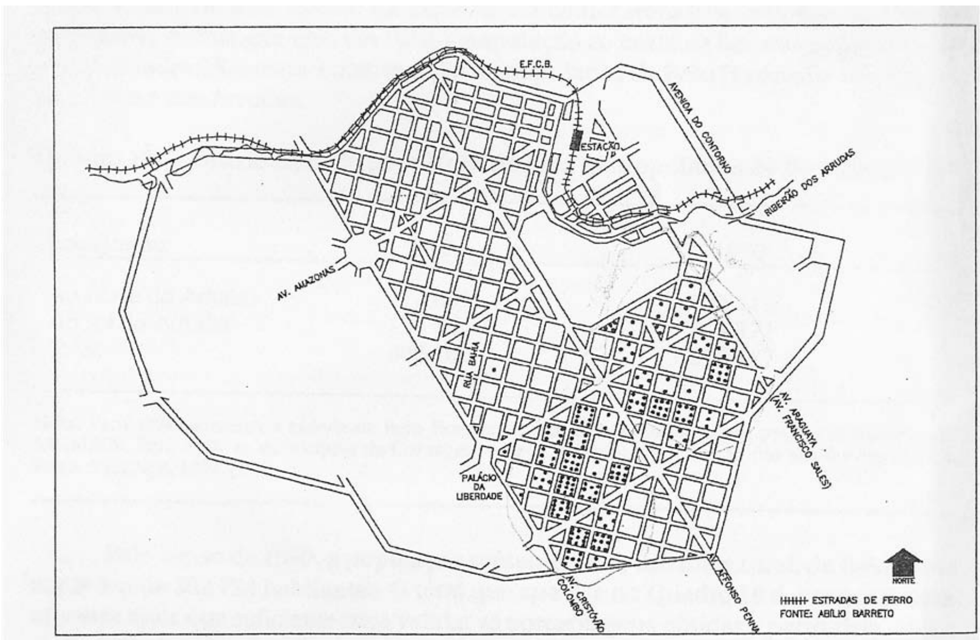


Figure 11.1 Initial occupation. Source: reproduction from Villaça (1998, p. 122).

11.2 RMBH municipalities at creation in 1974

See Figure 11-2

11.3 Principal components

Principal component analysis (PCA) is a statistics technique that has been largely used in data compression (Johnson and Wichern, 1998). PCA has been used in real estate context by Macedo (1996), and Hermann and Haddad (2005) in Brazil, and by Can and Megbolugebe (1997) in the United States, among many other to analyze real estate markets (Dale-Johnson, 1982; Abraham, Goetzmann et al., 1994; Archer, Gatzlaff et al., 1996; Bourassa, Steven C., Hamelink et al., 1999).

According to Mingoti (2005, p. 59):

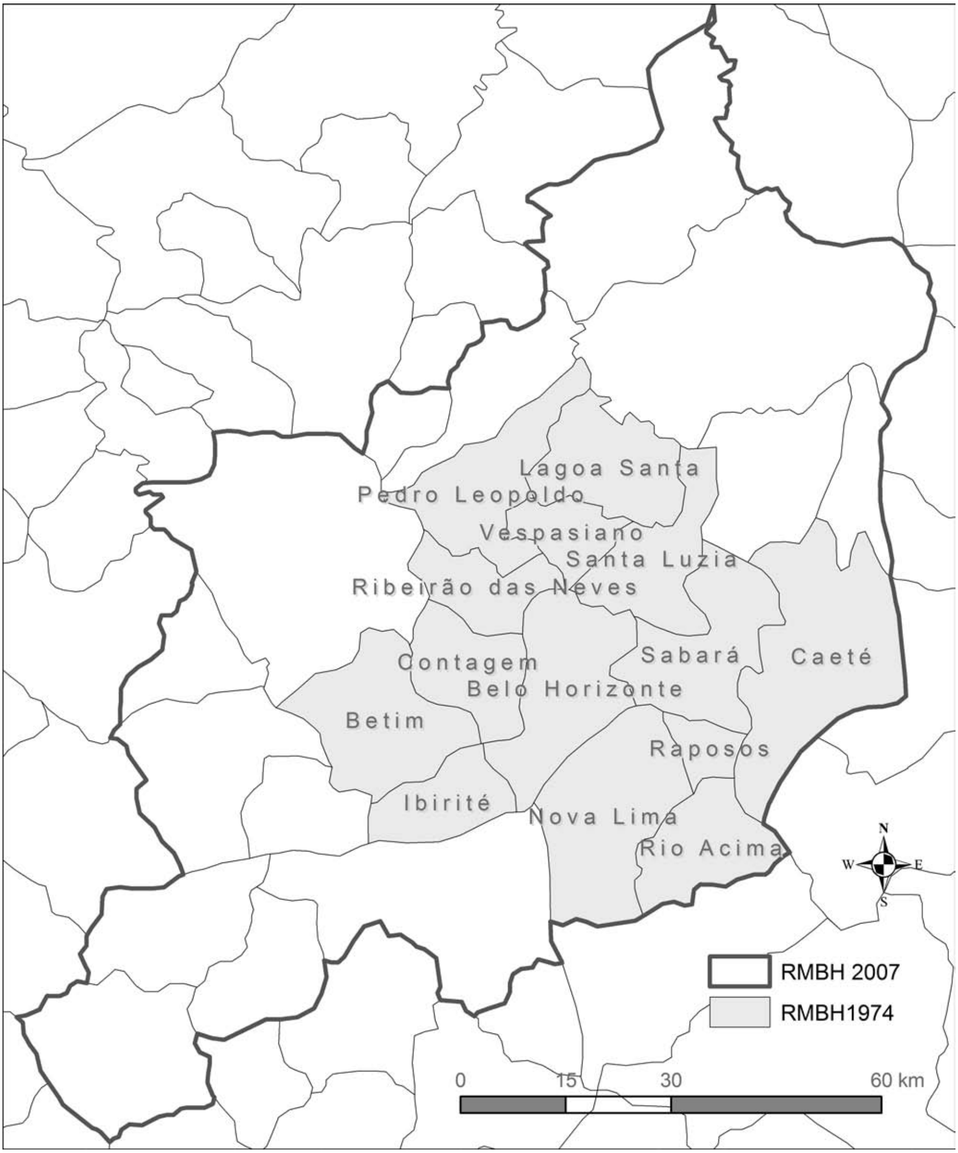


Figure 11.2 Municipalities of RMBH in 1974 Source: Furtado, 2003.

its main objective is to explain the structure of the variance and co-variance of a stochastic vector (…) through the construction of linear combinations of the original variables. These linear combinations are called principal components and are not correlated among each other.

Be $X=(X_1X_2,\ldots,X_p)$ a random vector; Σ_{pxp} its matrix of co-variance with $\lambda_1 \geq \lambda_2 \geq K \geq \lambda_p$, the eigenvalue of the matrix and $e_1,e_2,\ldots,e_p$, the eigenvectors normalized. The j^th^ principal component of the matrix Σ_{pxp} can be defined as (Mingoti, 2005, p. 60):

Equation 44

$$Y_j=e'_jX=e_{j1}X_1+e_{j2}X_2+\ldots+e_{jp}X_p$$

Table 11.1 BKW – test for multicollinearity

Belsley, Kuh, Welsch Variance-decomposition								
K(x)	**constant**	**lnarea**	**standard**	**age**	**d_ap**	**domic_n**	**pca_activ**	**pca_ind_n**
1	0.00	0.00	0.00	0.20	0.00	0.00	0.00	0.00
8	0.00	0.02	0.01	0.42	0.00	0.00	0.00	0.00
41	0.00	0.24	0.36	0.09	0.03	0.00	0.00	0.00
52	0.00	0.00	0.01	0.01	0.02	0.00	0.00	0.00
56	0.00	0.07	0.49	0.06	0.07	0.00	0.00	0.00
62	0.00	0.01	0.04	0.07	0.74	0.00	0.00	0.00
67	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.00
73	0.00	0.00	0.00	0.01	0.01	0.00	0.00	0.00
128	0.01	0.11	0.07	0.08	0.11	0.52	0.04	0.01
166	0.07	0.20	0.01	0.00	0.00	0.00	0.36	0.00
199	0.07	0.08	0.00	0.03	0.00	0.15	0.06	0.63
231	0.11	0.07	0.01	0.01	0.00	0.31	0.51	0.15
292	0.73	0.19	0.00	0.01	0.00	0.02	0.03	0.20
K(x)	**pca2_ino_n**	**shop_1500**	**slum_200**	**aven_200**	**f_melpub2**			
1	0.00	0.00	0.00	0.00	0.00			
8	0.00	0.00	0.00	0.00	0.00			
41	0.00	0.01	0.00	0.00	0.08			
52	0.00	0.01	0.01	0.85	0.01			
56	0.00	0.03	0.01	0.00	0.46			
62	0.00	0.03	0.02	0.04	0.14			
67	0.00	0.52	0.06	0.05	0.24			
73	0.00	0.12	0.86	0.01	0.00			
128	0.02	0.00	0.00	0.02	0.04			
166	0.17	0.13	0.00	0.00	0.03			
199	0.01	0.03	0.01	0.00	0.00			
231	0.25	0.11	0.00	0.01	0.00			
292	0.55	0.00	0.02	0.01	0.00			

The way the components are elaborated implies that the first component is always the most representative in terms of total variance. The numerical values of the principal components are called scores, and can be used in regression analysis.

The procedures, which are made automatically by the software ArcGIS or numerous other statistical commercial packages, are: the data is standardized and the matrix of correlation and the eigenvectors and eigenvalues are calculated. Those calculations enable the construction of the scores for each observation according to equation 37.

One of the advantages of PCA technique is that is not necessary to assume the normality of the distribution. A disadvantage is that some studies have shown that changes on the scale of the variables might alter the results (Naik and Khattree, 1996). The standardization of the data helps minimize this problem.

11.4 Other data

See table 11.1

11.5 Metrics on quantifying maps

Most of the metrics are based on 'patch size'. Patches are "groups of contiguous cells that are taken in by the same category" (Hagen-Zanker, 2006, p. 171). Of each patch, its size, perimeter, and edge length can be measured.

Other common (and popular) descriptive numbers are fractal dimension and the shape index (Batty, M. and Longley, 1994; Benenson and Torrens, 2004; Batty, M., 2005b).

Fractal dimensions provide a single number for a map which indicates how much of space is completely filled with a certain state or category. Two, for instance, represents a square filled with the same state, class or category. Among some other specifications (Batty, M., 2005b), it might be given as (Mcgarigal, Cushman *et al.*, 2002):

Equation 45

$$F = \frac{2\ln(0.25\,p_{ij})}{\ln a_{ij}}$$, where

F, and $F \in (1,2)$, is the fractal dimension; p_{ij} is the perimeter of each patch given by its location on an $i \times j$ plane, and a_{ij} is the area of a specific patch. The intuition behind it is that the fractal dimension represents the complexity of the shape.

The shape index, in turn, is calculated as the perimeter divided by the square root of the patch size. It is easy to see that "larger values indicate a more convoluted shape" (Riks, 2006, p. 33).

A number of software packages to calculate these measures is available such as The Map Comparison Kit (Visser and De Nijs, 2006) and Fragstats (Mcgarigal, Cushman *et al.*, 2002).

In the format applied here, what is seeked is the general structure pattern, therefore the indicators are calculated considering an average moving window of radius of four cells, as suggested by Hagen-Zanker (2006).

11.6 Methodological note on the data availability for 1991

Information for 1991 is only available in table format. This means that the polygons that describe each census tract are unavailable. The government foundation (IBGE) provided two correspondence tables (2000-1996 and 1996-1991), which enables the linking of census tract codes of 2000, for which polygons are available, to the census tract codes of 1996's population count. IBGE also provided census tract codes of 1996, which are linked to census tract codes of 1991.

The polygon map of 1991's 2,788 census tracts is the result of the aggregation of 4,121 census tract of 2000.

Some inconsistencies were noted because not all 2000 census tracts are disaggregations of census tracts existent in 1991. Actually, IBGE's reviews census tracts at every census occasion and, given the need for homogeneity and average size, a new tract might be an addition of parts of two previous census tracts. As a rule, for this thesis, every time this occurred, the maximum value was chosen. Visual comparison of the results with the results provided by Mendonça (2002) confirmed that these inconsistencies were minor and that they did not affect the overall pattern of spatial income distribution.

