PhD- FLSHASE-2014-14

The Faculty of Language and Literature, Humanities, Arts and Education

DISSERTATION

Defense held on 23/09/2014 in Walferdange

to obtain the degree of

DOCTEUR DE L'UNIVERSITÉ DU LUXEMBOURG

EN GÉOGRAPHIE

by

Marie-Line GLAESENER

Born on 29 August 1983 in Luxembourg, (Luxembourg)

GEOGRAPHICAL DETERMINANTS OF RESIDENTIAL LAND VALUES IN LUXEMBOURG

Dissertation defense committee

Dr Geoffrey Caruso, dissertation supervisor
Associate Professor, Université du Luxembourg

Dr Dominique Peeters
Professor, CORE, Université catholique de Louvain

Dr Markus Hesse, Chairman
Professor, Université du Luxembourg

Dr Mohamed Hilal
Senior Researcher, CESAER, UMR 1041, INRA, Dijon

Dr Luisito Bertinelli, Vice Chairman
Associate Professor, Université du Luxembourg

Contents

Introduction

The spatial expansion of cities is, among others, a consequence of the tremendous population growth observed mainly since World War Two. Since the 1950s, the European population increased by 33%, while urban areas have increased their size by about 78% (Nilsson et al., 2013, p.2). Artificial land cover was found to have increased by 3.4% (European Environment Agency, 2010) from 2000 to 2006, and it is predicted that in the next decades this land conversion continues to consume up to 0.7% of undeveloped land per year in Europe (Nilsson et al., 2013, p.406). This population growth is not distributed homogeneously through space but it is mostly orientated to the urban areas, which in consequence are expanding. Between 70% and 82% of the total population in Europe were living in urban areas in 2010 (UN, 2011).

Clear deconcentration trends from urban to rural areas have thus been observed in Europe, resulting in the irreversible conversion of "natural" land. Population growth is considered a main driver of urban growth, together with economic growth and decreasing transport costs, they are the fundamental causes of urban spatial expansion (Brueckner, 2001). These fundamental drivers are derived from the monocentric city model (Alonso, 1964), which explains the allocation of land-uses and the size of urban area as a result of market processes (Anas et al., 1998) and relates them to a trade-off made by individuals (between job accessibility and land consumption), aiming at maximising their utility under a budget constraint. Within the monocentric framework, urban spatial expansion is thus considered as the outcome of some fundamental changes at the global level and individual location decisions at the local level.

In this framework emerged the periurban concept which adds further complexity to the trade-off made by individuals when deciding to settle. The periurban area is defined as a morphologically and functionally distinct area between the urban core area, characterised by high population and residential land-use density, and the rural area defined by its agricultural activities. It is considered as a mixed urban and rural area, where individuals obtain utility from both, urban and rural amenities while maintaining strong functional links with the core urban area (Cavailhès et al., 2004; Caruso, 2005). This concept highlights thus the utility bearing characteristics of the periurban area, that emerge as a consequence of additional trade-offs made by land consumers towards local urban and rural amenities, besides job accessibility and land consumption.

Although spatial expansion is needed with regard to increasing population, it might be excessive and induce costs related to a non optimal allocation of land-uses (Brueckner and Fansler, 1983). Several market failures that might disturb the land-use equilibrium and result in excessive urban growth and socially undesired effects have been identified by Brueckner (2001). The land-use patterns generally associated with the periurban area are low-density settlements, discontinuous to the urban core area and emerging leapfrogging urban development said to waste and miss-allocate productive farmland (Mills, 1981). The negative consequences are in general summarised as the fragmentation of undeveloped land, the degradation of natural resources, the elimination of functional open-spaces, increased public service costs, and traffic congestion and pollution (Irwin and Bockstael, 2007). These processes and the resulting patterns are regarded as a non-sustainable way of urban development in urban planning practice today.

In this perspective, spatial planning measures developed since the 1990s, aim particularly at promoting a smarter growth of cities (Smarth Growth Network, 2014), relying mainly on the claim for a return to compacter urban forms. In the last decades, however critiques towards these measures have emerged (Neuman, 2005; Echenique et al., 2012), claiming that the "return" to the compact city is not necessarily yielding more liveable urban areas. Although these measures might tackle some of the global problems related to "excessive" urban growth (Brueckner, 2001)[1], they do not specifically account for quality-of-live at the local level. Policies fostering compacter urban development were even related to increasing housing prices, to fostering urban spatial expansion (Irwin and Bockstael, 2004; Glaeser, 2008), to increasing individuals' exposure to pollution and

[1]Who defines sprawl as "*spatial growth of cities that is excessive relative to what is socially desirable*" (Brueckner, 2011, p.69).

congestion (Echenique et al., 2012; Schindler and Caruso, 2014) and as ignoring individuals' preferences for local green amenities and low-density urban development (Burchell et al., 2000; Turner, 2005)[2].

Although the same fundamental drivers apply, regional variations within Europe have been observed in the patterns of urban spatial expansion (Caruso, 2002; Siedentop and Fina, 2012), generally associated with the specific historic, social and geographical context. In the case of Luxembourg, the fundamental drivers have encouraged urban expansion over the last decades and considerably challenged planning policies. With regard to the monocentric organisation of the Grand Duchy, the country's size and the specificities observed in the land market, there is a pressing need for further understanding the processes underling urban spatial expansion.

To understand land-use allocation, understanding how land prices are determined in a competitive economy is primordial (Fujita and Thisse, 2013). Urban economic theory suggests that individuals make rational choices that are not random in space (Glaeser, 2008). The land market and more specifically land prices capitalise spatial variations related to the local context, and are hence a means to further understand the drivers of periurbanisation. If consumers buy land, its price translates the consumers' willingness to pay for this particular piece of land. The hedonic pricing method considers land as a composite good (Rosen, 1974), whose price is the result of the different trade-offs made by consumers according to their preferences and budget. A better understanding of land prices is thus essential to develop more targeted planning policies.

The analysis of the local determinants and spatial variations of land prices requires to quantify the basic trade-off and local amenities that are assumed to impact on the purchase decision of land consumers. As was highlighted, the local scale is of particular importance, it is necessary to understand the benefits and disadvantages of different locations, to understand why and where land conversion takes place. The hedonic modelling approach is complementary to the analysis of the main drivers of urban expansion

[2] A more detailed discussion on the theoretical and conceptual underpinnings of this thesis and the related discussion is provided in appendix A.

or the observed land-conversion as it aims at better understanding the individual scale. Technically, the estimation of the consumers' willingness to pay for the characteristics composing the land price, requires to rely on advanced spatial econometrics estimation techniques.

Urban economic theory provides theoretical guidance on the determinants of urban spatial expansion and how its patterns are driven by individual decisions. The thesis' main objective is to identify the geographical determinants of developable land prices in Luxembourg.
Urban spatial expansion entails major challenges to spatial planning policies, we aim at highlighting the importance of considering the local scale and individuals' preferences to design policies allowing at the same time sustainable and liveable urban development.
Particular attention is turned to the local context and in particular green land-uses and urban amenities, quantified via geographical information systems. These geographical tools are combined to state-of-the-art spatial econometric techniques to account for spatial effects.
Results tend to confirm the critiques towards compaction policies and show that land consumers in Luxembourg value local periurban amenities and that their preferences vary with their socio-economic background and through space.

0.1 Luxembourgish context and challenges

The growth of Luxembourg-city as major employment and financial centre was sustained since the second half of the 20^{th} century. Zahlen (2012c) details the long term evolution of Luxembourg's economic situation, linking the different waves of immigration to the different stages of economic specialisation[3]. More recently, the general economic crisis (2008) was followed by a period of recession in Luxembourg. Nevertheless, the country

[3]Since the 1950s it was mainly determined by a need for labour force in the emerging steel industry (1950-1975), slowed down by the oil shocks and decreasing GDP (1976-1984). The second part of the 1980s is characterised by a strong increase in the GDP related to the boom of the financial and service sector (1985-2007) accompanied by an important population increase (Zahlen, 2012b,c).

continued to register an exceptional population increase, with an annual average of 1.9% between 2008 and 2011, mainly based on migration (Zahlen, 2012c). Economic growth is mainly concentrated in Luxembourg-city, that can be considered the main employment centre at national scale and beyond, with about 60% of the domestic jobs and being the destination of almost 50% of the internal commutes (Diederich, 2011b).

The prosperous economic condition was accompanied by a sustained demographic growth since the 1960s. Between 1960 and 2011, the countries population has increased from 313,050 to 524,853 which corresponds to an average annual increase of 1% (Zahlen, 2012b)[4]. This trend was accelerated in the last decade, with an average annual growth of 1.5% (global growth of 16.6% between 2001 and 2011) (Heinz et al., 2012b), not expected to relax since an increase of 45% of resident population between 2011 and 2060 is predicted (EUROSTAT, 2011). The dominant position of Luxembourg-city as population centre at national scale is strengthened in the same period, registering an increase of population of 24% (Heinz et al., 2012a). Since the 1960s, population growth is mainly determined by migration, the share of non-Luxembourgish residents has increased from 13% to 43% in 2011 (Zahlen, 2012a). Strongest population densities are mainly observed in the southern part of the country. However, more sustained relative population growth has been observed in the municipalities in the north and west and Luxembourg-city (Heinz et al., 2012a, p.4), and further analysed and discussed in Bousch and Decoville (2012) and in part I.

Hesse (2014b, p.4) highlights the specificities of urban and economic development in Luxembourg, namely the short period in which this growth took place, the small size of the country and the persistent internationalisation (migrant workers, European institutions and the banking sector). The rather small capital has specialised in functions usually found in major metropolitan areas[5], in particular with regard to the financial industries. Furthermore, its economic growth is nowadays widely based on a large share of daily transnational commuters, working in Luxembourg-city.

[4]Compared to UE-27, where the annual rate of growth was only of 0.43%.

[5]Although Luxembourg-city could not necessarily be considered a metropolis, it is subject to "metropolisation" (Hesse, 2014b, p.5).

Challenges for spatial planning and context related specificities

These evolutions comprise some substantial challenges for spatial planning in Luxembourg, at the national and local scale. Recent reviews on these challenges and the different policies, measures and instruments installed in the last decades can be found in Becker and Hesse (2011); Chilla and Schulz (2011); Hesse (2014b); Affolderbach and Carr (2014). The challenges faced by policy makers are particularly related to the sustained transformation of undeveloped land, the general high real estate prices and the consequences of increased traffic (congestion and pollution).

The part of the country covered by urban land has almost doubled (+ 43%) in the last 20 years, while in the same period, a population increase of 29% was registered (Chilla and Schulz, 2012), confirming the general trends observed in Europe described above (Caruso, 2002; Siedentop and Fina, 2012; Nilsson et al., 2013). Simultaneously the continuous increase in housing and residential land prices required measures towards more affordable housing. According to OECD (2007, p.90) prices of residential land have increased by 6.3% between 1981 and 2001, while at the same time housing prices and construction costs only increased by 2.6% and 0.3% respectively. Generally associated to the limited supply of available housing and residential land, the question of a speculative bubble on the different segments of the real estate market has been repeatedly raised (Blot, 2006; Licheron, 2013), further discussed in section 1.2. An overview and critique on these regulation measures has been recently presented by the Conseil Economique et Social (2013). They explicitly request further insights into price formation on the housing and land market and the determinants of demand and supply for housing (Conseil Economique et Social, 2013).

The Luxembourg-city centred organisation entails some major difficulties with regard to traffic congestion and pollution, fostered by Luxembourg's central position as employment centre of the "Greater Region". Traffic and congestion are aggravated by the international daily commutes to Luxembourg-city, that sum up to almost 150,000 per day (Diop, 2011b, p.3)[6]. The trans-border traffic related problems have been re-

[6] For recent studies on this topic we refer to Diop (2011a,b); Carpentier (2010).

cently accentuated by an increasing number of former residents of Luxembourg moving across the border, as a response to the high housing prices, entailing major impacts on the neighbouring countries' property market and land consumption. In this thesis, the international and "Greater Region" context is not further considered, mainly due to the unavailability and incompatibility of data.