11.7 CA model data

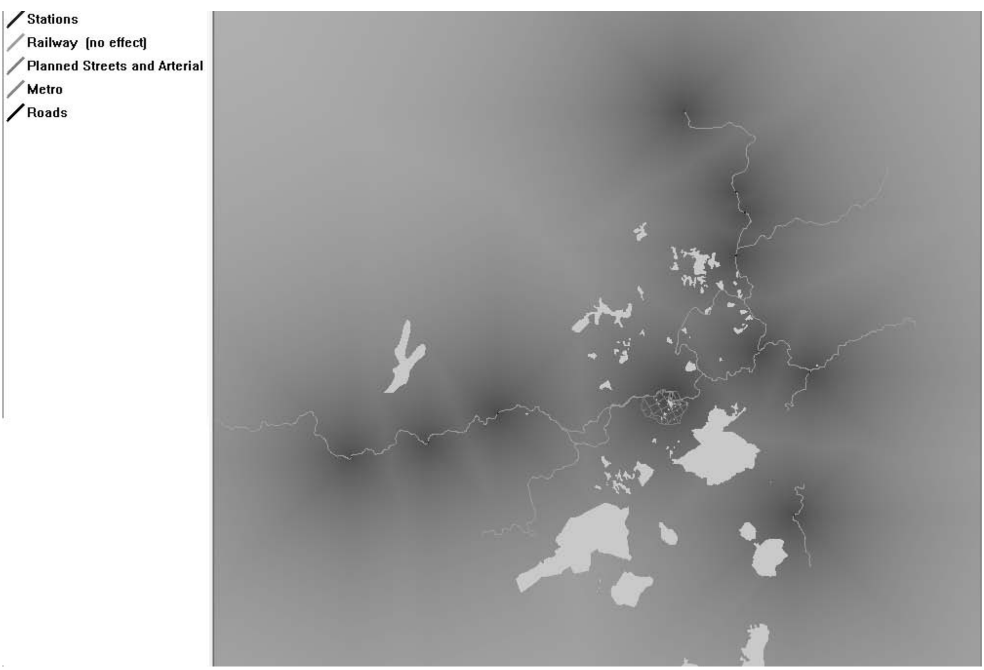


Figure 11.3 Example of accessibility parameters low-income residential area results at time 1897

Table 11.2 Matrix of correlation for sectors (ISS/PBH 2003)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	1.00	0.02	0.29	0.00	0.14	-0.08	0.67	0.04	0.66	0.41	0.46	0.70	0.43	0.05	0.73	0.67
2	0.02	1.00	-0.02	-0.01	0.00	-0.01	-0.04	-0.04	-0.03	0.01	0.20	-0.04	0.27	-0.02	-0.03	-0.03
3	0.29	-0.02	1.00	-0.03	0.03	-0.04	0.73	0.03	0.72	0.19	0.04	0.63	0.44	-0.02	0.55	0.76
4	0.00	-0.01	-0.03	1.00	0.34	0.20	0.03	0.08	0.07	0.21	0.00	-0.03	0.06	0.12	0.00	-0.01
5	0.14	0.00	0.03	0.34	1.00	0.60	0.14	0.04	0.15	0.13	0.07	0.15	0.21	0.05	0.16	0.14
6	-0.08	-0.01	-0.04	0.20	0.60	1.00	-0.06	-0.04	-0.04	-0.06	-0.07	-0.06	0.05	-0.03	-0.07	-0.05
7	0.67	-0.04	0.73	0.03	0.14	-0.06	1.00	0.01	0.97	0.37	0.40	0.94	0.46	-0.03	0.92	0.97
8	0.04	-0.04	0.03	0.08	0.04	-0.04	0.01	1.00	0.10	0.50	-0.07	0.01	0.11	0.87	0.03	0.05
9	0.66	-0.03	0.72	0.07	0.15	-0.04	0.97	0.10	1.00	0.49	0.50	0.92	0.52	0.07	0.91	0.96
10	0.41	0.01	0.19	0.21	0.13	-0.06	0.37	0.50	0.49	1.00	0.54	0.39	0.46	0.54	0.47	0.36
11	0.46	0.20	0.04	0.00	0.07	-0.07	0.40	-0.07	0.50	0.54	1.00	0.53	0.27	-0.06	0.52	0.38
12	0.70	-0.04	0.63	-0.03	0.15	-0.06	0.94	0.01	0.92	0.39	0.53	1.00	0.45	-0.03	0.93	0.95
13	0.43	0.27	0.44	0.06	0.21	0.05	0.46	0.11	0.52	0.46	0.27	0.45	1.00	0.10	0.53	0.49
14	0.05	-0.02	-0.02	0.12	0.05	-0.03	-0.03	0.87	0.07	0.54	-0.06	-0.03	0.10	1.00	0.00	0.00
15	0.73	-0.03	0.55	0.00	0.16	-0.07	0.92	0.03	0.91	0.47	0.52	0.93	0.53	0.00	1.00	0.91
16	0.67	-0.03	0.76	-0.01	0.14	-0.05	0.97	0.05	0.96	0.36	0.38	0.95	0.49	0.00	0.91	1.00
17	0.69	-0.03	0.72	-0.02	0.16	-0.04	0.97	0.03	0.96	0.35	0.42	0.96	0.47	-0.02	0.93	0.99
18	0.73	-0.03	0.66	0.05	0.17	-0.05	0.97	0.03	0.96	0.41	0.48	0.96	0.47	-0.01	0.94	0.97
19	0.18	-0.03	0.04	-0.04	0.03	-0.05	0.19	-0.02	0.29	0.50	0.73	0.26	0.01	0.00	0.24	0.19
20	0.21	-0.03	0.03	-0.07	-0.03	-0.07	0.01	-0.06	0.12	0.12	0.29	0.01	0.37	-0.05	0.09	0.00
21	0.70	-0.04	0.47	0.02	0.15	-0.07	0.87	0.03	0.90	0.56	0.74	0.88	0.37	0.01	0.88	0.85
22	0.86	-0.02	0.18	-0.01	0.17	-0.06	0.70	-0.02	0.69	0.41	0.66	0.75	0.27	-0.03	0.81	0.68
23	0.76	-0.02	0.12	0.15	0.19	-0.04	0.71	0.07	0.70	0.47	0.64	0.74	0.24	0.08	0.81	0.66
24	0.77	-0.02	0.55	0.01	0.17	-0.06	0.94	0.05	0.93	0.49	0.56	0.95	0.53	0.01	0.98	0.93
25	0.74	-0.05	0.22	0.05	0.18	-0.06	0.79	0.01	0.77	0.45	0.66	0.81	0.29	0.00	0.87	0.74
26	0.51	0.01	0.88	0.00	0.13	-0.01	0.90	0.04	0.91	0.29	0.23	0.84	0.53	-0.01	0.78	0.94
27	-0.02	-0.04	-0.04	-0.06	-0.05	-0.07	-0.03	-0.05	0.02	0.16	0.14	0.04	0.24	-0.02	0.18	-0.02
28	0.32	-0.04	0.86	-0.05	0.03	-0.05	0.72	0.03	0.74	0.23	0.12	0.70	0.56	-0.03	0.59	0.78
29	0.74	-0.01	0.76	-0.04	0.09	-0.06	0.84	0.04	0.84	0.33	0.24	0.78	0.57	0.01	0.77	0.87
30	0.76	-0.04	0.40	0.11	0.19	-0.06	0.85	0.08	0.85	0.53	0.59	0.87	0.46	0.06	0.95	0.83
31	0.57	0.06	-0.02	-0.03	0.00	-0.03	-0.04	-0.03	-0.02	0.09	0.00	-0.03	0.22	-0.02	-0.01	-0.02
32	0.65	-0.01	0.38	-0.04	0.13	-0.07	0.71	0.01	0.70	0.52	0.48	0.74	0.57	-0.03	0.78	0.69

17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
0.69	0.73	0.18	0.21	0.70	0.86	0.76	0.77	0.74	0.51	-0.02	0.32	0.74	0.76	0.57	0.65
-0.03	-0.03	-0.03	-0.03	-0.04	-0.02	-0.02	-0.02	-0.05	0.01	-0.04	-0.04	-0.01	-0.04	0.06	-0.01
0.72	0.66	0.04	0.03	0.47	0.18	0.12	0.55	0.22	0.88	-0.04	0.86	0.76	0.40	-0.02	0.38
-0.02	0.05	-0.04	-0.07	0.02	-0.01	0.15	0.01	0.05	0.00	-0.06	-0.05	-0.04	0.11	-0.03	-0.04
0.16	0.17	0.03	-0.03	0.15	0.17	0.19	0.17	0.18	0.13	-0.05	0.03	0.09	0.19	0.00	0.13
-0.04	-0.05	-0.05	-0.07	-0.07	-0.06	-0.04	-0.06	-0.06	-0.01	-0.07	-0.05	-0.06	-0.06	-0.03	-0.07
0.97	0.97	0.19	0.01	0.87	0.70	0.71	0.94	0.79	0.90	-0.03	0.72	0.84	0.85	-0.04	0.71
0.03	0.03	-0.02	-0.06	0.03	-0.02	0.07	0.05	0.01	0.04	-0.05	0.03	0.04	0.08	-0.03	0.01
0.96	0.96	0.29	0.12	0.90	0.69	0.70	0.93	0.77	0.91	0.02	0.74	0.84	0.85	-0.02	0.70
0.35	0.41	0.50	0.12	0.56	0.41	0.47	0.49	0.45	0.29	0.16	0.23	0.33	0.53	0.09	0.52
0.42	0.48	0.73	0.29	0.74	0.66	0.64	0.56	0.66	0.23	0.14	0.12	0.24	0.59	0.00	0.48
0.96	0.96	0.26	0.01	0.88	0.75	0.74	0.95	0.81	0.84	0.04	0.70	0.78	0.87	-0.03	0.74
0.47	0.47	0.01	0.37	0.37	0.27	0.24	0.53	0.29	0.53	0.24	0.56	0.57	0.46	0.22	0.57
-0.02	-0.01	0.00	-0.05	0.01	-0.03	0.08	0.01	0.00	-0.01	-0.02	-0.03	0.01	0.06	-0.02	-0.03
0.93	0.94	0.24	0.09	0.88	0.81	0.81	0.98	0.87	0.78	0.18	0.59	0.77	0.95	-0.01	0.78
0.99	0.97	0.19	0.00	0.85	0.68	0.66	0.93	0.74	0.94	-0.02	0.78	0.87	0.83	-0.02	0.69
1.00	0.98	0.22	0.00	0.87	0.74	0.71	0.95	0.78	0.91	-0.03	0.74	0.85	0.85	-0.02	0.69
0.98	1.00	0.25	0.02	0.91	0.77	0.78	0.97	0.83	0.88	0.01	0.70	0.82	0.90	-0.01	0.71
0.22	0.25	1.00	-0.01	0.59	0.39	0.30	0.25	0.35	0.12	0.00	0.04	0.12	0.26	-0.03	0.15
0.00	0.02	-0.01	1.00	0.06	0.09	0.06	0.12	0.07	-0.01	0.33	-0.01	0.08	0.11	0.27	0.30
0.87	0.91	0.59	0.06	1.00	0.83	0.83	0.91	0.89	0.72	0.07	0.50	0.69	0.89	-0.02	0.66
0.74	0.77	0.39	0.09	0.83	1.00	0.90	0.82	0.91	0.45	-0.03	0.20	0.59	0.85	0.24	0.65
0.71	0.78	0.30	0.06	0.83	0.90	1.00	0.85	0.96	0.43	0.06	0.15	0.46	0.91	0.00	0.65
0.95	0.97	0.25	0.12	0.91	0.82	0.85	1.00	0.90	0.79	0.10	0.61	0.76	0.96	0.00	0.80
0.78	0.83	0.35	0.07	0.89	0.91	0.96	0.90	1.00	0.51	0.08	0.26	0.53	0.93	-0.03	0.70
0.91	0.88	0.12	-0.01	0.72	0.45	0.43	0.79	0.51	1.00	0.01	0.89	0.87	0.65	-0.03	0.56
-0.03	0.01	0.00	0.33	0.07	-0.03	0.06	0.10	0.08	0.01	1.00	-0.03	-0.04	0.15	-0.03	0.17
0.74	0.70	0.04	-0.01	0.50	0.20	0.15	0.61	0.26	0.89	-0.03	1.00	0.77	0.45	-0.03	0.46
0.85	0.82	0.12	0.08	0.69	0.59	0.46	0.76	0.53	0.87	-0.04	0.77	1.00	0.65	0.39	0.60
0.85	0.90	0.26	0.11	0.89	0.85	0.91	0.96	0.93	0.65	0.15	0.45	0.65	1.00	-0.01	0.74
-0.02	-0.01	-0.03	0.27	-0.02	0.24	0.00	0.00	-0.03	-0.03	-0.03	-0.03	0.39	-0.01	1.00	0.12
0.69	0.71	0.15	0.30	0.66	0.65	0.65	0.80	0.70	0.56	0.17	0.46	0.60	0.74	0.12	1.00