The blending of the different political levels, between the local and national scale, and the "*unusually close and largely horizontal power distances*" (Affolderbach and Carr, 2014, p.11) are a further specificity of the Luxembourgish context. The role of private planning companies elaborating local development plans and their relation to, and the role of, real estate developers is further questioned by Hesse (2014b).

Spatial planning policies

Recent national planning policies in Luxembourg aim at promoting a "decentralised concentration" of the urban areas. These policies intend to compact urban structure in predefined regional urban centres, by redistributing economic activities to attain a more polycentric organisation. The different laws and instruments installed since 1999[7], established the "Programme directeur de l'aménagement du territoire" (*PDAT*) (MIAT, 2003) as guidance framework, introducing instruments at the national and local planning scale (Chilla and Schulz, 2014). The overall scope is a dynamic and harmonious regional development reflecting economic, social and spatial cohesion while reducing the consumption of natural resources, energy and undeveloped land (Diederich, 2011a). To reach these goals, policies aim at precisely influencing the residential location decisions of individual households (MIAT, 2003, p.98) and thus changing consumers' preferences.

The measures presented in this perspective aim at defining an urban development policy focussing on denser and renewed urban centres, a densification of the existing urban structure. This should be reached by fiscal measures and assistance to encourage the use of yet undeveloped land in urban areas and the renovation of existing housing stock. The definition of zoning tools at municipal scale aims at slowing down periurbanisation and protecting undeveloped land. Land-use allocation is generally organised at

[7] Mainly the Loi du 21 mai 1999 concernant l'aménagement du territoire (1999).

municipal scale via the "Plan d'Aménagement Général" (*PAG*), presented by Van Rijswijck and Wagner (2011). This zoning tool allows to regulate and control land-use at the local municipal scale. Although municipalities are quite autonomous with regard to fixing the boundaries and the functionalities of the different zones, the general objectives and limitations fixed in the national framework need to be respected.

The second fundamental planning tool is the "Integrative strategy for mobility and spatial development" (*IVL*) presented in 2004 (MIAT, 2004). Its overall aim is to foster a "decentralised concentration" of residential land-uses and economic activities. However, the *IVL* has a rather conceptual character without legally binding planning tools (Becker and Hesse, 2011). A monitoring of the *IVL* measures (CEPS/INSTEAD, 2008) recently revealed that the initial short term goals have not been reached. Mainly because the population and economic growth predictions had not foreseen the tremendous population growth experienced in the last decade (Hesse, 2014b). The monocentric position of Luxembourg-city has thus even been further strengthened and periurbanisation has been progressing in the last years. Further, this integrative approach has been recently criticised for its main focus being on economic growth rather than sustainable urban development (Hesse, 2014b). He points out the conflict between the national and the local scale, arguing that although "decentralised concentration" might be reasonable at the national/regional scale, the densification objectives promoted at the local scale are more problematic (e.g.: large scale urban projects). The main challenge of spatial planning in Luxembourg are nowadays to account for these local specificities and to develop "*strategies of adaptation and compensation, rather than to deploy "integrated" visions in an immanently disintegrated environment*" (Hesse, 2014b, p.11).

The periurban area is particularly targeted by these measures. The positive effects on individuals' quality-of-life related to larger housing and proximity to green amenities are secondary. The second "National Plan for Sustainable Development" (*PNDD*) (MDDI, 2010) fixed recently some basic principles for future urban development in Luxembourg, focussing on the preservation of high quality-of-life, respecting ecological, social and cultural diversity (Kohnen, 2011). It suggests, among others, to limit land consumption to one hectare per day till 2020 (MDDI, 2010). Although, the *PNDD* acknowledges and shows awareness of the quality-of-life bearing benefits carried by local green and urban amenities, these guidelines are criticised by the Conseil Economique et Social (2013) since they are expected to slowdown the production of new housing and hence rather foster high real estate prices.

The observations made in the European context, and in the Luxembourgish framework in particular, emphasise the importance of further understanding the local context and individuals' preferences to develop sustainable and acceptable spatial planning policies. The question arises if the failure of actual planning policies, to slowdown urban expansion and to promote more compact urban forms, is not related to the ignorance of what actually is improving consumers' quality-of-life and if these policies should not rather be re-framed in this regard. Urban economists have developed theoretical models that substantially contribute to understanding urban growth and periurbanisation, some of which will be presented in the next section.

0.2 Theoretical and conceptual framework

The most influential urban model is without doubt the monocentric city model developed by Alonso (1964). In general, it shows how urban spatial extension is triggered by urban growth, governed by some fundamental drivers: population increase, economic growth and/or decreased transportation costs (Brueckner, 2011). Further, it illustrates how consumers trade-off between job accessibility and land consumption to maximise the utility they obtain under a given budget constraint. Where individuals eventually settle is a complex process, as it results from their decisions, that are not entirely irrational, nor random in space (Glaeser, 2008). Urban form and land-use patterns are thus considered as the result of decisions taken by individuals. Understanding these choices, that are determined by the benefits and disadvantages consumers perceive from the characteristics of the available land and its location in space, is crucial to develop appropriate policies, that are able to cope with the negative consequences of periurbanisation at the global level while accounting for the needs and quality-of-life at the individual level. In this way, urban spatial expansion can be considered as the outcome of some fundamental changes at global level and the result of individual location decisions at the local scale. A main objective of urban economics is to explain the allocation of land-uses and the size of urban areas as a result of market processes (Anas et al., 1998). A major concern is the importance of space in economics, reduced in the monocentric model to a simple measure of distance from the urban centre (Irwin and Geoghegan, 2001). As highlighted by Nilsson (2014, p.46) "*the spatial distribution of amenities is an important determinant of urban development patterns and plays a major role in shaping the urban spatial structure*" and should thus be considered in the demand of residential land consumers.

The monocentric model is often criticised for lacking realism, but although it is an extreme simplification of reality, it manages well to describe urban spatial expansion and to identify the main drivers of urban growth (Brueckner, 2011). A main reproach to this model is that consumers' preferences are reduced to the sole desire for larger properties and being located close to the main employment centre. The consumers' decision to purchase land is acknowledged however to be more complex, as it relies on decisions made by heterogeneous households with heterogeneous preferences and needs. Consumers seek to improve their quality-of-life, largely determined by social interactions, economic activity and local amenities (Bockstael and Irwin, 2000). Several extensions to this model have been presented that aim at accounting for these consumers' and spatial heterogeneities. Neighbourhood qualities and landscape features add up to the standard trade-off and add further complexity to the spatial structure of land values. Recent theoretical advances have shown that the local arrangement of green space impacts on urban form and its scattered or leapfrogging nature (Cavailhès et al., 2004; Turner, 2005; Caruso et al., 2007, 2011). There has been an increasing number of theoretical and empirical analyses on the effect of green amenities in residential location choice. A review of theoretical models from urban economic and geographic literature considering the heterogeneity of urban patterns with regard to green amenities can be found in Bockstael and Irwin (2000); Caruso and Cavailhes (2010). Especially, the periurban neighbourhoods bear a variety of local characteristics expected to be valued by residential land consumers and therefore urban economists show increasing interest in land-use and its spatial distribution.

Some theoretical models aim at accounting for the consumers' preferences for such neighbourhood amenities. It is well-known since Tiebout (1956), that the provision of local public goods is an important aspect of residential competition. Further, it was shown that local urban amenities impact on the income sorting of households (Brueckner et al., 1999; Glaeser et al., 2001). Brueckner et al. (1999) provided evidence that the spatial pattern of exogenous amenities in a city impact on the location of different income groups and that the valuation of these amenities is rising rapidly with income.

The sustained growth of the periurban area suggests that residential land consumers are attracted to this area and obtain utility from the specific morphology and functionalities of these neighbourhoods. A growing body of literature (i.e. geography, urban economics, spatial and urban planning) emerged with the aim to define, explain and measure the causes, the extents and consequences of urban spatial expansion. The periurban concept and how it differs from other conceptions, for instance *urban sprawl* or *Zwischenstadt*, is further discussed in appendix A.1.

Cavailhès et al. (2004) model the emergence of the periurban belt, where residential and agricultural land-use coexist (Caruso, 2005). The periurban area is defined as a mixed area between residential and agricultural land-use, by accounting for the utility bearing properties that arise from the mix of rural morphology and in the meantime a variety of local urban amenities generated by the interactions among urban and rural agents. At the same time they consider consumers' preferences for being close to a major employment centre and their taste for larger properties.

Based on the periurban model and the observations made with regard to the causes and consequences of urban spatial expansion, the importance of the individual decisions in land-use allocation and how they shape urban form and patterns has been highlighted. Consumers' decisions are based on the preferences they have for periurban amenities and the standard trade-off identified in the monocentric model. Eventually the individual consumers' major aim is to maximise their quality-of-life with regard to the available budget.

0.3 Research questions and objectives

Recent urban growth patterns challenge sustainability and social goals that many urban and land use planning actions seek to address at the local scale (municipality or smaller). For the success of these urban policies, it is particularly important that the benefits of local amenities are well understood to design effective and acceptable neighbourhood plans.

In this perspective and with regard to the specificities of the Luxembourgish context, this research aims at evaluating **how the local geographical context is considered by Luxembourgish land consumers when purchasing developable land**. The first objective underlying this research is hence to find means to quantify and identify the geographical determinants of developable land prices. As illustrated in figure 1, the monocentric and the periurban model built the theoretical basis of this research. The preferences of consumers are further expected to vary with their socio-economic background and through space, important circumstances that need to be accounted for. The consumers' preferences will be approached from different viewpoints and the methods applied aim at accounting for restrictions related to data availability and aggregation as well as issues related to spatial effects.

Quantification of neighbourhood amenities

Based on the above mentioned theoretical framework, a first challenge of this research lies in the collection and generation of the dataset. To date, studies on the Luxembourgish real estate market mainly focus on the price evolution and the general high housing prices. The focus on the residential land market and the perspective from the demand side of the market has, in this way, not yet been considered in the Luxembourgish context. In addition, we contribute to quantitative research in Luxembourg by generating a large database considering, in addition to the municipal level, the section scale. The dataset is expected to be of further use for future studies in Luxembourg.

Besides the transaction specific and structural characteristics of developable land transactions, the focus was turned to the quantification of the local periurban amenities and some control variables. The location-specific amenities are summarised in four major categories: accessibility to Luxembourg-city, the provision of local urban and green amenities, as well as the socio-economic neighbourhood composition. Innovation is provided by some of the explanatory variables considered, going beyond those traditionally considered in hedonic models. For instance, the simultaneous consideration of access by car and public transport or net population density and, most of all, the land-use and local urban amenity diversity indices, requiring advanced GIS tools. A description of the transaction dataset and the explanatory variables is given in part I.

Estimating consumers' preferences

The determinants of individuals' residential location choice as identified in urban economic literature are not traded explicitly on a single market. They are assumed to be capitalised in the price consumers pay for land. Residential land should thus be seen as a composite good rather than a "*generalised housing commodity*" (Brueckner, 2011, p.117). To identify the benefits obtained by individuals and to evaluate the negative impacts of other amenities in the perceived neighbourhood of residential land parcel, we rely on the hedonic pricing method as developed by Rosen (1974). This method considers residential land as a composite good, composed of the different attributes from which consumers obtain utility. Econometric estimation techniques allow to decompose the observed transaction price with regard to its structural and local geographical specificities, and hence to estimate the implicit prices consumers are willing to pay for them. A general presentation of the hedonic pricing method and its assumptions is provided in the empirical part of this thesis, in section 3.2.

Caution has to be taken with regard to the spatial dimension of the data and processes analysed[8]. Spatial econometric techniques are used to obtain unbiased estimation results. In chapter 3, these techniques are presented and applied to the developable land transactions in Luxembourg. It is now commonly accepted that hedonic pricing models need to account for different forms of spatial dependence (Krause and Bitter, 2012) and

[8] Different forms of spatial effects are further discussed in the introduction to part II p.93.

many empirical examples can be found (e.g.: Brady and Irwin, 2011; Furtado, 2011b; Abelairas-Etxebarria and Inma, 2012; Seya et al., 2013, ...).

Preferences varying with consumers' socio-economic backgrounds and market segments

A major drawback of the monocentric model is that it assumes identical consumers with identical preferences for residential land. Several extensions to this model have been presented highlighting the importance of accounting for the heterogeneous socio-economic backgrounds and preferences of land consumers to further explain urban form. The consumers' willingness to pay for the attributes composing the good should be related "*to specific constraints of individuals and households, including economic status (employment and income), motorisation, family structure and life cycle*" Thériault et al. (2005, p.23). A second step of the hedonic pricing method (Rosen, 1974; Brown and Rosen, 1982) allows the estimation of the demand for different price determinants identified in the first step, taking the consumers' socio-economic characteristics into account. In other words, this second step allows to identify variations in the demand for certain attributes according to their income or household composition. With regard to the data available in the context of this research, Rosen's 1974 second stage could however not be estimated.

The second step of the hedonic pricing method will be approximated by spatial quantile regression techniques (Koenker, 2005; McMillen, 2013), still rather uncommon in the hedonic context. Quantile regression techniques allow to differentiate the implicit prices of attributes according to different price ranges and hence to estimate the marginal prices of attributes according to the consumers' ability to pay.

In addition, the land transaction dataset does not allow to distinguish between private or professional consumers of developable land. Further, professional developers are suspected to behave differently compared to private land consumers, purchasing land for their individual needs. The question that arises is if in their investment decision professional or public land developers, who aim at maximising their profit from developing the land, consider local urban amenities in the same way as individual "end-users".

Are these different actors in competition on the same market? Data privacy constraints denied essential information on the purpose of the acquisition, and it is hence the first attempt to account for such market segments in the developable land market in Luxembourg.

Spatial market segmentation and non-stationary marginal prices

Non-stationary marginal prices have been observed in the hedonic literature, mostly relying on geographically weighted regression (GWR) techniques. **The valuation of the structural (e.g.: parcel size) or location-specific (e.g.: land-use patterns) characteristics of land prices are expected to be valued differently through space**. Although it is today commonly accepted that hedonic models need to account for spatial dependence, estimation issues related to spatial heterogeneity are rarely considered at the same time.

The autoregressive functions developed in spatial econometric literature can be seen as "technical fixes" to the problems of modelling spatial data according to Orford (2000). Especially as they do not account for problems related to heteroskedasticity and spatial heterogeneity. Further, it is most likely that the Luxembourgish developable land market is segmented through space. Spatial heterogeneity results possibly in spatial heteroskedasticity or spatially varying parameters (De Graaff et al., 2001; Wilhelmsson, 2002; Le Gallo, 2004) and consequently the linear estimation imposing spatial homogeneity will be misspecified and affect the validity of diagnostic tests (Anselin and Lozano-Gracia, 2009). Moreover, the dataset considered represents three levels of aggregation, that are the transaction, the section and the municipal scale.

In this perspective, the multilevel modelling approach has recently been introduced in hedonic modelling context (Orford, 2000, 2002). A three level model will be implemented to identify and account for spatial variations in the valuation of the attributes of land prices. Further, this method should allow to account for the different levels of data available. This case study contributes to the empirical hedonic multilevel literature, especially by considering three levels.

The main research questions can be summarised as follows:

- **What local periurban amenities, besides the standard trade-off, are valued by developable land consumers?**
- **Are there different consumer-based market segments? How does the consumers' socio-economic background impact on their preferences for the price attributes, mainly local amenities?**
- **Is there spatial market segmentation and spatial heterogeneity in the valuation of the determinants of land price?**

0.4 Thesis outline

Figure 1 illustrates the organisation of the thesis and highlights its subdivision into two main parts: In part I, the different datasets, generated in the framework of this research and derived from urban economic theory, are presented. The objective of this part is twofold, a description of the context as well as stating the general expectations towards the different explanatory variables. This part provides insights into the dataset generated, it is divided into two main chapters. On the one hand the land transaction dataset will be presented (chapter 1) and on the other hand the location specific explanatory variables are discussed (chapter 2).

In part II, the implementation of the hedonic pricing method, relying on the dataset presented in part I, aims at providing answers to the questions raised above, while accounting for the presence of spatial effects through state-of-the-art econometric techniques. The objective of this thesis is to provide differentiated insights into the residential land market and the determinants of residential location choice in Luxembourg. A global spatial hedonic model is presented in chapter 3. In chapter 4 insights into consumer-based market segmentation and varying preferences for local periurban amenities based on the consumers' socio-economic background are provided. In chapter 5, spatial heterogeneity and spatial market segmentation are observed and discussed relying on the multilevel modelling approach.

Eventually, the main findings of this thesis, subject to some limitations, are presented in the conclusion and future perspectives and additional tracks for investigation are pointed out.

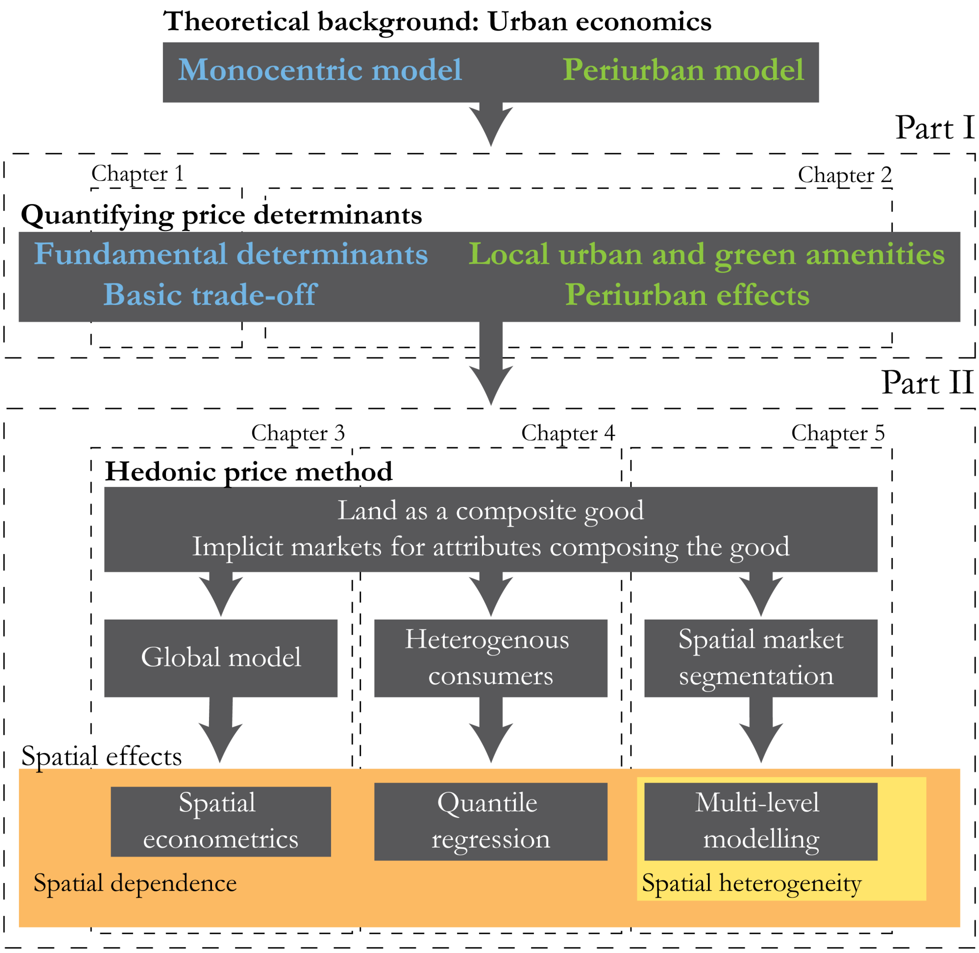


Figure 1: Thesis outline: illustration

Part I

Developable land transactions and explanatory variables: The $\boldsymbol{Y}$ and $\boldsymbol{X}$'s

Introduction to part I

Relying on the theoretical underpinnings of urban economic literature and the findings of previous hedonic studies, this part is dedicated to the presentation of the data generation process. The hedonic pricing method considers land price (Y) as composed of a wide variety of structural and location-specific attributes (X), which land consumers consider when deciding to buy residential land.

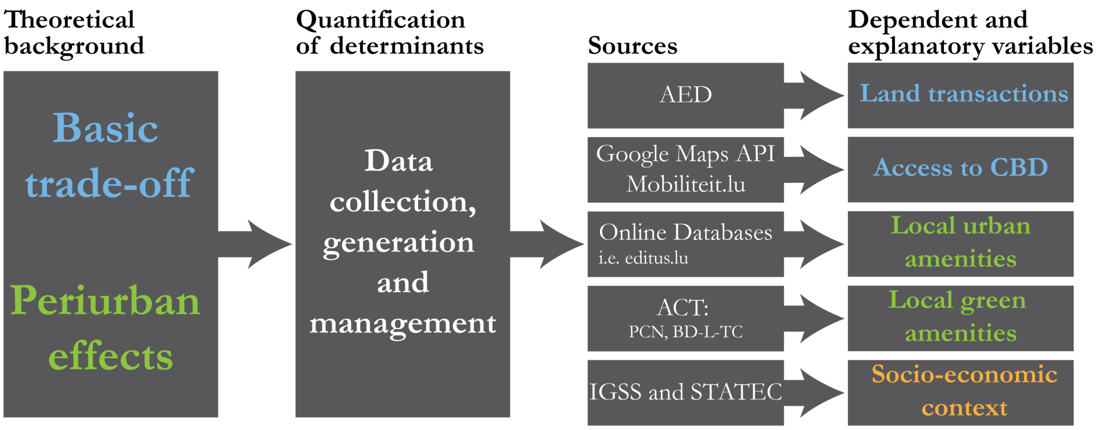


Figure 2: Considered datasets and sources

The quantification of the local amenities was an important and challenging part of this research, the wide variety of databases and sources relied on are summarised in figure 2. Although we had access to advanced geographical tools, the aggregated scale of the real estate dataset did not allow to account for the local transaction context. Numerous variables were generated at micro-scale (e.g.: landscape measures, local urban amenities), but eventually we had to rely on the more aggregated section scale. This contributes to the existing literature in Luxembourg, where quantitative research is generally at municipal scale. The challenge was to generate measures of the local transaction context at aggregated scale, by approximating transactions' neighbourhood statistics. In addition, this part provides a more general descriptive analysis of the Luxembourgish context at the finer, sub-municipal, scale. To identify the periurban area in

Luxembourg and to frame the descriptive part and to highlight regional differences, a typology at section scale is presented in appendix B.

In the following chapters, the different datasets underlying this research will be presented, providing a literature review on the theoretical and empirical findings related to the different determinants and the Luxembourgish context. In chapter 1, the Luxembourgish real estate market is discussed, in general and with a focus on the developable land transactions and the related structural variables in time and space. Location-specific variables, as derived from urban economic theory are presented in chapter 2. The main location-specific determinant of land price as identified in the monocentric urban model is access to the main employment centre, which is described in section 2.1. Besides the basic trade-off the focus is further turned on the periurban features: the local urban amenities, presented in section 2.2 and the land-use amenities, discussed in section 2.3. Eventually, the socio-economic context is presented in section 2.4.

This part aims thus at presenting:

- **the different datasets and variables.**
- **the related weaknesses and limitations.**
- **the expectations of their marginal impact on land prices.**

Chapter 1

Luxembourg's developable land market

The current challenges of periurbanisation and real estate market faced by spatial planning politics in Luxembourg request for more detailed insights into the individuals' demand and preferences towards housing and land. To date, analysis of the real estate market in Luxembourg is mainly conducted at the level of different administrations[1] and different research institutions[2] regrouped in the "Housing Observatory"[3].

In this framework, several studies have been published analysing the real estate market in Luxembourg and with a particular focus on the evolution of prices in time. On a trimester basis the housing observatory publishes indicators on the announced sales and rental prices of different dwelling types (Observatoire de l'Habitat, 2014a,b,c). In addition, the "*Notes de l'Observatoire*" and thematic reports focussing on the housing market (Observatoire de l'Habitat, 2009) and land supply (Observatoire de l'Habitat, 2007, 2010) have been published. In these studies the demand side is secondary and, to our knowledge, the residential land market is mostly ignored. This makes this research complementary to what has been done so far on the real estate market in Luxembourg.