Table 11.3 Results of the indices by UDHs

Name UDH [neighborhood] (FJP)	Level of activities	Quality of the neighborhood	Predominance of industry	Predominance of innovative sectors and education
1º DE MAIO-Minaslândia/Boa União	0.024	0.361	0.150	0.491
ABADIA/SÃO GERALDO	0.038	0.614	0.146	0.491
ACAIACA-Beira Linha	0.009	0.235	0.150	0.478
ADELAIDE/ENGENHO NOGUEIRA	0.084	0.633	0.268	0.550
ALPES	0.096	0.143	0.330	0.512
ÁLVARO CAMARGO/DOM BOSCO/PRIMAVERA	0.037	0.669	0.225	0.474
ALVORADA/SÃO MARCOS/NOSSA SENHORA DA PENHA	0.009	0.379	0.138	0.435
ALVORADA-Vila São Benedito	0.000	0.226	0.128	0.435
BAIRRO DAS INDÚSTRIAS/BAIRRO NOVO DAS INDUSTRIAS	0.032	0.366	0.228	0.500
BALEIA/PARAÍSO/VERA CRUZ/SAUDADE	0.034	0.426	0.158	0.491
BANDEIRANTES/S. LUIZ/PAQUETÁ/VILA MILITAR/UFMG	0.037	0.654	0.209	0.597
BARREIRO DE BAIXO/SANTA MARGARIDA/TEIXEIRA DIAS	0.113	0.646	0.221	0.510
BARREIRO-Mangueiras	0.005	0.334	0.137	0.450
BARRO PRETO/STO AGOSTINHO/CENTRO-Pça. da Estação	1.000	0.950	0.121	0.000
BARROCA/PRADO	0.224	0.941	0.532	0.627
BELVEDERE/MANGABEIRAS/COMITECO	0.053	0.971	0.121	0.495
BETÂNIA/PALMEIRAS	0.052	0.674	0.213	0.558
BOA VISTA/CASA BRANCA	0.072	0.402	0.251	0.489
BRASIL INDUSTRIAL/VILA SALES/CONJ. GETÚLIO VARGAS	0.019	0.573	0.173	0.472
BRAÚNAS/XANGRI-LÁ/SUL DO ZOOLÓGICO	0.012	0.603	0.140	0.538
BURITIS/ESTORIL/MANSÕES	0.040	0.905	0.132	0.857
CABANA PAI TOMÁS/FAV. NOVA GAMELEIRA/VILA DIVINÉIA	0.021	0.419	0.186	0.506
CAIÇARA/CONJUNTO PRESIDENTE JUSCELINO	0.115	0.832	0.324	0.594
CARLOS PRATES/LAGOINHA/BONFIM	0.213	0.800	0.766	0.523
CARMO/SION	0.199	0.984	0.089	0.530
CÉU AZUL-Vila dos Anjos	0.073	0.243	0.219	0.476
CHÁCARAS REUNIDAS SANTA INÊS-Jaqueline	0.014	0.318	0.137	0.476
CIDADE NOVA/SILVEIRA	0.130	0.829	0.125	0.512
CLÓRIS/SANTO INÁCIO/CAMPO ALEGRE/PLANALTO NOVO	0.096	0.773	0.316	0.488
COLINA-Vila Átila de Paiva	0.001	0.575	0.128	0.510
CONCÓRDIA/NOVA FLORESTA/GRAÇA	0.072	0.833	0.328	0.523
CONCÓRDIA-Vila Tiradentes	0.071	0.831	0.340	0.531
CONJUNTO ALÍPIO DE MELO/SERRANO	0.064	0.668	0.176	0.473

Name UDH [neighborhood] (FJP)	Level of activities	Quality of the neighborhood	Predominance of industry	Predominance of innovative sectors and education
CONJUNTO SÃO PAULO/JARDIM OURO BRANCO	0.059	0.391	0.208	0.481
CONJUNTO SERRA VERDE/HIPÓDROMO	0.019	0.318	0.159	0.474
CONJUNTO SERRA VERDE-Vila Serra Verde	0.019	0.318	0.162	0.494
COPACABANA/JARDIM LEBLON	0.026	0.609	0.139	0.479
CORAÇÃO EUCARÍSTICO/DOM CABRAL/ MINAS BRASIL	0.103	0.833	0.281	0.500
CRUZEIRO/ANCHIETA	0.164	1.000	0.187	0.531
DOM JOAQUIM-Vila de Sá	0.024	0.628	0.204	0.450
DOM SILVERIO/NAZARÉ/BELMONTE	0.009	0.297	0.149	0.477
DONA CLARA/JARDIM SANTA BRANCA/ ITAPOÃ	0.073	0.793	0.271	0.603
ESPLANADA/POMPEIA-Vila Nossa Senhora do Rosário	0.048	0.665	0.218	0.477
ESPLANADA/VERA CRUZ/POMPÉIA	0.115	0.669	0.242	0.493
FAVELA 31 DE MARÇO	0.049	0.729	0.313	0.494
FAVELA BARÃO HOMEM DE MELO	0.102	0.604	0.363	0.443
FAVELA CABEÇA DE PORCO	0.004	0.234	0.128	0.577
FAVELA CARLOS PRATES (Vila S.Francisco das Chagas)	0.298	0.500	1.000	0.501
FAVELA DA CEMIG	0.002	0.259	0.130	0.491
FAVELA DA SERRA	0.021	0.000	0.127	0.477
FAVELA DO AEROPORTO	0.045	0.338	0.207	0.683
FAVELA DO ATERRO SANITÁRIO (Vila da Paz)	0.018	0.666	0.210	0.479
FAVELA DO OURO PRETO	0.087	0.358	0.230	0.487
FAVELA DO PERRELA	0.750	0.756	0.281	0.601
FAVELA DO SION	0.033	0.775	0.086	0.493
FAVELA MONTE SÃO JOSÉ	0.067	0.997	0.128	0.511
FAVELA SANTA LÚCIA	0.073	0.120	0.104	0.490
FLAMENGO	0.003	0.227	0.138	0.474
FLORAMAR/JARDIM GUANABARA/JARDIM PAMPULHA	0.034	0.321	0.192	0.471
FLORESTA/COLÉGIO BATISTA	0.294	0.835	0.299	0.456
FUNCIONÁRIOS	0.666	0.999	0.165	0.557
GAMELEIRA/NOVA SUIÇA	0.132	0.843	0.418	0.470
GLALIJÁ/VILA SÃO JOSÉ	0.097	0.625	0.511	0.506
GLÓRIA/CONJUNTO FILADÉLFIA	0.069	0.665	0.285	0.478
GLÓRIA/SÃO SALVADOR/COQUEIROS	0.048	0.617	0.252	0.479
GOIÂNIA/PIRAJÁ	0.017	0.327	0.208	0.435
GRAJAU/GUTIERREZ	0.174	0.982	0.237	0.625
GUARANI/TUPI	0.044	0.354	0.211	0.479
HAVAÍ-Ventosa	0.093	0.398	0.320	0.443
HELIÓPOLIS/AVIAÇÃO	0.044	0.595	0.213	0.486
HIPERCENTRO-Praça Afonso Arinos/Praça da Liberdade	0.817	0.922	0.000	0.337

Name UDH [neighborhood] (FJP)	Level of activities	Quality of the neighborhood	Predominance of industry	Predominance of innovative sectors and education
HORTO/SANTA INÊS/NOVA VISTA	0.064	0.729	0.245	0.505
INDAIÁ/HOSPITAL INCONFIDENTES-Vila Santa Rosa	0.119	0.756	0.865	0.683
INDEPENDÊNCIA 1ª SEÇÃO	0.008	0.334	0.180	0.450
INDEPENDÊNCIA/MINEIRÃO	0.005	0.334	0.133	0.450
IPIRANGA/PALMARES	0.057	0.735	0.274	0.515
ITATIAIA/STA.TEREZINHA-Conj. Habitacional/SARAMENHA	0.036	0.648	0.179	0.498
JARAGUÁ/INDAIÁ/UNIVERSITÁRIO/ LIBERDADE	0.045	0.745	0.265	0.683
JARDIM DOS COMERCIÁRIOS/NOVA AMÉRICA 3ª SEÇÃO	0.018	0.243	0.152	0.474
JARDIM DOS COMERCIÁRIOS-Vila Mantiqueira	0.011	0.152	0.141	0.474
JARDIM EUROPA/NOVA YORK	0.026	0.279	0.134	0.483
JARDIM LEBLON-Vila Apolônia	0.030	0.274	0.137	0.479
JARDIM LEBLON-Vila Jardim Leblon	0.100	0.318	0.223	0.479
JARDIM MONTANHÊS	0.033	0.349	0.235	0.492
JATOBÁ/MARILÂNDIA/INDUSTRIAL 4ª SEÇÃO	0.012	0.572	0.137	0.476
JD. ATLÂNTICO/STA. AMÉLIA/CONJUNTO STA. MÔNICA	0.060	0.765	0.277	0.540
JOÃO PINHEIRO/CONJUNTO CALIFÓRNIA	0.029	0.688	0.242	0.492
JULIANA/JAQUELINE/CANAÃ	0.008	0.413	0.146	0.476
LAGOA/LAGOINHA LEBLON	0.038	0.318	0.187	0.478
LIBERDADE-Favela Vila Rica	0.063	0.757	0.249	0.683
LINDÉIA	0.028	0.575	0.147	0.476
LOURDES/SANTO AGOSTINHO	0.573	0.999	0.103	1.000
LUXEMBURGO/CID. JARDIM/CORAÇÃO DE JESUS/VILA PARIS	0.113	0.993	0.112	0.522
MANGUEIRA	0.067	0.267	0.218	0.477
MARAJÓ/HAVAÍ	0.020	0.759	0.157	0.465
MILIONÁRIOS	0.023	0.300	0.185	0.491
MILIONÁRIOS-Vila Nova dos Milionários	0.033	0.350	0.238	0.491
MONS.MESSIAS/CELESTE IMPÉRIO/AEROP. CARLOS PRATES	0.229	0.833	0.539	0.527
MONTIVIDEU/VILA MARIA/JARDIM VITÓRIA	0.005	0.185	0.129	0.474
MORRO DO QUEROSENE	0.060	0.155	0.189	0.490
NOVA CINTRA	0.025	0.403	0.198	0.505
NOVA GAMELEIRA/JARDINÓPOLIS/MADRE GERTRUDES	0.030	0.632	0.235	0.506
NOVA GRANADA/JARDIM AMÉRICA	0.094	0.833	0.320	0.479
NOVA SUIÇA-Vila Guaratã	0.115	0.773	0.404	0.443
PADRE EUSTÁQUIO/PROGRESSO	0.338	0.833	0.774	0.501
PADRE EUSTÁQUIO-Vila dos Marmiteiros	0.348	0.831	0.763	0.501

Name UDH [neighborhood] (FJP)	Level of activities	Quality of the neighborhood	Predominance of industry	Predominance of innovative sectors and education
PAQUETÁ-Vila Paquetá	0.021	0.000	0.128	0.487
PARQUE SÃO JOÃO BATISTA/SANTA CRUZ	0.066	0.662	0.356	0.484
PARQUE SÃO PEDRO/CONJUNTO SÃO PEDRO	0.092	0.461	0.208	0.485
PARQUE URSOLINA DE MELO/CASTELO/ RECREIO/ITAMARATI	0.046	0.731	0.237	0.489
PAULO VI-Beira Linha	0.023	0.156	0.154	0.473
PEDREIRA PRADO LOPES	0.020	0.235	0.206	0.478
PINDORAMA/COQUEIROS	0.032	0.405	0.233	0.479
PIRATININGA-Conquista União	0.015	0.574	0.137	0.476
PIRATININGA-Vila Santa Mônica	0.070	0.225	0.220	0.476
PROVIDÊNCIA/MATADOURO/1º DE MAIO/1º DE NOVEMBRO	0.023	0.393	0.183	0.488
RECANTO NOSSA SENHORA DA BOA VIAGEM	0.004	0.184	0.132	0.474
RENASCENÇA/CACHOEIRINHA/HUMAITÁ	0.110	0.729	0.457	0.501
RIBEIRO DE ABREU/CONJ. RIBEIRO DE ABREU	0.006	0.188	0.134	0.477
S. FRANCISCO/SENHOR BOM JESUS/ RIACHUELO/ERMELINDA	0.049	0.663	0.249	0.553
SAGRADA FAMÍLIA	0.289	0.833	0.321	0.522
SAGRADA FAMÍLIA-Pátio da Estação Ferroviária	0.142	0.830	0.223	0.521
SALGADO FILHO/GRANJAS REUNIDAS/ TIRADENTES	0.071	0.746	0.288	0.465
SALGADO FILHO-Ventosa	0.122	0.832	0.391	0.443
SANTA BRANCA/LEBLON-Vila Flamengo	0.042	0.301	0.196	0.481
SANTA CECÍLIA/CASTANHEIRA/DIAMANTE	0.007	0.417	0.167	0.462
SANTA CECÍLIA-Vila Formosa	0.002	0.334	0.162	0.450
SANTA EFIGÊNIA/NOVO SÃO LUCAS	0.512	0.688	0.239	0.596
SANTA HELENA/FLÁVIO MARQUES LISBOA/MIRAMAR	0.017	0.343	0.156	0.492
SANTA MARIA/VILA OESTE	0.079	0.682	0.442	0.501
SANTA MÔNICA	0.113	0.583	0.240	0.479
SANTA MÔNICA-Universo/Copacabana	0.067	0.301	0.196	0.479
SANTA MÔNICA-Vila Nossa Senhora Aparecida	0.060	0.299	0.314	0.479
SANTA TEREZA/SAGRADA FAMÍLIA	0.203	0.833	0.191	0.476
SÃO BENTO/SANTA LÚCIA	0.079	0.982	0.091	0.498
SÃO BERNARDO-Vila Aeroporto	0.046	0.665	0.209	0.484
SÃO GABRIEL/CONJUNTO SÃO GABRIEL	0.037	0.313	0.195	0.478
SÃO GABRIEL-Vila Três Marias	0.013	0.226	0.179	0.478
SÃO JOÃO BATISTA-Vila São João Batista	0.069	0.376	0.308	0.479
SÃO PAULO-Carioca/Modelo	0.080	0.309	0.111	0.462
SÃO PAULO-Vila São Paulo	0.080	0.317	0.111	0.435
SÃO PEDRO/SANTO ANTÔNIO	0.297	1.000	0.125	0.605