This chapter aims at providing in section 1.1, an overview of the market for land available for construction in Luxembourg via the developable land transaction dataset provided by the Administration of Deeds (AED). With regard to the restricted amount of information available on the structural characteristics and future use of these parcels, some assumptions on the consumers and sellers of land in Luxembourg will be presented.

[1]Generally public authorities as for example the Ministry of Housing.

[2]STATEC, BCL and CEPS/Instead.

[3]The "*Observatoire de l'Habitat*", created in 2003, has three main objectives: the collection, the analysis and the communication of information relevant for the spatial planning and housing policy in Luxembourg. Their research focus is turned on the analysis of different segments of the housing market considering announced and observed prices. Further, land-use change, the supply of land and the comfort of housing units are at the centre of attention. The observation of spatial developments in Luxembourg has been recently entrusted to the "Observatoire du Développement Spatial" (ODS), with the mission to follow and improve spatial development by collecting and analysing relevant data in the spatial context.

Further, the evolution of transactions and prices in time are described in section 1.2, while the variations in space are eventually discussed in section 1.3.

1.1 Developable land transactions

The real estate transaction dataset was provided by the Administration of Deeds. The dataset includes all transactions registered by the notaries at the AED in the period between January 2007 and December 2011, the latest update was provided in June 2012[4]. The quality of this administrative database is limited for the purpose of statistical and econometric analysis[5].

In Luxembourg, real estate transactions have to be declared authentic by a notary, registering them at the AED. Although no parcel can be sold without a deed (*acte notarié*), there is yet no explicit inventory on what characteristics are required to figure in the deed at registration. Therefore, the AED dataset is very heterogeneous and lacking detail, especially for dwelling transactions (i.e. houses and apartments), where important characteristics are not always indicated at registration.

A first, and very pragmatic, reason for concentrating on the developable land market, rather than the housing market, is thus related to the quality of the available dataset. With regard to the complexity of built structures, hedonic studies analysing preferences of house or apartment consumers have in general information on a goods' structural characteristics at hand (e.g.: amount of bedrooms, date of construction). The price of undeveloped land may be free of the complicating effects of construction structures, while it is most certainly not free of the presence of local effects (Cheshire and Sheppard, 1995).

The main and second reason for concentrating on land prices, rather than built property transactions, is related to urban economic theory and our overall research context. Land-use change mainly occurs if the equilibrium residential bid-rent is higher than the agricultural bid-rent, in this case farmland is transformed to residential land.

[4]Despite this update, some transactions might still be missing, due to delays in registration by the notaries at the AED.

[5]Some substantial clean-up was necessary to make the database suitable for our research, the different steps of data management as well as the limitations faced are detailed in appendix C.1.

Considering the demand for land is of special interest as spatial planning policies mainly work by regulating and limiting its supply (Cheshire and Sheppard, 1998). In order to understand the geographical determinants that push individuals to locate outside the urban area analysing the demand for residential land and the extent to which its price is determined by the local context is of major interest.

As defined by the AED, a developable land transaction is a parcel of land available for construction, located inside the municipal building perimeter and not occupied by any construction at the moment of the transaction[6]. In general, these land transactions are assumed to mostly represent land foreseen for residential land-use, however some of the transactions might be affected to other uses (e.g.: commercial, office space). This is a limitation to our analysis, since consumers of non-residential land are expected to value location-specific attributes differently.

Since all transactions get a unique ID by the AED, accounting for repeated sales was not possible. Therefore, no insights into the speculative and hence anticipatory behaviours of some actors could be gained.

A further limitation related to the AED dataset is the missing information on both landowners and consumers. As presented in the introduction, theory and empirical findings strongly suggest that the socio-economic background of consumers affects their preferences for properties' characteristics. A limitation that might be partly overcome by concentrating on land transactions, since it is assumed that consumers of residential land are less heterogeneous and that there are less potentially endogenous variables (McDonald and McMillen, 2011). However it was impossible to distinguish directly between private, professional and public actors (e.g.: individuals, real estate promoters, the state or municipalities), a shortcoming further addressed in section 1.1.2.

The main limitation faced, with regard to the AED dataset, is however related to the aggregated location of the transactions. The finest scale available in the framework of this research[7] is the sub-municipal scale[8]. Consequently, it was not possible to undertake

[6]Land with other uses or for other purposes (e.g.: forest, public green space) has not been considered.
[7]Based on the decisions of the AED and the National Data Privacy Commission (CNPD).
[8]The illustration of the spatial distribution of the different transaction related information will only

studies as sophisticated as those presented among others by Brossard et al. (2008); Des Rosiers and Thériault (2006) or Jim and Chen (2009), to name a few. Throughout this research, the aggregated data level entailed methodological and data issues further detailed in the following sections.

1.1.1 Structural parcel characteristics

After price deflation and clean-up[9], the general descriptive statistics of land transactions are summarised in table 1.1. The average transaction price is of 327,724€ and it is sized 6.07 are (607 square metres). As expected, land prices are increasing with parcel size (fig. A 1.1) and with increasing size the per unit price is decreasing (fig. B 1.1).

Variable	Mean	Min	Max
Total price	327,724	14,454	2,878,744
Price per are	68,444	9,131	510,384
Parcel size	6.07	0.63	50.52

Table 1.1: Developable land transactions: Basic statistics

However, regressing the size as well as squared size on the total land price reveals that above 22.77 are, the unit price is increasing with size (fig. B 1.1). These very large parcels are assumed to be rather bought by professional developers, in a perspective of investment or speculation. They are expected to either concern large scale residential development projects, that might also be dedicated to non-residential construction projects (e.g.: commercial or office space).

Parcel size is the main structural determinant of land price, as shown in urban economic theory (Alonso, 1964) and in hedonic price literature (among many others: Cheshire and Sheppard, 1995; Cavailhès, 2005). An increase in parcel size is expected to have a positive marginal impact[10]. Deriving additional structural parcel attributes (e.g.: orientation, relief or parcel form) was impossible due to the aggregated location of the data.

be provided at aggregated municipal scale in respect with potential data privacy issues.

[9]Presented respectively in section 1.2.1 and appendix C.1.

[10]In regression the mean centred and natural log of size will be considered, to account for the non-linear relationship and to ease the results interpretation.

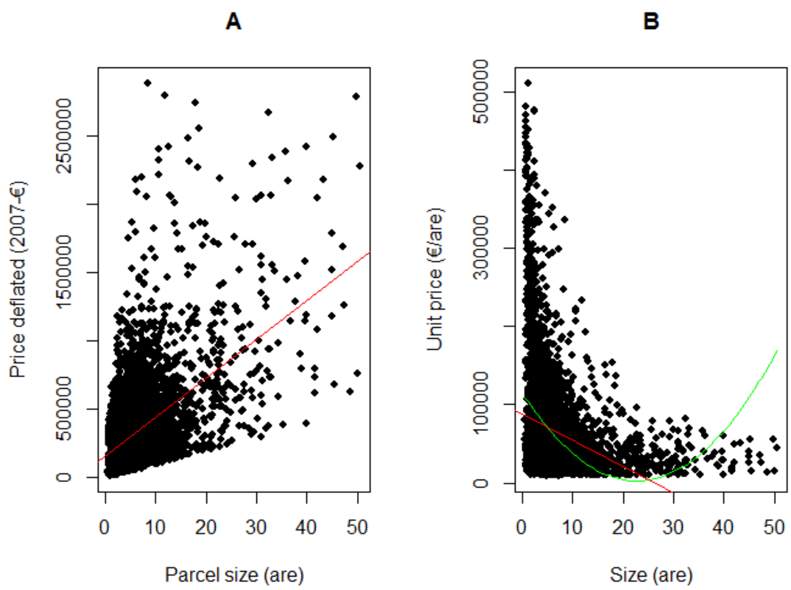


Figure 1.1: Transaction price versus parcel size

1.1.2 Assumptions on consumers and sellers based on transaction types

The socio-economic context of land consumers is expected to have an impact on their preferences for the structural and local amenities composing the good (Brown and Rosen, 1982; Palmquist, 2006). This is further complicated as the transactions are not exclusive to residential land-use by private "end-users". Land might also be bought by professional investors, real estate developers (RED) or public authorities. Although we assume them to mostly cover residential development projects, this adds the possibility of non-residential uses. In general, distinction should be made between individual residential "end-users" and those that purchase land in an investment logic. While professional land developers choose the density and location of development to maximise profits (Wu and Plantinga, 2003), individual land consumers are mainly concerned about maximising their quality-of-life, as already discussed. Accounting for these different types of consumers in the hedonic model is important since the implicit preferences for the attributes are most likely to vary among consumer types. First, some general insights into the Luxembourgish land and housing consumer will be presented. Second, attempts to account for the different actors are discussed.

With regard to the existing studies on the Luxembourgish residential land market, little general information on the owners of developable land is available from official sources. In 2007, 78.94% of the parcels available for construction were owned by private individuals while 10.62% were held by corporations (i.e. real estate developers, associations, state or municipalities) (Observatoire de l'Habitat, 2012). Between 2007 and 2010 a change in the residential landownership structure was observed and the part of developable land belonging to corporations increased to almost 15% (Observatoire de l'Habitat, 2012). The implication of professional and public developers has thus increased, implying an increasing amount of transactions from private landowners to developers.

Although the prevailing residential structure in Luxembourg is the single family house, 83.5% of all residential buildings of which 37.9% are isolated houses. Zahlen (2012d) showed that between 1990 and 2009, apartments represented around 70% of the built dwellings, confirming the findings of the last census, that 21.2% of all apartment buildings existing in 2011 were built after 2001 (Allegrezza et al., 2014), while in total 11.5% of the residential buildings did not exist in 2001.

Based on the findings of the 2011 census some coarse insights into residents' households and their preferred dwelling type are provided here. Almost 64% of households are found to live in single family buildings, while 27.6% live in isolated single family houses. On average, 73% of the residents own their home, while 84.5% of people of Luxembourgish nationality are home owners (Allegrezza et al., 2014). In general, the ownership of different property types depends largely on the living standard, age and the employment status of the head of household (Berger, 2004). Ownership rates increase from the age of 30 onwards and vary with the households' composition (Allegrezza et al., 2014). Less than 55% of single person households own their dwelling, while the average ownership rate is above 70% for households with a minimum of two persons and increases as household size goes up.

The AED dataset provides some information on the type and legal framework of the transaction, allowing some assumptions on the consumer types (private or professional), on the buyer and on the seller's side.

Big sized transactions (BigSize): As presented in section 1.1.1, the unit price per are generally increases for transactions larger than 22.77 are, representing 2% of the observations (fig. B 1.1). In the perspective of gaining further insights into the actors involved in land transactions, a dummy variable was generated for these large-size transactions (*Bigsize*). These parcels are assumed to be mostly bought by professional land developers (real estate developer (RED) or public actor) rather than individual "end-users", and thus most likely bought in the perspective of future developments that may or may not be residential. While the average size of these 128 parcels is 31 are the average unit price is almost divided by two for the big sized parcels (table 1.2).

Since the objective of this thesis is to identify individuals' preferences when purchasing residential land, these transactions should be controlled for in the hedonic model[11], we expect them to have a positive impact on prices, after size and all else is held constant.

	Count	Price/are	Mean price	Mean size
Big Size = 0	6,239	69,178	313,456	5.56
Big Size = 1	128	32,681	1,023,180	31.15
Total	6,367	68,444	327,724	6.07

Table 1.2: Transactions of *Big Size*: statistics

Development projects (dVFA): The increased part of land belonging to professional and public developers has been discussed in the beginning of this chapter. It is assumed to be mainly purchased with the objective of future residential development and sold to individuals. In the AED dataset a dummy informs on the presence of such projected development plans.

Beside "normal sales" (VEN)[12], these transactions with development plans (*dVFA*)

[11] A limitation related to this dummy is that it is not possible to control for consumers of larger parcels in the case adjacent parcels were bought in several transactions from different owners and/or at different points in time.

[12] 79% of the transactions are registered as normal sales.

are the most frequent type, with 19% of the transactions. According to the AED a transaction of type *dVFA* includes a contract by which the seller assigns his rights on the land immediately to the purchaser, while the planned constructions become property of the purchaser as the project progresses, while the seller remains the “maître d’ouvrage” until final inspection[13]. As summarised in table 1.3, these transactions register higher mean prices per are, while the mean parcel size is on average smaller for *dVFA* transactions. *dVFA* transactions are assumed to be mostly sold by real estate promoters to individual households, generally in the framework of a larger development project; this variable should provide insights into the marginal effect on price of the implication of professional developers, which is expected to be positive.

	Count	Price/are	Mean price	Mean size
dVFA = 0	5,128	63,569	330,562	6.55
dVFA = 1	1,239	88,623	315,977	4.11
Total	6,367	68,444	327,724	6.07

Table 1.3: Transactions of type *dVFA*: statistics

Transactions with non-residential building (dTerrain): Besides the price available for all transactions (*Montant*), supposed to consider pure price paid for the land, a specific *Terrain* price was indicated for 516 transactions. For these the ordinary *Montant* was then dismissed, since for these observations, *Terrain* indicates the price paid for the land, while some additional features have been sold at the same time, most likely existing constructions of non-residential use like for instance barns or ruins. In general, the average price, per unit and the size of *dTerrain* observations are lower than for all other transactions (table 1.4). Additional fees are expected to arise related to the need of additional development costs, associated with the destruction of the built structure.

[13] There is no explicit definition of the type and progression of the development project for the *dVFA* transactions (e.g.: first administrative steps, building permit request, or the connection to the municipal infrastructure). Regardless the extent of preparations done for parcel development, we assume that some fees (e.g.: administrative, architect) have been raised, although these should not be added to the land price.

	Count	Price/are	Mean price	Mean size
dTerrain = 0	5,851	69,550	338,724	6.25
dTerrain = 1	516	55,902	203,003	4.08
Total	6,367	68,444	327,724	6.07

Table 1.4: Transactions of type *dTerrain*: statistics

Expected results

The structural and type variables are assumed to explain a considerable part of the total transaction price. The overview provided in table 1.5 summarises the expected marginal effects on the price of a "typical" transaction. Increased size is expected to have a positive impact on price, as are *BigSize* and *dVFA* dummies, while *dTerrain* is expected to have a negative impact. Regression results should thus represent the preferences revealed by private individuals as residential "end-users" or professional developers investing in smaller, mostly residential, projects, bought from private landowners.

Variable	Unit	Description	Impact
Transaction price	€	Total price	Dep. Var.
Size	are	Total size	+
dBigSize	dummy	1 if size $\succ$ 22.27 are	+
dVFA	dummy	Existing development project	+
dTerrain	dummy	Complex transactions	-

Table 1.5: Transaction specific variables: Expectations

1.2 Land prices in time

The general high and constantly increasing real estate prices in Luxembourg are a major political challenge. Frequently the question of a speculative bubble is raised in this context and some studies dealing with this debate are briefly presented here. A recent analysis of a potential speculative bubble[14] in the Luxembourgish real estate market has been undertaken by Blot (2006), distinguishing between housing, non-residential and residential land markets. The observation of an important price increase is not a

[14] A speculative bubble is defined as an excessive and abnormal evolution of an assets price (Blot, 2006).

sufficient condition for a speculative bubble, an analysis of such a price evolution cannot be done independently of what would be considered as the "normal" evolution of prices, with regard to the fundamentals[15]. Blot (2006) tends to reject the bubble hypothesis for the land market as prices were not systematically disconnected from the fundamentals and their evolution was most probably related to the supply side.

Since the publication of this study in 2006, and considering that the continuously increasing real estate prices, hardly slowed down in the events of the economic crisis, the hypothesis of a speculative bubble have reappeared in the last years. In 2013, Licheron (2013) analysed the possibility of a speculative bubble in the real estate market in Luxembourg. He mainly confirms the results of Blot (2006) that actually no major speculative movements can be observed and that the high prices are mainly explained by the strong demand for housing and the limited supply of dwellings (Licheron, 2013).

To account for the temporal dimension of the transaction data in the model, prices were deflated (section 1.2.1) and their evolution will be described in in section 1.2.2.

1.2.1 Price deflation

Some temporal variables were generated based on the date of transaction (e.g.: year dummies, seasonal dummies) to identify and account for the evolution and seasonal fluctuations within land prices in the regression models through year dummies for instance. However, it is common practice in hedonic pricing studies today to rely on standardised prices at a specific date (Geoghegan, 2002; Treg, 2010; Carruthers et al., 2010; Kadish and Netusil, 2012; De Bruyne and Van Hove, 2013). Deflating the price to a fixed date, relies on the assumption that the marginal effects remain constant in time and this assumption seems appropriate for short time periods according to Anselin and Lozano-Gracia (2009). Further, the time-series approach to capturing the impact of geographical characteristics on property prices is less appropriate as these factors remain rather constant over time (De Bruyne and Van Hove, 2013). Therefore, based on preliminary estimations considering the time dummies and with regard to literature, the

[15] As fundamentals Blot (2006) considers economic and financial variables that can exert a significant influence on real estate supply and demand, and thus on the prices (e.g.: construction costs, financing conditions).

transaction price was deflated to 2007-Euros.

To deflate the transaction price to January 2007-Euros ($P_{Jan2007}$), we accounted for the (natural) inflation during the reference period by considering the consumer price index (CPI) generated on a monthly basis by the STATEC. The CPI measures national inflation, considering residents' final consumption expenditures (STATEC, 2014).

$$P_{Jan2007} = P_i * (CPI_{Jan2007}/CPI_i) \quad (1.1)$$

In equation 1.1 the transaction prices (P_i) at month (i) was multiplied by the ratio of the CPI in 2007 ($CPI_{Jan2007}$) and the consumer index at transaction date (CPI_i).

1.2.2 Land prices between 2007 and 2011

The evolution of the average transaction price per are between 2007 and 2011 is illustrated in figure 1.2. The mean price per are before deflation is represented in light red, while the deflated price accounting for the CPI is illustrated in dark red.

In this rather short period of five years, the number of transactions (blue bars) is rather constant, only a slight increase can be observed (light blue line). Some seasonal fluctuations and peaks become however apparent in the number of land transactions registered. In general, a certain slowdown is noticed in 2008 and 2009, most probably related to the general economic crisis. What is more, the number of transactions is generally lowest in August and in January and highest in July and December. Main peaks in the number of transactions can be identified in the fourth trimester of 2007 and in December 2010. The summer season slowdown is likely related to the holiday season and a general slowdown of economic activities in this period. The regular increase of the number of transactions in December is mainly related to a seasonal effect generally observed at the end of a year before a slowdown of transactions in January (Observatoire de l'Habitat, 2011). No clear relationship between the unit price and amount of transactions can be identified.

Despite the deflation to January 2007-Euros a slight general increase of prices is still noticed as well as some seasonal fluctuations, probably related to the increased

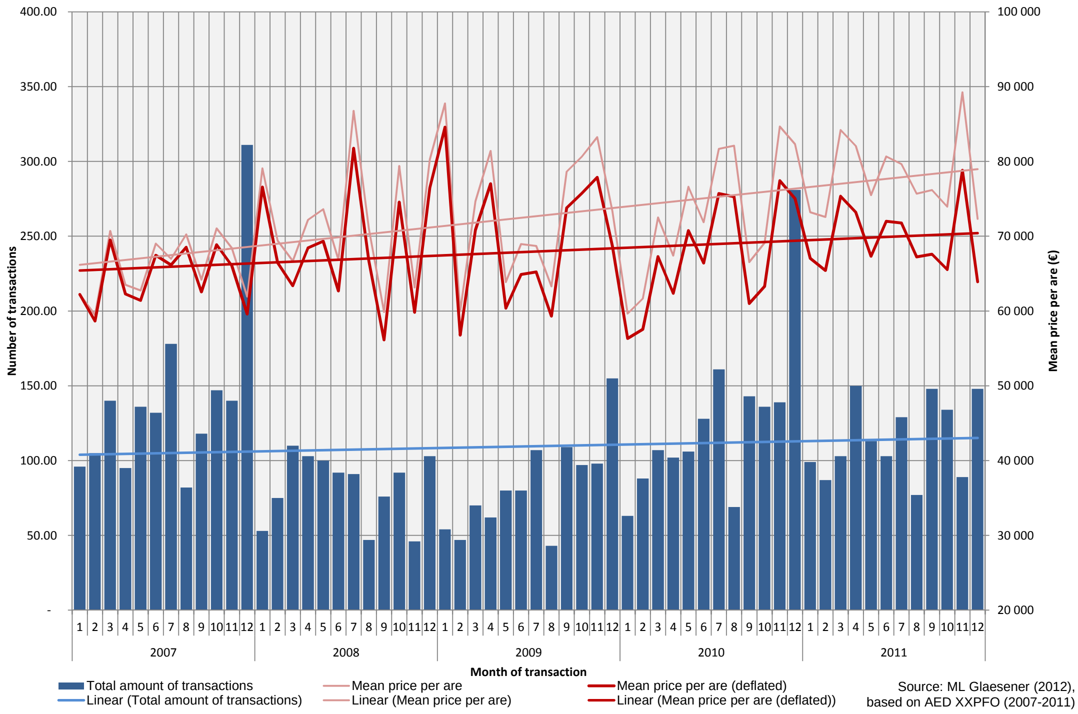


Figure 1.2: Mean price per are and number of transactions in time

demand, as suggested by Licheron (2013). The first peak is assumed to be related to the more favourable tax regulations for selling, in force until the 31^{st} of December 2007[16]. To benefit from these lower taxes on real estate sales, we assume that mainly private landowners seized this opportunity for selling, supported by the observation that less than 7% of transactions in December 2007 were sold under type VFA (fig. 1.3). Although less pronounced, this was also the case for the second peak in December 2010, where the suppression of interest subsidies and tax credits on foreseen deeds, incited consumers to buy (Observatoire de l'Habitat, 2011). Seasonal fluctuations will not be further discussed because of the short observation period and since according to Cavailhès (2005) such price variations, mostly observed on housing markets, do not seem as pronounced in the case of residential land prices.

[16]Installed by the Loi du 30 juillet 2002 déterminant différentes mesures fiscales destinées à encourager la mise sur le marché et l'acquisition de terrains à bâtir et d'immeubles d'habitation (2002).

In figure 1.3, distinction was made between the evolution of unit prices and the amount of transactions with existing development plans ($dVFA$) and all other transactions. During the reference period the part of transactions registered as $dVFA$ increased, which confirms the general observations made at the beginning of this section. Their average unit price, after deflation, increases during the reference period, while on the contrary the unit price of all other transactions remains rather constant. Hence the general increase of land prices still observed might be explained by the larger share and evolution of $dVFA$ transactions and be related to profit maximisation approach of the professional developers.