Name UDH [neighborhood] (FJP)	Level of activities	Quality of the neighborhood	Predominance of industry	Predominance of innovative sectors and education
SERRA/SÃO LUCAS	0.235	1.000	0.262	0.547
SOLAR DO BARREIRO/PILAR	0.003	0.327	0.139	0.459
SOLIMÕES/JARDIM FELICIDADE	0.009	0.188	0.136	0.473
STO. ANDRÉ/BOM JESUS/NOVA ESPERANÇA/S. LEOPOLDO	0.069	0.670	0.352	0.553
TAQUARIL/CASTANHEIRAS	0.004	0.077	0.132	0.474
TIROL/COLINA/MALDONADO/CONJ. TÚNEL DE IBIRITÉ	0.034	0.516	0.243	0.479
TUPI	0.058	0.317	0.190	0.470
TUPI-Vila Ribeiro de Abreu	0.004	0.318	0.135	0.470
UNIÃO/DOM JOAQUIM/FERNÃO DIAS/ EYMARD	0.052	0.618	0.226	0.462
VALE DO JATOBÁ	0.012	0.334	0.234	0.450
VENDA NOVA-Lagoinha/Nova América/ SESC	0.020	0.266	0.167	0.477
VENDA NOVA-Letícia/Candelária/Rio Branco	0.070	0.548	0.207	0.479
VENDA NOVA-São João Batista/Sinimbu	0.090	0.605	0.289	0.477
VILA CACHOEIRINHA	0.064	0.461	0.308	0.550
VILA MARIA	0.012	0.168	0.128	0.479
VILA SÃO JOSÉ	0.023	0.315	0.209	0.477
VILA SÃO TOMÁS-Aeroporto da Pampulha	0.032	0.503	0.186	0.491
VILA SUMARÉ/BOM JESUS	0.010	0.488	0.160	0.553

Table 11.4 Parameters used for Neighborhood Influence (N) for other actors

	Small service and commerce										
Distance (in cells)	0.00	1.00	1.41	2.00	2.24	2.83	3.00	3.16	3.61	4.00	4.12
High-income	10.00	4.35	2.01	1.61	1.50	1.21	1.13	1.05	0.84	0.65	0.59
Average-income	5.50	1.60									
Low-income	0.00	0.00									
Small service and commerce	100.00	10.00									
Small industry											
In the market											
Large industry											
Schools and Campuses											
Water and Reservation Areas											
Large Service and Commerce	10.00	6.67	5.29	3.33	2.55	0.57					
Institutional buildings	10.00	5.00	2.93								
Airports											
	Small industry										
Distance (in cells)		1.00	1.41	2.00	2.24	2.83	3.00	3.16	3.61	4.00	4.12
High-income	-1.00	-0.29									
Average-income	-1.00	-0.59	-0.42	-0.18	-0.09						
Low-income	1.00	0.63	0.47	0.25	0.16						
Small service and commerce											
Small industry	55.00	5.50									
In the market											
Large industry											
Schools and Campuses											
Water and Reservation Areas											
Large Service and Commerce											
Institutional buildings											
Airports											

4.24	4.47	5.00	5.10	5.39	5.66	5.83	6.00	6.80	6.32	6.40	6.71	7.00	7.70	7.21	7.28	7.62	7.81	8.00
0.53	0.42	0.16	0.11															
																0.01	0.03	0.05
																0.01	0.04	0.06

4.24	4.47	5.00	5.10	5.39	5.66	5.83	6.00	6.80	6.32	6.40	6.71	7.00	7.70	7.21	7.28	7.62	7.81	8.00
															0.07	0.97	1.49	2.00

	In the market										
Distance (in cells)		1.00	1.41	2.00	2.24	2.83	3.00	3.16	3.61	4.00	4.12
High-income	1.00	0.29									
Average-income	0.55	0.16									
Low-income											
Small service and commerce											
Small industry											
In the market	120.00										
Large industry											
Schools and Campuses											
Water and Reservation Areas											
Large Service and Commerce											
Institutional buildings	5.90	2.95	1.73								
Airports											

4.24	4.47	5.00	5.10	5.39	5.66	5.83	6.00	6.80	6.32	6.40	6.71	7.00	7.70	7.21	7.28	7.62	7.81	8.00

Resumo em português (Portuguese extended summary)

Esta tese aplica conceitos oriundos de dois campos da literatura no estudo do desenvolvimento urbano: o da economia urbana – baseado na noção de monocentrismo – e o de sistemas complexos auto-organizáveis e modelagem com autômatos celulares (CA), que permitem noções de multicentrismo e desenvolvimento evolucionário. O mercado de bens imóveis é analisado, pois que reflete a complexidade do meio urbano e está intimamente imbuído em suas múltiplas causalidades e determinações. Heterogeneidade socioeconômica e a noção de vizinhança – vista como essencial em estudos urbanos – são introduzidos nos nossos modelos. Ou seja, vizinhanças são consideradas elementos essenciais para explicar preços de bens imóveis e desenvolvimento urbano local de longo prazo na metrópole de Belo Horizonte, Brasil. O melhor entendimento desses processos pode contribuir na melhoria do trabalho de analistas de políticas públicas ao provê-los com maior arsenal de ferramentas hermenêuticas.

Desenvolvimento urbano nas comunidades epistêmicas de pesquisa

As cidades são fenômenos socioeconômicos complexos que se espraiam fisicamente no espaço e têm características de permanência, continuidade e evolução no tempo (Jacobs, 1970; Serra, 1987; Soja, 2000).

Ambientes urbanos podem ser analisados a partir de diferentes pontos de vista conceituais. Após identificar e definir a cidade como objeto de estudo no século XIX (Benevolo, 1980), várias disciplinas científicas debruçaram-se em teorias, conceitos e metodologias para explicar e descrever os intrincados processos urbanos.

Sociólogos urbanos propuseram modelos ecológicos no início do século XX (Park, Burgess *et al.*, 1925), almejando melhor entendimento da segregação social e espacial na cidade. Nos anos 1960, o estudo da economia urbana ganhou avanço teórico com o trabalho seminal de Alonso sobre uso do solo e localização (1964). Sua abordagem cartesiana serviu como subsídio para a rica discussão teorética que se seguiu (Capello and Nijkamp, 2004b), que considera distância ao centro como o fator primordial a explicar a configuração econômica da cidade. Contemporaneamente, Jane Jacobs (1970) aborda a (macro)economia da Cidade de forma distinta, alegando que a diversidade e a proximidade dos atores – que ocorrem exclusivamente na cidade – são o que permite o desenvolvimento urbano.

Morfologia e estrutura urbanas receberam maior atenção desde os anos 1990, com ênfase em fatores econômicos implícitos (White and Engelen, 1993), em questões segregacionistas (Portugali, 2000); em morfologia (Batty, M. and Longley, 1994) ou em transporte e uso do solo (Waddell, 2002). De forma ampla, cientistas sociais e geógrafos urbanos confirmam a configuração socialmente dividida da cidade (Caldeira, 2000). A complexidade das cidades pode ser vista ainda como muito diferenciada em termos de composição social; poder econômico desigual e características segregacionistas espaciais (Lemos, Ferreira *et al.*, 2004), que são mais pronunciadas em aglomerações urbanas em países em desenvolvimento. Esses fatores nos inspiram a incorporar heterogeneidade espacial e social como fator relevante na nossa análise urbana.

Mercados imobiliários urbanos e complexidade espacial

Os mercados de imóveis urbanos influenciam e são influenciados pela complexidade urbana e pela percepção dos habitantes da cidade. Essa causalidade e influência dual fazem com que o mercado seja tão complexo e multifacetado quanto as cidades o são. Os efeitos de atração e repulsão entre atores são dinâmicos, nos quais a localização do domicílio é determinado em parte pela localização do mercado de trabalho e vice-versa (Mills and Nijkamp, 1987) ou, como alguns sugerem, em um padrão no qual "os empregos seguem os domicílios" (Steinnes, 1982).

Ademais, o mercado de imóveis representa parcela entre 45% e 75% do estoque e fluxos de riqueza em países desenvolvidos ou áreas metropolitanas de países em desenvolvimento (Ibbotson, Siegel *et al.*, 1985). DiPasquale e Wheaton (1996) demonstram que residências representam o maior item de consumo no orçamento dos cidadãos. No contexto urbano, no qual a terra é um bem diferenciado, economias e deseconomias de aglomeração podem ser vistas como competição por espaços mais bem localizados, compensada pelos preços que aumentam com a competição. O preço do imóvel urbano é crucial na determinação dos fatores dispersivos que agem contra a aglomeração (Krugman, 1996). A dinâmica da configuração urbana espacial pode ser descrita como o resultado de efeitos retro-alimentadores positivos e negativos (Batty, M., 2005b). De um lado, a aglomeração exerce efeito de atração para empresas e residências, de outro, a maior competição por espaço encarece a terra e serve como força dispersiva.

Mercados imobiliários também apresentam características específicas tais como durabilidade, processo de ajuste lento e custoso, rigidez espacial, informação limitada e inconsistências de oferta (Whitehead, 1999, p. 1565). Todas estas peculiaridades do mercado e seu ambiente levam a práticas de avaliação que dependem fortemente do conhecimento individual das idiossincrasias do mercado. A literatura brasileira (Liporoni, Neto *et al.*, 2003) e a legislação técnica (Abnt, 2004) sobre o tema são baseados na prática cotidiana, nas quais os "valores de mercado" determinados pelos corretores experientes regulam o sistema de avaliação. A literatura internacional também aponta para abordagens discricionárias (Kauko, 2002).

Identidade dos lugares: o papel da vizinhança

Fator crucial a ser considerado ao analisar o mercado de imóveis urbanos é o fato de que ele reflete componentes de percepção pela população do que é a cidade e o que suas partes representam. Isto é relevante para corretores de imóveis. A maioria dos estudos econômicos inclui grande número de variáveis, porém, sem considerar as noções de identidade dos lugares. Variáveis comumente utilizadas, tais como distância à rodovia ou ao posto de saúde, não incluem informação referente à unidade espacial na qual a observação se encontra. As pessoas, por outro lado, aprendem sobre espaço e localização de forma intuitiva, o que é equivalente a dizer que associam sentido a porções ou segmentos da cidade (Tversky, 2003). Nesta tese defende-se a idéia de que influências locais e a identidade da vizinhança e seus reflexos para a determinação dos preços de bens imóveis não é otimamente tratada em pesquisas anteriores e, assim, tentamos contribuir nessa direção.

Surpreendentemente, a disciplina de economia urbana ou os manuais de finanças do mercado imobiliário não mencionam o tema de vizinhança no sentido de que representam uma comunidade com determinada identidade (Durlauf, 2004). Alguns trabalhos de determinação de preços imobiliários enfatizam submercados (Bourassa, Steven C., Hamelink *et al.*, 1999), no intuito de identificar aglomerados de residências com características homogêneas por meio de técnicas exploratórias de dados. Megbolugbe (1996) afirma que submercados homogêneos são uma das definições de vizinhanças. Acreditamos, entretanto, que a definição que estabelece vizinhança como aquela que "possui uma identidade ou coesão social" (também proposta por Megbolugbe) mais provavelmente influencia preços imobiliários. De fato, a vizinhança como construção sócio-espacial é relevante para a modelagem econômica do mercado imobiliário.

Cientistas sociais produzem conceitos que explicitam por que vizinhanças, identidade e padrões de segregação espacial são importantes (Blakely, 1997; Galster, 2001). No entanto, não os relacionam explicitamente aos preços dos imóveis.

A natureza dinâmica e baseada nos atores em modelos de economia urbana

Em relação a análise *cross-section* usualmente utilizada em análises de economia urbana, nosso entendimento é de que os preços de imóveis urbanos em vizinhanças influenciam preços na vizinhança no próximo período, isto é, o mercado reflete mecanismos retro-alimentadores positivos e negativos. Esse fato indica que a análise de mercado deva ser dinâmica. Entretanto, a literatura que estuda modelos dinâmicos de configuração urbana, tais como Benenson e Torrens (2004) não enfatiza as diferenças de comportamento observadas em atores com diferentes níveis de renda. Usualmente trabalha-se com a noção de uso do solo residencial, não apresentando diferenciação maior do que essa. Mesmo a distinção ampla de atores urbanos do estudo de Hagoort (2006) inclui apenas um ator chamado de 'residencial'. Tal divisão não é de forma alguma suficiente para uma análise que se propõe intraurbana. A literatura inclusive já demonstrou que o comportamento de cidadãos com níveis de renda distintos apresentam padrões diversos (Zietz, Zietz *et al.*, 2008). Nesse caso, disparidades sociais são cruciais para o entendimento da influência da cidade nos preços dos bens imóveis.