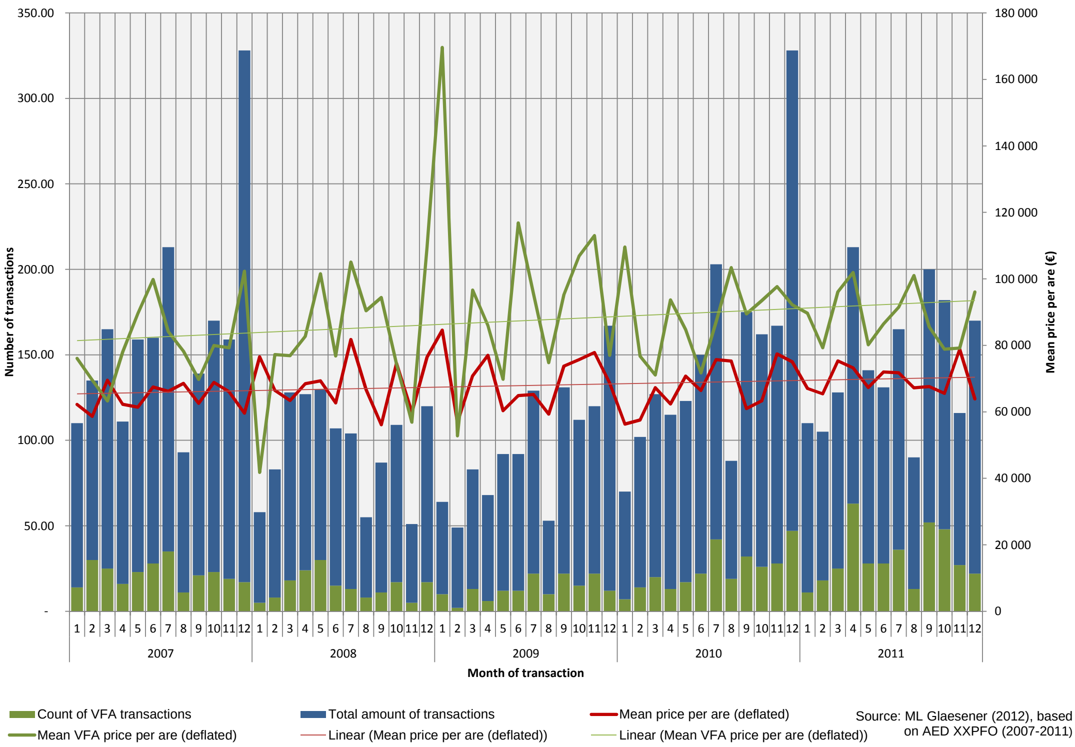


Figure 1.3: Total transactions vs. VFA transactions

1.3 Heterogeneity of transactions in space

Some limitations were faced since not all sections and municipalities registered transactions during the reference period. Hence, only transactions were kept that are located within sections registering more than three transactions. Under these conditions 200 sections of 521 did not figure in the analysis. While of those sections, 94 did not register any transactions, between one and three transactions were registered in 106 of the ignored sections[17]. In addition, only municipalities with more than six transactions have been considered in the regression analysis, the same threshold is considered for instance by the housing observatory for apartment sales (Observatoire de l'Habitat, 2014c) or by Djurdjevic et al. (2008) in the Swiss context. Eventually, the 6,367 transactions kept are distributed over 321 sections located within 113 municipalities. There are on average 56.35 land transactions registered per municipality, while a section registers on average 19.83 transactions for the considered reference period.

1.3.1 Spatial variations in price and size

The distribution of the transactions across Luxembourg is illustrated in Map A figure 1.4. Fewer transactions are registered in the municipalities in the north of the country compared to the south, where a concentration of residential land transactions is observed. At the same time, as illustrated in Map B figure 1.4, the average unit price also varies through space. While highest in the municipalities of the southern part of the country, lower averages are observed in the north. Developable land prices per are are highest for Luxembourg-city and its periurban area (as well as in the south-west). In the areas that witness an important population growth, prices are still below the national mean. In general, prices decrease with distance to the capital[18], confirmed in graph A figure 1.5 and it also confirms the basic findings of the monocentric model (Alonso, 1964; Muth, 1969; Mills, 1967). In spite of this, some more remote municipalities (generally regional urban centres) register slightly higher average unit prices.

The distribution of average parcel sizes varies between municipalities (C fig.1.4), however no clear spatial pattern arises. Graph B in figure 1.5, confirms that transaction

[17]Figure E.1 in appendix shows the location of these sections, those sections without a critical low number of transactions are represented in grey. We observe that for instance no transactions have been registered in the historical city centre of Luxembourg-city and that the northern part of the country registers many sections without transactions.

[18]Distance being measures by travel time by car, these distance measures are detailed in section 2.1.

size slightly increases with distance to the capital.

On average transaction size is smallest within sections of the Agglomerated South[19], with 4.83 are on average and a median of 2.99 are. These observations should be considered as a particularity of this region and are assumed to be related to differences in land-use traditions related to its industrial past. Table 1.6 shows that about 65% of the transactions are located in the (close and distant) periurban zone, with almost 41% of transactions located in the close periurban area (30% of the total area of Luxembourg). Further, 10% of the transactions took place within the agglomeration of Luxembourg, while only 6% in the southern more densely populated area (Agglomerated South)[20].

C15	Count	of total	Price/are Mean	Price/are Median	Size Mean	Size Median
Rural	1,121	17.61%	41,826	30,425	6.37	5.53
Distant periurban	1,556	24.44%	59,139	45,031	6.10	4.79
Close periurban	2,607	40.95%	70,501	56,272	6.14	4.89
Agglo South	422	6.63%	98,150	62,110	4.83	2.99
Agglo Luxembourg	661	10.38%	108,415	85,205	6.02	4.15
Total	6,367	100.00%	68,444	52,000	6.07	4.80

Table 1.6: Average prices by regions

1.3.2 Transaction types in space

In this section, the distribution of different transaction types, as identified in section 1.1.2, will be illustrated. The distribution of *dVFA* transactions is not homogeneous throughout Luxembourg (fig. 1.6). Compared to the total number of transactions, on the one hand the part of *dVFA* is observed highest around the capital and the southern part of the country, with a large amount of municipalities with over 25% of transactions of this type. On the other hand, the 14 municipalities without transactions of type *dVFA* or with a low ratio are mostly located in the north and hence in more remote municipalities. From map B (fig. 1.6), significant clusters of high values are identified in the municipality of Luxembourg as well as the northern neighbouring municipalities. A second high-high cluster has been identified in the south-west, including the municipalities

[19] Identified in appendix B.

[20] The increased part of transactions located in these areas is related to the relative size of this regions.

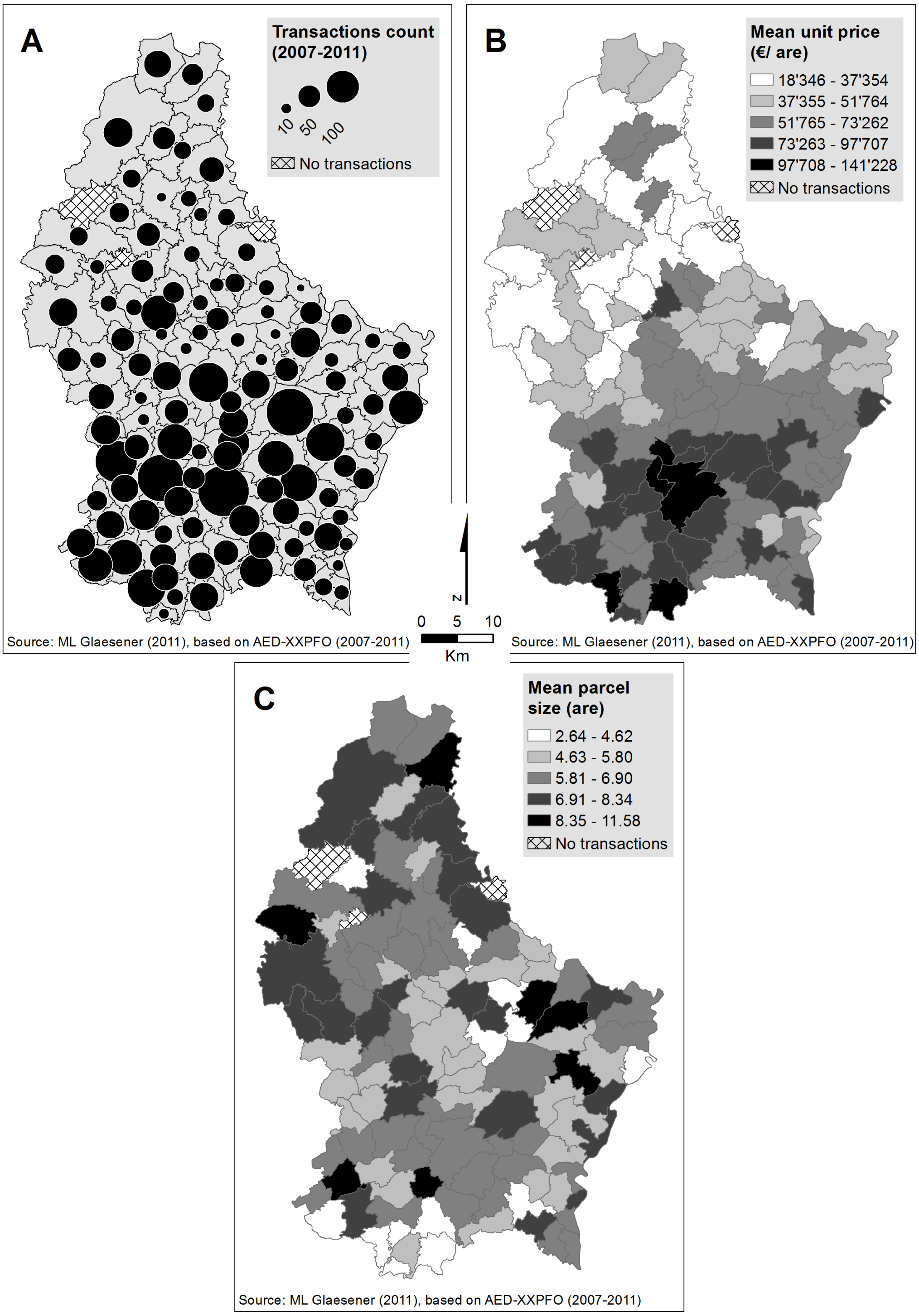


Figure 1.4: Count of observations, mean unit price and average parcel size

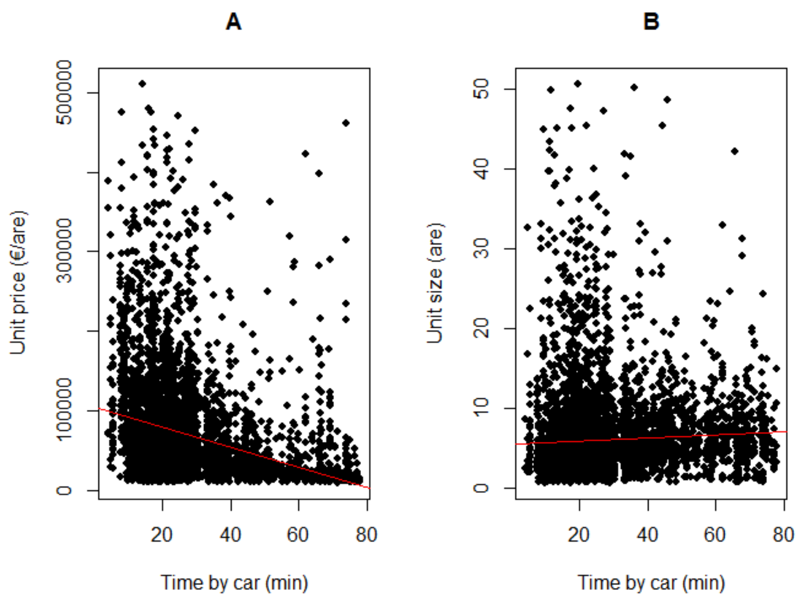


Figure 1.5: Unit price and size vs. travel-time to Luxembourg

of Dudelange, Bettembourg, Mondercange and Schifflange. The amount of transactions sold with development projects is most important in the denser urban areas.

The distribution of *dVFA* transactions through the different morphological and functional areas is quite heterogeneous, as observed in table 1.7. Almost 30% of the transactions located in the agglomeration of Luxembourg are sold with existing development plans, in the southern agglomeration and the close periurban area, 20% of all transactions are of this type while the proportion is slightly higher in the distant periurban area. The rural area seems least attractive for transactions of this type. 41% of all *dVFA* transactions have been registered in the periurban area. The urban denser areas and especially the periurban area hence appear particularly attractive for real estate developers. In the distant periurban area, the average size of *dVFA* transactions is even smaller than in the close periurban area, although their size is generally on average bigger in these sections. These transactions are assumed to be sold to individuals by real estate developers (RED), and they tend to be located close to Luxembourg-city. These observations might be explained by time, the *dVFA* in the distant periurban area might be more recent and reply to the requirements of recent planning policies and modifications in the market.

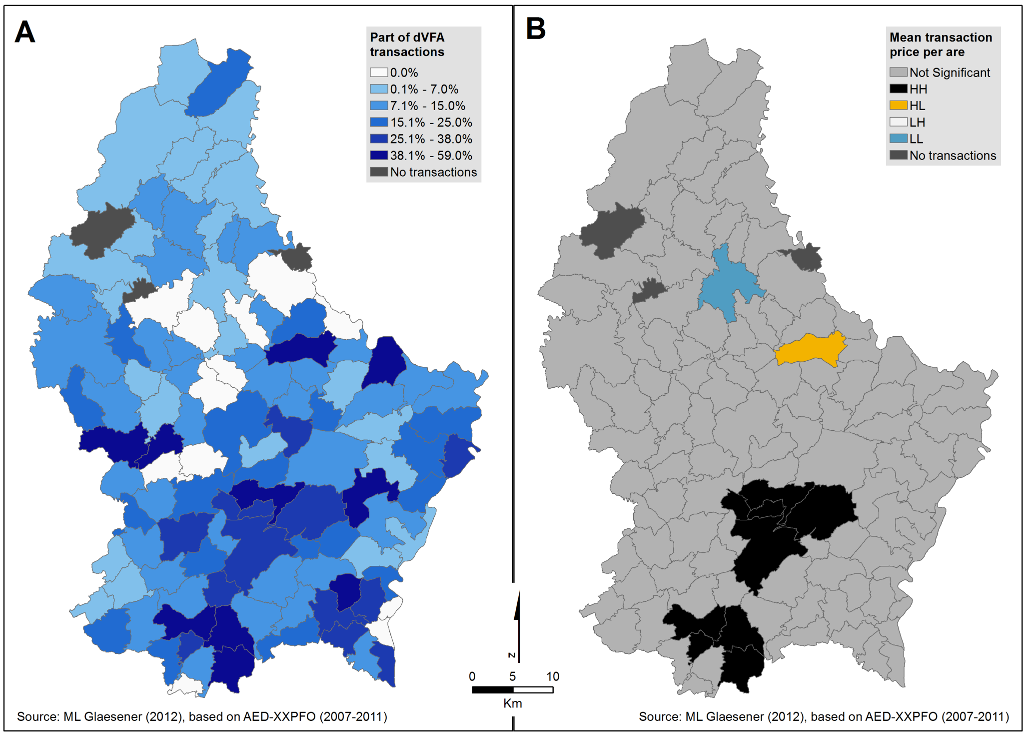


Figure 1.6: Part of *dVFA* transactions

	Count	Part of total	Price/are	Mean Size
Rural	98	8.74%	49,708	5.12
Distant periurban	327	21.02%	78,980	4.01
Close periurban	531	20.37%	87,124	4.27
Agglo South	86	20.38%	115,936	3.15
Agglo Luxembourg	197	29.80%	116,106	3.78
Total	1,239	19.46%	88,623	4.11

Table 1.7: *dVFA* transactions by region

Regarding the large size development projects identified in section 1.1.2, map A (fig. 1.7) shows above mean ratios of *BigSize* transactions in isolated municipalities all over the country. If one compares this map to the national road and railway network, it appears that especially good connected but remote municipalities seem to register such projects. From table 1.8 we learn that the unit price for these transactions is below the overall means (table 1.6), which was expected since below 27.77 are the unit prices are in general decreasing with size.

	Count	Part of total	Price/are	Mean Size
Rural	11	0.98%	18,954	30.66
Distant periurban	34	2.19%	27,603	29.92
Close periurban	52	1.99%	36,404	31.42
Agglo Esch	11	2.61%	31,185	30.17
Agglo Luxembourg	20	3.03%	40,005	33.36
Total	128	2.01%	32,681	31.15

Table 1.8: *BigSize* transactions by region

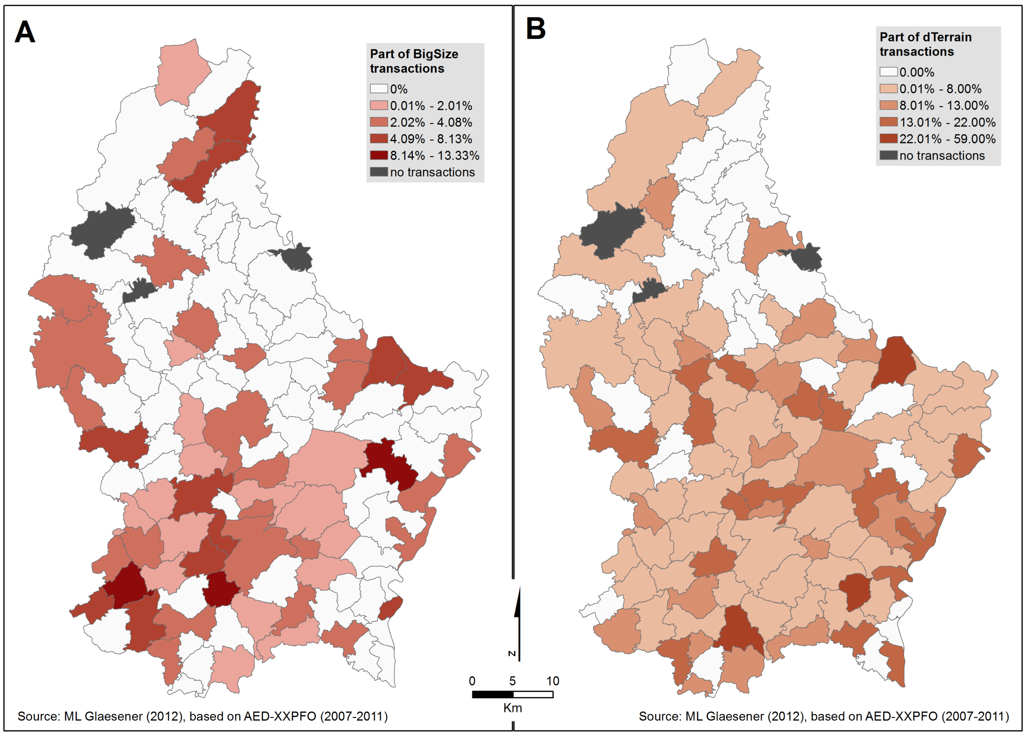


Figure 1.7: Part of *dTerrain* and *BigSize* transactions

Map B fig. 1.7, shows lowest ratios of *dTerrain* transactions in the northern part of the country. With regard to the national average of 8% of the transactions, very high values are observed in the municipalities of Berdorf, Waldbredimus and Bettembourg,

	Count	Part of total	Price/are	Mean Size
Rural	44	3.93%	33,483	5.07
Distant periurban	155	9.96%	48,762	4.04
Close periurban	194	7.44%	55,070	4.23
Agglo South	61	14.45%	63,457	2.98
Agglo Luxembourg	62	9.38%	84,829	4.08
Total	516	8.10%	55,902	4.08

Table 1.9: *dTerrain* transactions by region

observations that would request an in depth analysis of the *dTerrain* variable, that is however beyond the scope of this thesis. According to table 1.9, the agglomerated south registers the highest part of *dTerrain* transactions, while in the rural area less than 4% of transactions are of this type. The average unit price of these transactions is, with 55,902€/are, below the national mean. In addition, parcel size of these transactions are smaller than the national mean, in particular in the Agglomerated South.

Concluding remarks

A major question raised in this chapter is on the role of different types of agents on the real estate market. Given the limitations related to the AED dataset some assumptions on the consumers and sellers of the developable land transactions were made in section 1.1.2. Different legal types of transactions are presented and we expect that they allow, at least to some extent, to control for the marginal effects on prices if different actors are involved (e.g.: professional, public or private consumers).

A brief description of the evolution of transactions in time and their distribution throughout the country was presented in sections 1.2 and 1.3. After this presentation of the structural and transaction specific variables, and since the objective of this thesis is to identify the geographical determinants of land price, chapter 2 focusses on the quantification of the location-specific transaction context. Besides accessibility measures to Luxembourg-city, local urban amenities and land-use variables will be presented.

Chapter 2

Location specific attributes

Besides the structural characteristics of developable land, consumers obtain utility by the neighbourhood specific characteristics of the location in which land is located. The basic urban model (Alonso, 1964; Fujita, 1989) postulates the importance of good access to employment as a main determinant of land prices, beside its size. As an extension to this model, the emergence of the periurban area is explained by Cavailhès et al. (2004) as the result of consumers' preferences for local urban and green amenities. A wide variety of hedonic studies accounting for the local geographical context have been undertaken since Cheshire and Sheppard (1995). Who showed that including both the accessibility aspect of a location as well as a range of neighbourhood characteristics is primary to getting reliable and stable estimates of either.

In this perspective and with regard to the main objective of this thesis, to identify the geographical determinants of residential land prices in Luxembourg, this chapter aims at presenting the local amenities considered in the hedonic models of part II. As presented in figure 2 (p.21), a wide variety of different sources were considered, beside the AED dataset presented in chapter 1. Access to Luxembourg-city was quantified in section 2.1 and local urban amenities in section 2.2. Further, the quantification of green amenities is discussed in section (2.3). Eventually, the variables considered to control for the impact of the socio-economic context on land prices is provided in section 2.4.

2.1 Access to Luxembourg-city...

Similar to the observations made worldwide, motorised individual transport modes have become very prominent in Luxembourg since the middle of the last century. With 654 vehicles per 1,000 inhabitants in 2003, the degree of motorisation is one of the highest in Europe (Petit, 2007, p.2). For 2012, Hansen (2012) identified an increase of this ratio to

666 vehicles per 1,000 inhabitants. This underlines the continuously growing importance of individual transport in Luxembourg. Further, Epstein (2010) confirmed a rise in the average travelling speed since the 1960s in Luxembourg. While at the same time an increased modal part of individual transport is observed, linked partly to the evolution of the road network and the highway network in particular since the 1980s. Considering population growth and travel-time to Luxembourg-city, Epstein (2010) further observed an increased population growth in the municipalities located in a 30 minutes radius around the capital[1]. Moreover, the area covered by road infrastructure has increased by 0.7% between 2000 and 2007 (Klein, 2010).

Based on the data available in this thesis, road infrastructure covers an area of 81 km^2, while approximately 2 km^2 are covered by railway infrastructure (figure 2.7 p. 61), together these two uses cover 3.22% of the country. In map A (fig. 2.1) the main road infrastructure (highway and national roads) are illustrated, while the public transport network (bus and train stations and railway network) are presented in map B. The highway network is concentrated almost in a radio-concentric manner around Luxembourg-city, supporting the assumption of a monocentric organisation around the capital and indicating further a good accessibility of the neighbouring countries. In the north, the coverage in national roads is still relatively dense, but except for the highway sections linking the periurban area to the *Nordstad*, it appears disconnected from the southern part of the country.

With regard to the motorised transport modes (individual and public transport) a modal split of 85.5% is found for individual transport in 2009 (Ministère du Développement durable et des Infrastructures, 2012). This share indicates the relative importance of car use compared to public transport in Luxembourg. In the evaluation of the *IVL* (CEPS/INSTEAD, 2008), it was highlighted that only the southern part of the country is covered in an efficient way. Rural municipalities are often weakly covered by the public transport network and hence lack accessibility to the rest of the country by this means.

The public transport infrastructure dataset (map B fig.2.1) relies on the land registry

[1]Some observations on the evolution of the population will be detailed in section 2.4.