Análises do mercado imobiliário como as de Dowall (1995), Lucas e Rossi-Hansberg (2002) e Wheaton e Nechayev (2005) trazem abordagens que focam a cidade de forma estática. Outros autores, tais como Henderson (1974), Fujita, Krugman e Venables (1999) e Capello (2002) discutem relações entre cidades, hierarquias e tamanho ótimo de cidades. Estes estudos têm em comum a análise da cidade como objeto homogêneo. Se considerarmos a proposta da Pós-Metrópole de Soja (2000), segundo a qual a cidade é heterogênea e multifacetada, uma escala de análise maior – no nível intraurbano – faz-se necessária.

Motivações para o estudo

Potencialmente, as duas vertentes da literatura (economia urbana e modelagem complexa) juntas podem criar valor-adicionado. De um lado, em relação a análise da economia urbana na escala intraurbana, distância ao Centro Comercial (CBD) ou a outros pontos de interesse somente parecem não ser capazes de capturar a rica heterogeneidade espacial observada na cidade. De outro, a consideração de que os cidadãos percebem o espaço de forma intuitiva e adicionam valores a ele sugere que bairros – os prováveis recipientes desses atributos – devam ser explicitamente entendidos, descritos e incluídos na análise do mercado de imóveis urbanos. As ciências sociais proveem interessantes indicativos da importância de bairros, identidade espacial e padrões de segregação espacial. Entretanto, elas não as relacionam diretamente aos preços de bens imóveis.

Em primeiro lugar, a idéia essencial de repulsão exercida por fatores como o preço da terra na economia urbana não estão claramente definidos, embora implicitamente presentes, em modelos de autômatos celulares. Em segundo lugar, influências locais, centrais na modelagem com CA, estão presentes somente marginalmente na economia urbana por meio de submercados homogêneos e análise espacial baseada em distâncias. Em terceiro lugar, heterogeneidade está presente no nível teórico da economia urbana, já que agentes com preferências diferenciadas são permitidos, porém, não no ambiente urbano, usualmente caracterizado como uma planície homogênea. Modelos de CA, por outro lado, não impõem homogeneidade aos atores, mas seus trabalhos empíricos utilizam-se somente de uma categoria residencial. Em quarto lugar, no intuito de ressaltar as diferenças, pode-se dizer que se o foco da economia urbana é na análise do equilíbrio, sistemas complexos e auto-organizáveis focalizam essencialmente a dinâmica do processo.

Finalmente, a visão desta tese é a de que para um melhor entendimento do ambiente urbano complexo contemporâneo e, consequente, intervenção nele, analistas de políticas públicas em especial e cientistas em geral deveriam obter uma compreensão da cidade que vai além da morfologia urbana e está intrinsecamente ligada a fatores econômicos: os indicadores apresentados pela economia urbana de um lado, e a atuação social dos atores, no espaço, de outro. Para isso, este trabalho enfoca uma melhor inclusão de bairros cognitivamente definidos em modelos econômicos, bem como melhor modelagem de efeitos negativos que considerem heterogeneidade social em modelos de autômatos celulares, ambos ausentes na literatura corrente.

Portanto, o foco do trabalho se divide em três partes: (a) no papel dos bairros e sua identidade, (b) no comportamento diferenciado de atores urbanos heterogêneos e (c) na gama de atributos urbanos associados aos imóveis dinâmica e estaticamente. Em conjunto, estes fatores conceitualmente indicam que a literatura em economia urbana e em geografia urbana não apresentam bairros como fatores explicativos em modelos econômicos, bem como modelagem que inclua explicitamente efeitos negativos e heterogeneidade social.

Perguntas da pesquisa

Esta tese, portanto, trabalha com a hipótese de que a economia urbana e a abordagem de autômatos celulares são mutuamente relacionadas e que ambas as vertentes de pesquisa se beneficiam ao incorporar elementos explicativos uma da outra. A análise empírica utilizada estende modelos em ambas vertentes utilizando fertilização-cruzada e mecanismos causais. As duas perguntas da pesquisa são:

1. Modelos de preços imobiliários urbanos se beneficiam da incorporação de conceitos de identidades de vizinhança e heterogeneidade social?
2. Modelos de autômatos celulares de desenvolvimento urbano aumentam seu entendimento da dinâmica urbana, quando se incluem heterogeneidade social e mecanismos econômicos negativos de retroalimentação?

Desenho da pesquisa e metodologias

No intuito de responder a essas questões, o seguinte trajeto teórico, metodológico e empírico é observado.

A primeira parte da tese revê as principais questões teóricas de uma perspectiva da economia urbana e, subsequentemente, de abordagens alternativas que enfatizam interações locais. Baseado nesta revisão, o estudo propõe um modelo de autômatos celulares que especifica efeitos negativos dinâmicos.

A segunda parte enfoca algumas aplicações baseadas nas discussões teóricas feitas. A análise empírica em um ambiente de rápido crescimento e contexto socioeconômico extremamente heterogêneo na Região Metropolitana de Belo Horizonte, Brasil, é conduzido no intuito de facilitar a elaboração das respostas colocadas.

Técnicas de geoprocessamento e análise de componentes principais (PCA) – usualmente utilizadas no contexto da geografia física e estudos do ambiente – são aplicados para agregar grande quantidade de informações socioeconômicas distribuídas espacialmente.

Um exercício de análise econométrica é apresentado na sequência. Visa distinguir melhor o componente espacial daquele intrínseco ao imóvel na composição do preço. O modelo é testado para seu componente espacial e a análise quantílica é incorporada e testada por meio de uma estimação por variáveis instrumentais (IVQR). O método de estimação se mostra robusto a heterocedasticidade e a valores atípicos e tem desempenho superior ao do método de momentos generalizados (GMM) ou máxima verossimilhança (ML).

Os dados necessários para a validação do modelo de autômatos celulares são detalhados. Dados por setores censitários são utilizados na produção da configuração espacial de atores residenciais em Belo Horizonte, classificados por renda. Algumas medidas de quantificação de mapas são apresentadas e dão apoio à comparação dos resultados simulados com aqueles observados.

Uma aplicação do modelo estendido proposto de autômatos celulares é aplicado ao caso da Região Metropolitana de Belo Horizonte. O objetivo da aplicação é incorporar a interação espacial entre atores dinamicamente e, portanto, oferecer novos indicativos dos processos que levam à determinação dos preços dos imóveis. Os parâmetros escolhidos e seu poder de explicação em conjunto com os resultados do modelo original e o adaptado são apresentados. Os resultados são validados utilizando-se de comparação com dados da configuração espacial e de preços de imóveis observados.

Conclusões

Concluímos que o modelo econométrico apresenta boa performance e permite subsídios em relação a quais fatores (e em qual proporção) influenciam a formação de preços de imóveis urbanos. O modelo é preciso e consistente e possibilita sua utilização para inferências estatísticas. É descritivo no sentido em que dada certa conjuntura, o modelo pode ser aplicado (e responde) imediatamente. O modelo, entretanto, não contribui para discernir quais elementos levam dinamicamente a dada configuração.

O modelo de autômatos celulares por outro lado fornece exatamente isso. O modelo indica que características da vizinhança complementadas com informações gerais sobre acessibilidade permitem a modelagem de configuração de preços e uso do solo comparáveis àquelas empiricamente obtidas, ainda que ao fim de período especialmente longo. Comparativamente, entretanto, sua previsão é menos precisa, uma vez que a abordagem enfoca a estrutura geral e não a evolução de porções específicas da cidade. Portanto, o modelo de autômatos celulares não inclui alguns dos fatores capturados pelo modelo econométrico, tais como mudanças rápidas e influências específicas. Alguns destes fatores poderiam ter sido contemplados se o estudo de caso fosse mais detalhado e aplicado em menor amplitude temporal.

Acreditamos que a tese apresentada realça a importância da consideração de influências locais e a vizinhança como elementos centrais do mercado imobiliário em relação à complexidade urbana. Também demonstramos que heterogeneidades espaciais e sociais são elementos inerentes à estrutura urbana e seus processos determinantes. Efeitos negativos também contribuem para o entendimento do desenvolvimento urbano e deveriam ser incluídos e estudados em modelos futuros de modelagem urbana. Esperamos ainda ter avançado na interface entre campos da ciência próximos (teórica e metodologicamente). E, finalmente, a investigação que teve Belo Horizonte e sua região metropolitana, como estudo de caso fornece dados e análises que podem potencialmente contribuir no seu planejamento e execução de políticas públicas.

Especificamente, em relação à primeira pergunta da pesquisa: 'Modelos de preços imobiliários urbanos se beneficiam da incorporação de conceitos de identidades de vizinhança e heterogeneidade social?', podemos responder afirmativamente. Há indicativos de que informações relevantes não são incluídas em modelos nos quais heterogeneidade social – expressa pela regressão quantílica do capítulo 6 – não são incorporadas. Ademais, demonstramos que ao nível de análise da vizinhança – a incorporação da percepção do espaço pelos cidadãos – adiciona valores específicos ao preço do imóvel urbano. A descrição da vizinhança por meio de um grupo de índices permite a captação de similaridades e idiossincrasias do tecido urbano que melhora nosso entendimento dos processos urbanos.

A heterogeneidade especial influencia os preços dos imóveis em escala até maior do que no nível da vizinhança. Esta conclusão sugere – contrariamente ao observado na literatura – que modelos (espacialmente) detalhados de formação de preços são necessários, ao invés de modelos gerais que assumem no nível teórico a cidade como elemento homogêneo e não heterogêneo.

Quanto à heterogeneidade social, o estudo confirma que escolhas e preferências dos consumidores são substancialmente distintos quando considerados níveis de valor dos imóveis distintos. Análises que não considerem tais disparidades e enfatizem resultados na média (quantile.5) podem indicar erros não só de magnitude como até de tendência invertida. Esse fato é relevante para executores de políticas públicas, negociantes e corretores quando ponderando o impacto de mudanças na configuração urbana e seu rebatimento nos preços.

A segunda pergunta: 'Modelos de autômatos celulares de desenvolvimento urbano aumentam seu entendimento da dinâmica urbana quando se incluem heterogeneidade social e mecanismos econômicos negativos de retroalimentação?' é respondida pelo modelo de autômatos celulares do capítulo 3. O modelo proposto simula a dinâmica do desenvolvimento urbano e explicitamente distingue entre forças positivas e negativas de aglomeração. Ao fazê-lo, avança na proposta de modelagem desenhada por White e Engelen (1993) na qual todos os efeitos de influência local são modelados de maneira agregada. A modelagem dos efeitos de forma separada (aglomeração e desaglomeração) permite a contribuição de obter-se como resultado a configuração espacial dinâmica de preços. Isso permite uma segunda rodada de validação do modelo baseada em um segundo conjunto independente de dados, como recomendado por Batty e Torrens (2005).

Adicionalmente, a aplicação do modelo proposto feita no capítulo 8 mostra que é factível a comparação da configuração empírica com o resultado da simulação, ainda que em uma grande área, com período de tempo ampliado e sob influência de incrementos significativos de população.

Especificamente, a análise de sensibilidade demonstra a influência da vizinhança como importante elemento de explicação da dinâmica urbana para o caso analisado. Heterogeneidade espacial também mostra-se central uma vez que atores que entram no sistema ponderam as alternativas de forma não-linear. A configuração proposta resulta em uma ocupação do solo urbano que é complexa, e em sintonia com a cidade observada empiricamente.

Heterogeneidade social também se comprova relevante. A análise de sensibilidade demonstra que negligenciar a diversa dimensão social dos atores urbanos resultaria em configuração não compatível com a observada. Recomenda-se que futuros modelos de autômatos celulares enfatizem o nível intraurbano a ponto de incluirem heterogeneidades sociais, exceto em regiões altamente homogêneas ou nas quais a regulação governamental seja estritamente aplicada (como parece ser o caso da Holanda, por exemplo).

Este trabalho traz influências locais e bairros cognitivamente percebidos para o centro da discussão. Em segundo lugar, considera explicitamente heterogeneidades espaciais e sociais em um contexto de rápido crescimento. Ademais, formula um modelo adaptado de autômatos celulares que inclui efeitos negativos, testa e aplica uma nova metodologia (IVQR) em um estudo de caso. Busca com isso a melhoria da qualidade da transição entre economistas urbanos, sociólogos e geógrafos urbanos e planejadores urbanos no intuito de oferecer aos formuladores de políticas públicas, em específico, e à literatura, em geral, novos enfoques hermenêuticos. Adicionalmente, fornece descrição extensiva do estudo de caso em uma escala que ainda não tinha sido feita nem para Belo Horizonte, nem para sua região metropolitana.

Samenvatting in het Nederlands (Dutch summary)

Dit proefschrift richt zich op de integratie van twee wetenschapsvelden die zich bezig houden met stedelijke ontwikkeling, namelijk de stedelijke economie – welke van oorsprong vertrekt vanuit een monocentrisch ruimtelijk-economisch raamwerk – en dat van de zelforganiserende systemen en cellular automata modellen, hetgeen zich toespitst op het verklaren van muti-nodale en evolutionaire stedelijke ontwikkeling. In het onderzoek staan stedelijke vastgoedmarkten centraal, omdat deze markten de complexiteit en heterogeniteit van stedelijke ontwikkeling goed weerspiegelen. In de geïntroduceerde modellen betrekken we sociaal-economische heterogeniteit en wijkkenmerken, welke als zeer belangrijk worden beschouwd in de planologie en stedelijke geografie maar in beide onderzoekstradities geen expliciete plaats hebben. De centrale hypothese van het onderzoek is dat wijken als *identiteitsgemeenschappen* een belangrijke factor vormen bij het analyseren en verklaren van vastgoedprijzen en lokaalgebonden stedelijke ontwikkeling op de lange termijn. Dit wordt getest voor de casestudy van de Braziliaanse metropool Belo Horizonte. Een beter inzicht in deze processen zal beleidsmakers een aantal concrete en tegelijkertijd verdiepende inzichten verschaffen.

Steden zijn complexe sociaaleconomische verschijnselen, die zich fysiek over de ruimte verdelen en continu een (evolutionaire) ontwikkeling doormaken. Stedelijke vastgoedmarkten beïnvloeden en worden beïnvloed door de eigenschappen van een stad en de percepties van haar inwoners. Bovendien is een groot deel van het prijsniveau op stedelijke vastgoedmarkten afhankelijk van de locatie van het vastgoed in het stedelijk netwerk. Deze wederzijdse beïnvloeding tussen vorm en functie van vastgoed consumenten, maakt de vastgoedmarkt complex. Het is ook een economisch belangrijk element voor grootstedelijke regio's, want het stedelijke vastgoed vormt tussen de 45% en 75% van het kapitaal en uitgavenpatroon.

De dynamiek van deze stedelijke structuur kan het best omschreven worden als het resultaat van positieve en negatieve terugkoppelingen die de prijsniveaus in steden beïnvloeden en tevens door diezelfde prijsniveaus worden beïnvloed. Ondanks een omvangrijke literatuur over de facetten van stedelijke dynamiek in het algemeen en stedelijke vastgoedmarkten in het bijzonder, stellen we dat lokale effecten en wijkidentiteit in onvoldoende mate zijn meegenomen in eerder onderzoek.

De literatuur die zich bezig houdt met het dynamisch modelleren van stedelijke patronen richt zich niet zozeer op de verschillen in gedrag van actoren op basis van hun inkomensniveau, maar op residentiële bebouwing in het algemeen (in tegenstelling tot andere vormen van grondgebruik zoals bosbouw, landbouw of industriële bebouwing). Het is duidelijk dat dit laatste onderscheid niet volstaat, omdat in de literatuur is aangetoond dat het residentiële bezettingsgedrag verschilt tussen de diverse inkomensgroepen. In feite kan een analyse die zich op het intra-stedelijk niveau richt, de sociale verschillen die er binnen een stad bestaan en hun effect op de prijsstelling niet negeren.

De stedelijke economie en de complexiteits- en landgebruiksmodellen kunnen vooral in samenhang toegevoegde waarde leveren. Aan de ene kant – met betrekking tot de stedelijke economie op het binnenstedelijk schaalniveau – lijkt de afstand tot het zakencentrum of andere oriëntatiepunten *an sich* de grote ruimtelijke heterogeniteit die aanwezig is in de stad niet te kunnen vangen. Aan de andere kant nemen burgers ruimte cognitief waar en kennen zij er waarde aan toe. Dit suggereert dat wijken, waaraan deze waarde wordt toegekend, uitdrukkelijk begrepen, beschreven en ingebed dienen te worden in de analyse van stedelijke vastgoedprijzen. Sociale wetenschappen hebben al eerder beschreven waarom wijken, plaatsidentiteit en gesegregeerde ruimtelijke patronen er toe doen. Die discipline legt echter niet het verband met vastgoedprijzen.

Er is een aantal aanknopingspunten waaruit de complementariteit van de twee benaderingswijzen van stedelijke ontwikkeling blijkt. Het wezenlijke idee van negatieve (fysieke) terugkoppelingen van en naar de grondprijzen wordt in de stedelijke economie niet onderscheiden. Dit gebeurt echter wel impliciet in cellular automata modellen. Andersom vormen lokale effecten de kern van cellular automata modelleren, maar kunnen deze slechts in geringe mate teruggevonden worden in de stedelijke economie via homogene submarkten en de op afstand gebaseerde ruimtelijke analyse. En hoewel stedelijke economische theorieën vaak onvoldoende rekening houden met sociale heterogeniteit (aangezien actoren verschillende preferenties kunnen hebben), wordt dit niet van toepassing geacht op het binnenstedelijk niveau, wat als een homogene ruimte wordt beschouwd. Cellular automata modellen leggen de homogeniteit van actoren niet op, maar in empirisch werk worden vaak wel homogene 'residentiële' categorieën onderscheiden. Tenslotte richt de stedelijke economie zich op cross-sectionele evenwichten, terwijl complexe en zelforganiserende systemen zich voornamelijk bezighouden met de dynamiek van het veranderingsproces.

In dit proefschrift wordt gesteld dat om een beter begrip te krijgen van de complexe hedendaagse stedelijke omgeving en veranderingen in deze omgeving, stedelijke onderzoekers en beleidsmakers de stad niet louter moeten benaderen vanuit een stedelijk morfologische invalshoek, maar dat men tevens meer economische factoren zou moeten meenemen. Het gaat om een combinatie van inzichten vanuit de stedelijke economie enerzijds en de sociale actieruimte van actoren anderzijds. Om dit te bewerkstelligen, richten we ons op een betere inbedding van cognitief afgebakende wijken in verklarende economische modellen alsmede een betere modellering van (lokale) negatieve terugkoppelingen waarin sociale heterogeniteit wordt meegenomen. Derhalve richt de focus van dit proefschrift zich op (a) de rol van de wijk en haar identiteit, (b) verschillen in gedrag tussen residentiële actoren op basis van verschillen in sociale achtergrond en (c) de mate van heterogeniteit van stedelijk ruimtelijke kenmerken die gerelateerd kunnen worden aan de locatie van vastgoed. Dit laatste past in zowel een statische als dynamische visie op de stedelijke economische en ruimtelijke structuur.

Twee onderzoeksvragen worden derhalve in de studie beantwoord:

Verbeteren stedelijk vastgoedprijs modellen wanneer wijkidentiteit en sociale heterogeniteit worden meegenomen?

Kunnen cellular automata modellen meer bijdragen aan het begrip van stedelijke dynamische processen wanneer sociale heterogeniteit en negatieve terugkoppelingsmechanismes worden meegenomen?

Het eerste gedeelte van dit proefschrift bespreekt de belangrijkste theoretische benaderingen vanuit een stedelijk economisch perspectief en vanuit benaderingen die lokale interacties benadrukken. Aan de hand van deze discussie introduceren we een aangepast cellular automata model, welke op een dynamische manier omgaat met negatieve terugkoppelingen. Het tweede deel richt zich op toepassingen in Belo Horizonte, gebaseerd op de gevoerde theoretische discussies. Deze Braziliaanse metropool wordt gekenmerkt door een snel groeiende sociaal-economische ongelijkheid.

Geo-processing technieken en principale componenten analyse – welke vaak worden toegepast in de fysische geografie en mileuwetenschappen – worden gebruikt om een grote hoeveelheid sociaaleconomische, ruimtelijk gedetailleerde data te aggregeren naar het wijkniveau. Vervolgens wordt een ruimtelijk-econometrische analyse gepresenteerd. Deze heeft tot doel om de relatieve invloed van vastgoedkenmerken versus stadskenmerken op het prijsniveau nader te bestuderen. De resultaten leveren ook empirische indicatoren voor de parameters in het cellular automata model. Hierbij wordt een kwantiele regressie toegevoegd aan het model, omdat er aanwijzingen zijn dat de vastgoedmarkt voor dure en goedkope woningen verschillende patronen volgt. Een Instrumental Variable Quantile Estimation (IVQR) of Spatial Autoregressive Model (SAM) wordt toegepast op de dataset. Deze schattingstechniek heeft aangetoond robuust te zijn voor heteroskedasticiteit en uitbijters en presteert beter dan de General Moment Methods (GMM) of Maximum Likelihood (ML). In dit model worden kwantiele regressie en ruimtelijke analyse simultaan toegepast.

We introduceren de benodigde data voor de validatie van het cellular automata model. Deze komen vanuit de lokale volkstellingen, en we gebruiken die om een fysieke configuratie van inwoners in Belo Horizonte te creëren op basis van inkomen. Metriek betreffende het kwantificeren van kaarten wordt gepresenteerd, omdat dit ons kan helpen wanneer we een referentiekaart met de simulaties vergelijken. We passen het aangepaste cellular automata model toe op de metropool Belo Horizonte. Het model streeft ernaar de dynamiek van ruimtelijke interactie tussen actoren mee te nemen en hiermee een groter inzicht te verschaffen in de totstandkoming van prijzen. De resultaten van de toepassing worden gevalideerd tegen zowel de ruimtelijke configuratie van actoren als de geobserveerde vastgoedprijzen.

We concluderen dat het econometrische model goed functioneert en inzicht verschaft de factoren die bijdragen tot de ontwikkeling van vastgoedprijzen. Het is consistent en maakt causale afleidingen mogelijk. Het is beschrijvend in de zin dat, gegeven een nulalternatief, het model eenvoudig toegepast kan worden. Het model houdt echter geen rekening met factoren die zich dynamisch ontwikkelen in een specifieke stedelijke structuur. Het cellular automata model houdt hier juist wel rekening mee. Dit model wijst uit dat wijkkenmerken aangevuld met algemene informatie over bereikbaarheid een vergelijkbare empirische configuratie van actoren en prijzen in de stad genereert. In vergelijking met het econometrische model is de voorspelling van het cellular automata model echter minder nauwkeurig. De reden hiervoor is dat het cellular automata model zich op de algemene structuur richt en niet op de ontwikkeling van specifieke cellen (wijken). Om deze reden ontbreken in het cellular automata model factoren die wel gevangen worden door het econometrische model, zoals snelle, innovatieve ontwikkelingen en complexe netwerkeffecten (zie Hoofdstuk 9). Enkele van deze factoren kunnen wel toegevoegd worden in een meer gedetailleerde studie die zich uitstrekt over een kortere tijdsperiode.

We concluderen dat de gepresenteerde modellen het belang van lokale effecten en wijkkenmerken als centrale elementen in de complexe stedelijke vastgoedmarkten, aantoont. We tonen ook aan dat ruimtelijke en sociale heterogeniteit sterk ingebed zijn in de stedelijke structuur en de ontwikkeling van deze stedelijke structuur. Negatieve terugkoppelingen kunnen ook bijdragen aan het proces van stedelijke ontwikkeling en zouden verder uitgewerkt moeten worden in cellular automata modellering. De hypothese wordt bevestigd dat sociale en ruimtelijke heterogeniteit potentieel een grote invloed op de marktstructurering hebben.

De toegevoegde waarde van deze uitkomsten aan bestaande inzichten zitten in vijf elementen. Ten eerste plaatst dit proefschrift locale effecten en wijken in het middelpunt van de discussie. Ten tweede wordt in dit proefschrift uitdrukkelijk rekening gehouden met ruimtelijke en sociale heterogeniteit in een snel groeiende omgeving. Ten derde introduceert het een modelaanpassing om negatieve terugkoppelingen mee te nemen. Ten vierde wordt in het proefschrift een nieuwe methodologie getest en toegepast in de context van een casus (IVQR). Ten vijfde probeert het de verschillen tussen stedelijk economen, stedelijk sociologen, stedelijke planologen en stedelijke geografen te overbruggen. Hierbij biedt het aan stedelijk onderzoekers en beleidsmakers een meer substantieel hermeneutisch alternatief. Daarnaast wordt in het proefschrift een uitgebreide beschrijving gegeven van de metropool Belo Horizonte, waarbij moet worden opgemerkt dat zowel voor deze stad als haar omliggende gebieden dit nog nooit op deze schaal gebeurd is.

Curriculum Vitae

Bernardo Alves Furtado has a B.Sc. in Architecture and Urbanism from the Federal University of Minas Gerais (UFMG). In 2003, he finished his Master's degree in Geography with an emphasis on geoprocessing from Pontifical University of Minas Gerais (PUC-MG), and started working as an assistant professor. In 2005, he started his PhD in Regional Economics from CEDEPLAR at UFMG. In 2008, he did a sandwich-doctorate at the Research Institute of Knowledge Systems, in Maastricht, Netherlands. Immediately after that, he was accepted to the PhD Geosciences program at the Urban and Regional Research Center (URU) at the University of Utrecht (UU) where a co-tutorship agreement with UFMG was signed. He has been working as an assistant professor and course coordinator at Centro Universitário UNA since 2006. In March 2009, he was approved in first place for a position as Adjunct Professor at the Urbanism Department at UFMG. In May 2009, he was also approved for a researcher position to Institute of Applied Economic Research (IPEA), within the Brazilian National Secretary of Strategic Affairs. His main interests are urban development, spatial analysis, regional and urban economics, cellular automata models, spatial econometrics, territorial-public policy issues, and geoprocessing and multivariate analysis tools.

Endnotes

1 This was developed theoretically, empirically and analytically using software and case studies (Anselin, 1990; 1992a; Kelejian and Robinson, 1995; Anselin and Bera, 1998; Kelejian and Robinson, 1998; Anselin, 1999; Kelejian and Robinson, 1999; Anselin, 2003).

2 A good review of the modeling with mathematical details and derivations explained didactically can be found in Fujita (1989), especially part I.

3 For a recent review of succession and a case study in Brazil, see Furtado (2006)

4 Alonso details this in appendix D: *Urban rent as monopoly income*.

5 Subscripts are used to denote partial derivatives.

6 The complete notation $P_f(t)||G_o$ indicates that the price paid p by each firm f in any given location t is established such that for an optimal quantity of land, the level of profit G_o is constant. Therefore, the firm is indifferent to location.

7 Note the similarity to equation 5.

8 p is potential price and P is price actually paid.

9 This would be equivalent to equation 16, generalized for the last location t_n: $P_n = P_n(t_n)||t_{n+1}, P_{n+1}$

10 Discussion on the effects of different minimum and maximum lot size is in section D of chapter 6, pp. 117.

11 Wheaton, W. C., A comparative analysis of urban spatial structure. **Journal of Economic Theory**, v.9, p.223-237. 1974.

12 For Brazilian cities like, Belo Horizonte, in particular, Brueckner's starting point of "observed regularities" should be taken with care because, fast-growing, populous, developing-country cities are usually described as irregular and differentiated in a number of different parameters. Regularities might be observable only at greater scales, but not at the urban, neighborhood level (see items 2.2, and 2.3). Glaeser (2007) in a recent and didactic explanation of urban economics to non-economists reiterated that states that their approach is usually a general view that might not apply to specific cases. That is, the scale of analysis preferred by economists, according to Glaeser, is the larger scale.

13 Brueckner (1987, p. 825) reminds us that "$\partial q/\partial y$ bears no relation to the regular income effect since utility is held fixed". In this case, a decrease in x has the same effect of an increase in y.

14 $XD\ (x,t,y,u)\ dx = L$, where θ is the number of radians of land available for housing at every distance x (Brueckner, J., 1987, p. 829).

15 For an approach of agglomeration, scale diseconomies, and their importance to the success or decline of cities, see Glaeser (1998).

16 In this case, if t/q increases with income, then the accessibility effect is dominant and the chosen location is near the CBD. Similarly, if the ratio t/q decreases, preference for space is dominant and individuals with higher income locate in the suburbs.

17 See page 102, equation 11.

18 Wheaton says: "The facts simply do not fit a 'monocentric' model" (2004, p. 418).

19 See also Leven et al. (1977).

20 Bartik and Smith quote Diamond and Tolley's (1982) definition that defines "urban amenities [as] location-specific goods that are often non-excludable once access to their location is acquired" (1987, p. 1226).

21 A recent study (Santiago, 2006) based on census information revealed that vacancy rates for apartment buildings in Belo Horizonte average 13%.

22 Abramo (2001, p. 90) analyzes these aspects when he explains the Wingo model.

23 For a review on urbanism history, see Choay (1965).

24 See the social-urban characterization of this period in, for example, Oliver Twist by Charles Dickens (1949).

25 See Hobsbawm (1975; 1987).

26 For a formal demonstration of Schelling's model see Batty (2005, p. 52).

27 "The stochastic term *S* has a highly skewed distribution, so most values are near unity, and much larger values occur only infrequently" (White and Engelen, 1993, p. 1179).

28 The model is extended, detailed, and its internal mechanisms are described in the following chapter.

29 The estimation the authors calculate is standard in the real estate literature and includes the distance to points of interest and density of residences where the estate is located. In chapter 6, we present a different approach to both spatial characteristics and the importance of valuation of attributes according to different socioeconomic levels.

30 As an illustration, consider the case of Belo Horizonte. The city has a population of 2.4 million inhabitants (Table 4-3) with a metropolitan area of 4.4 million (Table 4-2). Yet, it has one single line of metro of about 16 kilometers. Most intra-metropolitan journeys are made by bus, which are a much more flexible form of transportation to implement, and does not necessarily require complex highway networks. Buses also have a smaller effect than a structural mode of transport. However, institutional bodies (such as Millennium Objectives Observatory – Belo Horizonte ODM-BH) are starting to work together in order to obtain fully integrated, comprehensive, spatial, and common databases and financial restraints that might diminish in the future, which will allow Brazil (and Belo Horizonte) to plan and implement transport systems before capacity is reached and overloaded.

31 Detailing the pace of evolution of Brazilian urban population and the case of the Greater Area of Belo Horizonte is the theme of chapter 4.

32 "[E]ven sprawl is far from homogeneous, and geographers have perceived patterns that conform to the mathematics of highly irregular structures such as fractals" (Anas, Arnott *et al.*, 1998, p. 1427).

33 First term on the right-hand side of equation 31.

34 Second term on the right-hand side of equation 31.

35 In the case of Belo Horizonte, for instance, although the majority of the early occupation of the lower class was removed from their original sites and *succeeded*, some areas have maintained their original presence, amidst upper class occupation for over 60 years (Furtado, B.A., 1995). For further information on neighborhood succession, see Grigsby (1987).

36 Adapted by the author, from METRONAMICA model description (Riks, 2007b).

37 Adapted by the author, from METRONAMICA model description (Riks, 2007b).

38 Licensed by RIKS BV, available at http://www.riks.nl, Maastricht, The Netherlands.

39 Other instances of rapid urban growth have been observed elsewhere (Brueckner, J., 1990).

40 See also Santos (1994).

41 For further details on the early network of Minas Gerais during the mining colony years, see Santos (2001).

42 Singer suggests a process of strong "alienation from economic life" (1977, p. 213).

43 Singer cites Professor Francisco Iglesias as the source for this information.

44 Adjective: natural of, or from the state of Minas Gerais.

45 In the American literature, this notion, detached from the social aspect, is discussed as development. The actors who implement it are known as developers.

46 The author refers to appropriation of space in a manner of socially taking possession of the space. Residents slowly develop a sense of belonging that is essential for their quest for improvements, which is part of the logic of the dynamics.
47 For a good compilation of Belo Horizonte's early spatial, economic history, see also Canuto (2005).
48 Law 6,303 of 30 April, 1974.
49 Superseded by State Law 12,153 of 21 May 1996.
50 And complementary Laws 88 and 89 in 2006.
51 Its composition is 5 from the State, 2 from the State Legislative Chamber, 2 from the City of Belo Horizonte, 1 from Contagem, 1 from Betim, 3 from remaining municipalities and 2 from organized civil society.
52 For data on the urbanization growth of other countries, see the Population Division of the Department of Economic and Social Affairs of the United Nations Secretariat, **World Population Prospects: The 2006 Revision and World Urbanization Prospects: The 2007 Revision**, available at http://esa.un.org/unup.
53 See also Cullen (1961) and Yi-Fu Tuan (1980; 1983) for similar approaches in the fields of micro-scale and human geography, respectively.
54 Lynch formally proposes five analytical elements which are: (path/high)ways, limits, marks, neighborhoods and nodes (1997).
55 Results of the neighborhood proposed by the foundation are also available for the cities of Salvador, Recife, Manaus and Rio de Janeiro, in Brazil.
56 This heterogeneity is easily assessed via an analysis of income, which is for example, made explicit in the results of the census. For a visual analysis of Belo Horizonte, see Figure 7-3.
57 See item (d) of the criteria used by Fundação João Pinheiro (p. 91): to "have a common and recognizable identity".
58 We would like to acknowledge the advice and the provision of the estimation program by Professor Zhenlin Yang, School of Economics Singapore Management University and Professor Liangjun Su, Guanghua School of Management, Peking University.
59 As discussed in section 4.3, in Belo Horizonte, Brazil, a large polygon can be considered the CBD.
60 It is not the objective of this paper to discuss the underlying theory. For further details, see Anselin (1988).
61 The spatial error model was originally proposed by Whittle (1954).
62 Other spatial models include average moving spatial models, spatial Durbin models, mixed models with spatial lag, and error moving averages of the first degree (Almeida, 2004).
63 According to Anselin, "due to the simultaneity implied by the spatial nature of the dependence, these procedures [estimated general least square] are not applicable in the spatial case and a full maximum likelihood estimation must be carried out" (1992b, p. 214).
64 For example, as in the case of a mosquito flying autonomy.
65 See Buchinsky (1997) and Koenker and Bassett (1978) for details.
66 $F(x) = P(X \leq x)$, where F(x) is the probability that the variable X presents value inferior or equal to x such that the probability of X is in an interval (a, b) is $F(b) - F(a)$ if $a \leq b$.
67 The authors originally used the notation λ. However, to avoid confusion with the parameter for the error it was replaced with ρ.
68 The dataset of real estate households was provided by the Finance Department of the City Administration of Belo Horizonte.
69 For individual houses, this value is always 1. For apartments or multifamily units, the value is proportional to the number of domiciles within the parcel. The value is 0 if there is no building on the parcel.
70 Detailed explanation of the factors is available in the legislative decree (Decreto 10925 of 2001 from City of Belo Horizonte).

71 The exchange rate was R$ 2.60 on March 7, 2008.

72 The k-nearest-neighbor, in which the *k* number of nearest neighbors are considered, is not evaluated because it lacks theoretical (Smirnov and Anselin, 2001; Anselin, 2002) and empirical intuition. If one thinks spatially, one's five nearest neighbors might have a totally different configuration and distance from one another and should not be considered part of the same pattern. Moreover, Anselin (2005) suggests that results have not proven to be consistent using this matrix and maximum likelihood methodology.

73 Because the variable *standard* is ordinal, the recommended procedure (Wooldridge, 2002, p. 214) would be to have dummy variables assigned to all categories but one. This would more accurately capture the differentiated impact from moving from a specific standard to another. However, for the sake of clarity and simplicity and because the standard plays the role of the control variable, only one dummy variable that represents an average influence is applied in the model. As a robustness check, an estimation with all variables was performed and little difference was observed in the parameters. The results are available on request. Finally, the quantile approach implicitly considers differentiation in quality of the estate and, by the construction of the standard variable criteria by the government, it is not possible to have an estate of standard 1 that, at the same time, is embedded in the highest quantile of price. The results would reflect this, which would make it more difficult to interpret locational variables (which are the focus of our study).

74 (c), the quality of public service offered within, and (e), proximity (such as accessibility to shopping or employment centers), as described in chapter 5.

75 All variables, except proximity to slums (see Table 6-11).

76 See examples in Lillesand et. al. (2004).

77 In this study, the following neighborhoods areas serve as reference for high-income classes: Lourdes, Sion, Mangabeiras, Belvedere, Cidade Jardim, Funcionários, some condominiums in Nova Lima, Cidade Nova, São Luís and Bandeirantes in Pampulha. For average-income actors, the references were parts of Caiçara, Padre Eustáquio, Sagrada Família, Carlos Prates and Jaraguá. Finally for the low-income classes, parts of the municipality of Vespasiano, in its conurbated part with the north of Belo Horizonte, portions of the municipality of Ribeirão das Neves (Justinópolis) and parts of the three main slum agglomerations of the city (Morro das Pedras, Serra/Cafezal and Santa Lúcia/Papagaio) were used.

78 The classes of income available are: (1) families whose income are (1) above 20 times the minimum-wage salary, (2) between 15 and 20 times the minimum-wage salary; (3) between 10 and 15 times the minimum-wage salary; (4) between five and ten times the minimum-wage salary; (5) between three and five times the minimum-wage salary; (6) between two and three times the minimum-wage salary; (7) between one and two times the minimum-wage salary; (8) between one-half a minimum-wage salary and one; (9) up to half of one minimum-wage salary; and (10) no income

79 The referred municipalities are (1) Betim, (2) Contagem, (3) Ibirité, (4) Nova Lima, (5) Ribeirão das Neves, (6) Sabará, (7) Santa Luzia and (8) Vespasiano.

80 The clustering procedure is applied only to domiciles. The other data, which is used in the CA model, is shown only to facilitate reading maps and comparing results in chapter 8.

81 The city center was chosen *ad hoc* at the cell at x:325, y:361 which is in the central area of the central ring. As discussed before, the choice of where exactly the city center is in an intra-urban analysis is always discretionary.

82 "High-income" refers to domiciles that earn more than 20 times the minimum-wage salary. "Average-income" represents families with salaries ranging from three to 20 times the minimum-wage salary. "Low-income" refers to families below three times the minimum-wage salary.

83 Available at www.ipead.face.ufmg.br.

84 Vacant is the baseline function when implementing the model using the software GEONAMICA. It plays no role in the results and is always what is left after all demand for all functions has been fulfilled in a given year.
85 Capital of Minas Gerais state, in Brazil. See **Figure** 4-1 for a contextualization map.
86 Represented as an asterisk (*) in Table 8-6.
87 There are only two major parks in the city of Belo Horizonte. One is centrally located within a heavily commercial service area and is considered popular (in a negative sense). It is close to a polluted, fetid river. The other has exclusive access, with an entrance at the end of a high-income area. Other green areas, such as Parque Lagoa do Nado, are missing infrastructure or maintenance.
88 Most intra-urban transportation consists of buses or private vehicles (Gouvêa, 2005). Large freight mining companies, such as Vale, use their private rail.
89 Because of size, see the legend in Figure 8-5.
90 This information is exogenously supplied to the model, which should capture the locational choice of newcomers rather than just their aggregate numbers.
91 The results of Table 7-7 are reproduced on the left side of Table 8-9 for easier comparison.
92 Price data for comparison basis is only available for the city of Belo Horizonte (Figure 7-7). The simulation, however, produces results for the entire study area, which is composed of the nine municipalities.
93 An author basing his findings solely on observed experience would suggest that steeper locations would more favorable for this kind of occupation. Detailed hypsometric (10 meters distance) information for the study area was not available for this study.
94 The parameters in Table 8-6 are set to have similar numbers for the interaction among residential actors and others, rather than different numbers. This would be equivalent to saying that there is only one residential income-level actor, as is in studies that apply urban x non-urban classifications.
95 The *Simulation* value is drawn from the land price map (Figure 8-6) at the specific locations of the observations (Figure 6-1).
96 Implicitly, it would incorporate parts of Tiebout's (1956) seminal work in the proposed model.
97 In their experiment, agents optimize their consumption according to resources available using adaptive trade rules as the environment changes.

NEDERLANDSE GEOGRAFISCHE STUDIES / NETHERLANDS GEOGRAPHICAL STUDIES

350 Z FÖLDI Neighbourhood dynamics in Inner-Budapest; A realist approach -- Utrecht 2006: Knag/Faculteit Geowetenschappen Universiteit Utrecht. 345 pp, 72 figs, 10 tabs, 34 pics. ISBN: 978-90-6809-391-9, Euro 35,00
351 S AGTERBOSCH Empowering wind power; On social and institutional conditions affecting the performance of entrepreneurs in the wind power supply market in the Netherlands -- Utrecht 2006: Knag/Copernicus Institute. 246 pp, 36 figs, 8 tabs. ISBN: 978-90-6809-392-6, Euro 27,50
352 C RINZIN On the middle path; The social basis for sustainable development in Bhutan -- Utrecht 2006: Knag/Copernicus Institute. 204 pp, 18 figs, 37 tabs, 5 box. ISBN: 90-6809-393-2, Euro 22,50
353 M VAN WIJK Airports as cityports in the city-region; Spatial-economic and institutional positions and institutional learning in Randstad-Schiphol (AMS), Frankfurt Rhein-Main (FRA), Tokyo Haneda (HND) and Narita (NRT) --Utrecht 2007: Knag/Faculteit Geowetenschappen Universiteit Utrecht. 323 pp, 59 figs, 18 tabs. ISBN: 978-90-6809-394-0, Euro 35,00
354 A C HELDERMAN Continuities in homeownership and residential relocations -- Utrecht/Amsterdam 2007: Knag/ Faculteit der Maatschappij- en Gedragswetenschappen Universiteit van Amsterdam. 125 pp, 15 figs, 14 tabs. ISBN: 978-90-6809-395-7, Euro 19,00
355 W BORREN Carbon exchange in Western Siberian watershed mires and implication for the greenhouse effect; A spatial temporal modeling approach -- Utrecht 2007: Knag/Faculteit Geowetenschappen Universiteit Utrecht. 125 pp, 36 figs, 17 tabs. ISBN: 978-90-6809-396-4, Euro 13,00
356 S O NEGRO Dynamics of technological innovation systems; The case of biomass energy -- Utrecht 2007: Knag/Copernicus Institute. 166 pp, 24 figs, 17 tabs. ISBN: 978-90-6809-397-1, Euro 18,00
357 R NAHUIS The politics of innovation in public transport; Issues, settings and displacements -- Utrecht 2007: Knag/Copernicus Institute. 184 pp, 9 figs, 40 tabs, 4 box. ISBN 978-90-6809-398-8, Euro 20,00
358 M STRAATSMA Hydrodynamic roughnesses of floodplain vegetation; Airborne parameterization and field validation -- Utrecht 2007: Knag/Faculteit Geowetenschappen Universiteit Utrecht. 180 pp, 55 figs, 28 tabs. ISBN: 978-90-6809-399-5, Euro 23,00
359 H KRUIZE On environmental equity; Exploring the distribution of environment quality among socio-economic categories in the Netherlands -- Utrecht 2007: Knag/Copernicus Institute. 219 pp, 76 figs, 25 tabs. ISBN: 978-90-6809-401-5, Euro 25,00
360 T VAN DER VALK Technology dynamics, network dynamics and partnering; The case of Dutch dedicated life sciences firms -- Utrecht 2007: Knag/Copernicus Institute. 143 pp, 23 figs, 25 tabs. ISBN: 978-90-6809-402-2, Euro 15,00
361 M A SCHOUTEN Patterns in biodiversity; Spatial organisation of biodiversity in the Netherlands -- Utrecht 2007: Knag/Copernicus Institute. 152 pp, 16 figs, 20 tabs. ISBN: 978-90-6809-403-9, Euro 20,00
362 M H J W VAN AMSTEL – VAN SAANE Twilight on self-regulation; A socio-legal evaluation of conservation and sustainable use of agrobiodiversity by industry self-regulation -- Utrecht 2007: Knag/Copernicus Institute. 167 pp, 15 figs, 13 tabs, 5 box. ISBN: 978-90-6809-404-6, Euro 20,00
363 S MUHAMMAD Future urbanization patterns in the Netherlands, under the influence of information and communication technologies -- Utrecht 2007: Knag/Faculteit Geowetenschappen Universiteit Utrecht. 187 pp, 82 figs, 20 tabs. ISBN: 978-90-6809-405-3, Euro 20,00
364 M GOUW Alluvial architecture of the Holocene Rhine-Meuse delta (The Netherlands) and the Lower Mississippi Valley (U.S.A.) -- Utrecht 2007: Knag/Faculteit Geowetenschappen Universiteit Utrecht. 192 pp, 55 figs, 14 tabs. ISBN: 978-90-6809-406-0**,** Euro 22,00
365 E HEERE & M STORMS Ormeling's cartography; Presented to Ferjan Ormeling on the occasion of his 65th birthday and his retirement as Professor of Cartography -- Utrecht 2007: Knag/Faculteit Geowetenschappen Universiteit Utrecht. ISBN: 978-90-6809-407-7, Euro 20,00
366 S QUARTEL Beachwatch; The effect of daily morphodynamics on seasonal beach evolution -- Utrecht 2007: Knag/Faculteit Geowetenschappen Universiteit Utrecht. 125 pp, 39 figs, 7 tabs. ISBN: 978-90-6809-408-4, Euro 12,50
367 R O VAN MERKERK Intervening in emerging nanotechnologies; A CTA of Lab-on-a-chip technology regulation -- Utrecht 2007: Knag/Copernicus Institute. 206 pp, 19 box, 35 figs, 12 tabs. ISBN: 978-90-6809-409-1, Euro 20,00
368 R M FRINGS From gravel to sand; Downstream fining of bed sediments in the lower river Rhine -- Utrecht 2007: Knag/Faculteit Geowetenschappen Universiteit Utrecht. ISBN: 978-90-6809-410-7**,** Euro 25,00
369 W IMMERZEEL Spatial modelling of the hydrological cycle, climate change and agriculture in mountainous basins -- Utrecht 2008: Knag/Faculteit Geowetenschappen Universiteit Utrecht. 147 pp, 54 figs, 12 tabs. ISBN: 978-90-6809-411-4, Euro 25,00

370 D S J MOURAD Patterns of nutrient transfer in lowland catchments; A case study from northeastern Europe -- Utrecht 2008: Knag/Faculteit Geowetenschappen Universiteit Utrecht. 176 pp, 44 figs, 19tabs. ISBN: 978-90-6809-412-1, Euro 20,00
371 M M H CHAPPIN Opening the black box of environmental innovation; Governmental policy and learning in the Dutch paper and board industry -- Utrecht 2008: Knag/Copernicus Institute. 202 pp, 41 figs, 30 tabs. ISBN: 978-90-6809-413-8, Euro 22,50
372 R P ODDENS & M VAN EGMOND Ormelings atlassen; Catalogus van atlassen geschonken aan de Universiteit Utrecht door de hoogleraren F.J. Ormeling sr. en jr. – Utrecht 2008: Knag/Faculteit Geowetenschappen Universiteit Utrechts. ISBN: 978-90-6809-415-2, Euro 15,00
373 R VAN MELIK Changing public space; The recent redevelopment of Dutch city squares – Utrecht 2008: Knag/Faculteit Geowetenschappen Universiteit Utrecht. 323 pp, 47 figs, 32 tabs. ISBN: 978-90-6809-416-9, Euro 20,00
374 E ANDRIESSE Institutions and regional development in Southeast Asia; A comparative analysis of Satun (Thailand) and Perlis (Malaysia) -- Utrecht 2008: Knag/Faculteit Geowetenschappen Universiteit Utrecht. 250 pp, 42 figs, 43 tabs, 18 box. ISBN: 978-90-6809-417-6, Euro 25,00
375 E HEERE GIS voor historisch landschapsonderzoek; Opzet en gebruik van een historisch GIS voor prekadastrale kaarten -- Utrecht 2008: Knag/Faculteit Geowetenschappen Universiteit Utrecht. 231 pp, 73 figs, 13 tabs. ISBN: 978-90-6809-418-3, Euro 30,00
376 V D MAMADOUH, S M DE JONG, J F C M THISSEN & J A VAN DER SCHEE Dutch windows on the Mediterranean; Dutch Geography 2004-2008 -- Utrecht 2008: Knag/International Geographical Union Section The Netherlands. 104 pp + cd-rom, 38 figs, 9 tabs. ISBN: 978-90-6809-419-0, Euro 10,00
377 E VAN MARISSING Buurten bij beleidsmakers; Stedelijke beleidsprocessen, bewonersparticipatie en sociale cohesie in vroeg-naoorlogse stadswijken in Nederland -- Utrecht 2008: Knag/Faculteit Geowetenschappen Universiteit Utrecht. 230 pp, 6 figs, 31 tabs. ISBN: 978-90-6809-420-6, Euro 25,00
378 M DE BEER, R C L BUITING, D J VAN DRUNEN & A J T GOORTS (Eds.) Water Wegen; Op zoek naar de balans in de ruimtelijke ordening -- Utrecht 2008: Knag/VUGS/Faculteit Geowetenschappen Universiteit Utrecht. 91 pp, 18 figs, 1 tab. ISBN: 978-90-6809-421-3, Euro 10,00
379 J M SCHUURMANS Hydrological now- and forecasting; Integration of operationally available remotely sensed and forecasted hydrometeorological variables into distributed hydrological models -- Utrecht 2008: Knag/Faculteit Geowetenschappen Universiteit Utrecht. 154 pp, 65 figs, 12 tabs. ISBN 978-90-6809-422-0, Euro 15,00
380 M VAN DEN BROECKE Ortelius' *Theatrum Orbis Terrarum* (1570-1641); Characteristics and development of a sample of *on verso* map texts -- Utrecht 2009: Knag/Faculteit Geowetenschappen Universiteit Utrecht. 304 pp + cd-rom, 9 figs, 65 tabs. ISBN 978-90-6809-423-7, Euro 30,00
381 J VAN DER KWAST Quantification of top soil moisture patterns; Evaluation of field methods, process-based modelling, remote sensing and an integrated approach -- Utrecht 2009: Knag/Faculteit Geowetenschappen Universiteit Utrecht. 313 pp, 108 figs, 47 tabs. ISBN 978-90-6809-424-4, Euro 30,00
382 T J ZANEN 'Actie, actie, actie…'; De vakbeweging in regio Noord-Nederland, 1960-1992 -- Utrecht/ Groningen 2009: Knag/Faculteit Ruimtelijke Wetenschappen Rijksuniversiteit Groningen. ISBN 978-90-6809-425-1, Euro 30,00
383 M PERMENTIER Reputation, neighbourhoods and behaviour -- Utrecht 2009: Knag/Faculteit Geowetenschappen Universiteit Utrecht. 146 pp, 10 figs, 19 tabs. ISBN 978-90-6809-426-8, Euro 15,00
384 A VISSER Trends in groundwater quality in relation to groundwater age -- Utrecht 2009: Knag/Faculteit Geowetenschappen Universiteit Utrecht. 188 pp, 47 figs, 24 tabs. ISBN 978-90-6809-427-5, Euro, 20,00
385 B A FURTADO Modeling social heterogeneity, neighborhoods and local influences on urban real estate prices; Spatial dynamic analyses in the Belo Horizonte metropolitan area, Brazil -- Utrecht 2009: Knag/Faculteit Geowetenschappen Universiteit Utrecht. ISBN 978-90-6809-428-2, Euro 25,00

For a complete list of NGS titles please visit www.knag.nl. Publications of this series can be ordered from KNAG / NETHERLANDS GEOGRAPHICAL STUDIES, P.O. Box 80123, 3508 TC Utrecht, The Netherlands (E-mail info@knag.nl; Fax +31 30 253 5523). Prices include packing and postage by surface mail. Orders should be prepaid, with cheques made payable to "Netherlands Geographical Studies". Please ensure that all banking charges are prepaid. Alternatively, American Express, Eurocard, Access, MasterCard, BankAmericard and Visa credit cards are accepted (please specify card number, name as on card, and expiration date with your signed order).