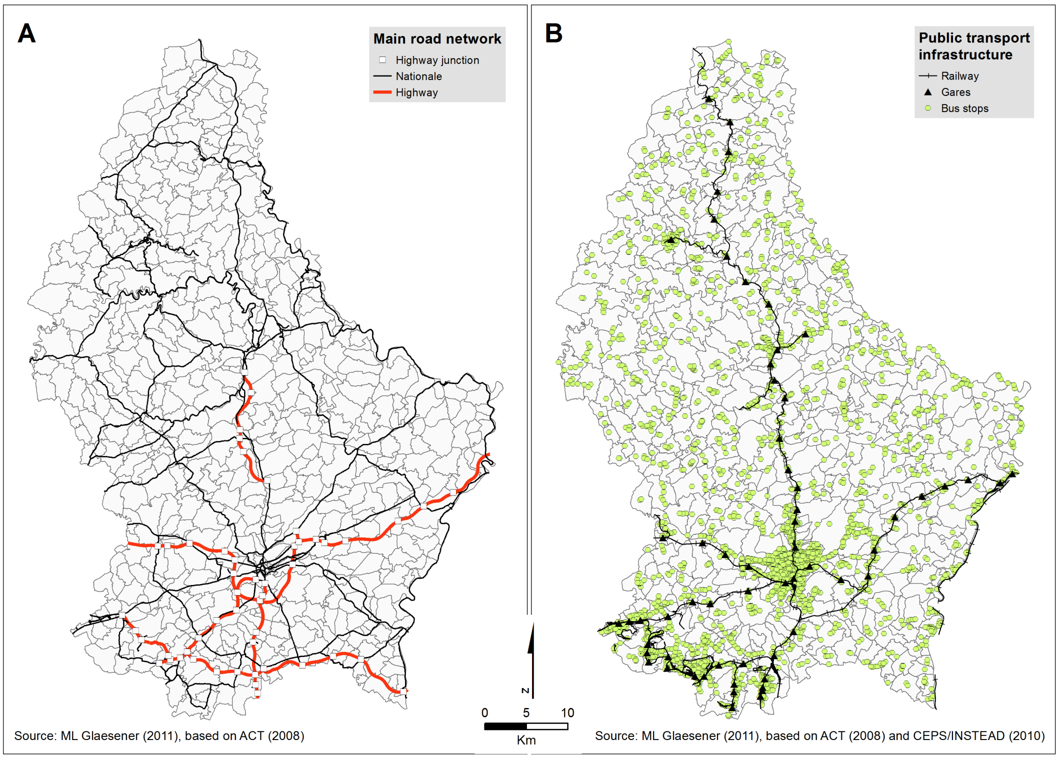


Figure 2.1: Transport infrastructure in Luxembourg

database (Administration du Cadastre et de la Topographie, 2008a) and was combined to the online information of the *Chemin de fer du Luxembourg* (CFL) for the train station dataset; the bus-stops were provided by CEPS/INSTEAD. According to Klein (2010), in 2007, the public transport network in Luxembourg is composed of 236 lines (bus and rail). 95% of the urban centres are at less than 500 m to a bus stop or railway station; with a total of 2,150 stops.

Only 8% of the sections have direct access to the highway network, while 12% have access to the rail network via a train station. However, with regard to access to the bus network, the sections are much better covered, with on average of almost five bus stops per section with 92% of the section registering at least one stop. With regard to this observation it is important to highlight that the frequency of the different bus lines varies from line to line and the sole presence of a bus stop does not allow any conclusions on the availability of public transport. The different local infrastructure information (e.g.: Highway junctions, bus stops) was not further considered in the analysis as fine scale measures were not possible to generate.

In hedonic pricing literature, including access to the main employment centre is most common. As highlighted in Sirmans et al. (2005) review on the composition of hedonic pricing models (125 studies), distance to CBD is one of the most frequently included characteristics. In general, distance to CBD is found to have a negative impact on price, hence consumers benefit from being located close to the main employment centre and are willing to pay more for shorter commutes. Some recent hedonic studies accounting for distance to CBD can, among others, be found by Ahlfeldt (2011); Koschinsky and Lozano-Gracia (2012); Melchiar and Kaprová (2013).

LeRoy and Sonselie (1983) highlighted that distinction should be made between travel distance and time. Especially since the technological improvements on infrastructure and cars themselves have significantly increased travel speed in the last decades and have thus reduced transport costs with regard to time. The decrease in travel-time due to the omnipresence of cars was discussed among others in Glaeser and Kahn (2004). Hence, rather than considering distance to the CBD one should account for travel time, as undertaken for instance by Thériault et al. (1999); Kestens et al. (2006); De Bruyne and Van Hove (2013) in the hedonic context. Further, the impact of different transport modes on the distribution of individuals has been discussed among others by Glaeser (2008).

With regard to the importance of Luxembourg-city as a destination for job related commuting flows, time-distance to the capital should be tested for in order to measure the marginal impact of increased distance on land values.

2.1.1 ... by road network

Travel time and distance by car were computed using a Google Maps® application developed by (Medard de Chardon and Caruso, 2010). Generalising the itinerary function of Google Maps, it allows the input and output of coordinates for a large number of origins and destinations. The section centroids served as origin and the central train station of Luxembourg-city as destination.

On average, it takes 37 minutes from a section to Luxembourg central station, average distance being 34km (table 2.1). Comparing Euclid distance to distance by car confirms the importance of considering road network, distance as-the-crow-flies indicating on average much shorter commutes.

And as can be read from table 2.2, close periurban sections are located on average in a 22 minutes distance to Luxembourg-city, similar to what is observed for the sections of the Agglomerated South. In map A figure 2.2 the cartographic illustration of time-distance is provided at section scale, showing as expected increasing travel-time with distance to Luxembourg-city.

Eventually, travel-time by car was kept in the hedonic model to account for the travel-time cost consumers face when commuting from the section in which the parcel is located to Luxembourg-city. A limitation to this data is that it does not account for variations in commuting-time related to traffic congestion, frequent at rush hour.

Variable	Unit	Mean	Min	Max
Euclid distance	km	25.38	0.66	63.75
Distance by car	km	34.40	0.79	77.71
Time-distance by car	minutes	36.52	2.72	80.47
Time-distance by public transport	minutes	50.76	6.00	173.00

Table 2.1: Travel-time from sections to Luxembourg-city

2.1.2 ... by public transport

Travel-time by public transport was generated via “Mobiliteit.lu®”[2]. The bus-stops or train station closest to section centroid served as origin and the central train station of Luxembourg as destination. A request was send to the website to obtain the time by public transport from origin to Luxembourg-city on a Tuesday morning with expected arrival time around 8.30 a.m.[3]. The aim was to capture the average commuting time by public transport at section level, found to be on average of 50.76 minutes (table 2.1).

[2]Thanks to Anissa Zerarga for her precious help in generating this dataset at the section level.

[3]“Mobiliteit.lu” provided generally five alternative journeys and from these we selected the fastest one.

As illustrated in map B (fig. 2.2), travel-time by public transport shows a more heterogeneous pattern compared to time-distance by car. Furthermore, substantial differences can be observed between the morphological and functional regions (table 2.2). As expected, the agglomeration of Luxembourg, the close periurban and the Agglomerated South show below average means, while in the rural area the claims of CEPS/INSTEAD (2008) regarding the poor connection of the northern sections to the capital are confirmed.

2.1.3 Comparing transport modes

Time to Luxembourg-city by individual (map A) and public (map B) are illustrated in figure 2.2, and as expected they are structured along the the transport infrastructure (fig. 2.1). At national scale, it is observed that mean travel-time to Luxembourg-city increases from 37 to 51 minutes if using public transport. Table 2.2 illustrates a similar trend for both modes with travel-time increasing from the urban to the rural areas, while the averages for the close periurban and the Agglomerated South rather the same.

	Time to Luxembourg by...	
	... car	... public transport
Rural	55.87	69.57
Distant periurban	34.74	52.31
Close periurban	21.85	35.21
Agglo South	21.04	38.70
Agglo Luxembourg	11.02	20.25
Total	36.52	50.76

Table 2.2: Travel-time to Luxembourg-city by region and mode

To compare both travel modes, figure 2.3 illustrates the efficiency of public transport over individual transport at section scale. A value above 1 suggests that using public transport is faster than relying on individual transport. Compared to map B (fig. 2.1, p.45) it is obvious that this pattern is governed by the transport infrastructure. In the north, in the sections located close to the railway network, both transport modes are competitive with regard to travel-time, with public transport sometimes even being faster. While in the south the better coverage in highways is assumed the major reason

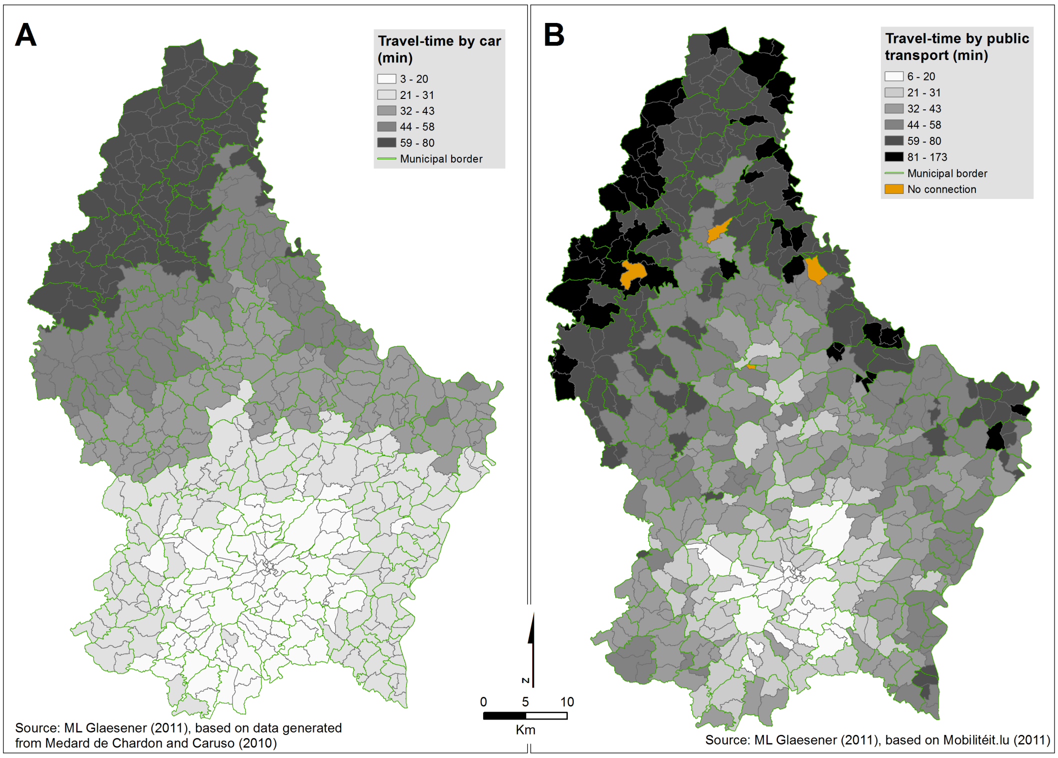


Figure 2.2: Time-distance to Luxembourg-city

for the car prevailing as a faster transport mode in almost all sections, whether connected to railways or not[4].

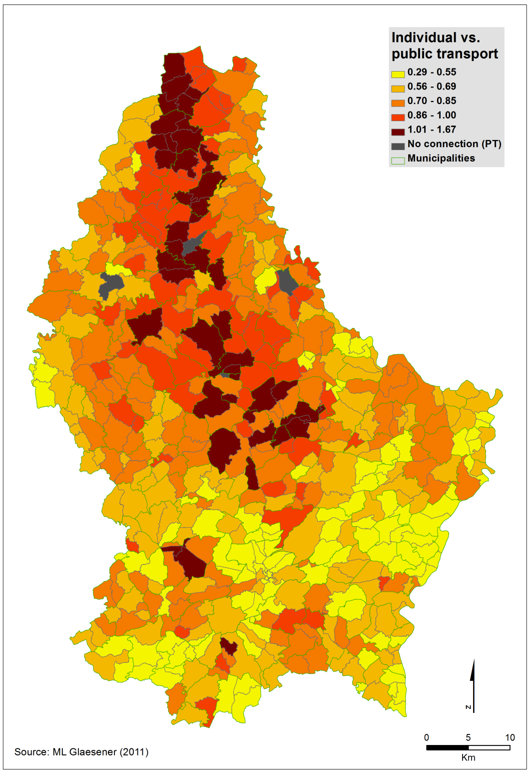


Figure 2.3: Car vs. Public transport

Expected results

Longer commuting time to Luxembourg-city is expected to have a negative marginal impact on land prices. Whatever transport mode considered, consumers are thus expected to value increasing distance from the employment centre negatively (table 2.3). Single distance to the CBD is not sufficient to asses the effect of urban spatial expansion on land prices, for instance accessibility to local public amenities should also be tested (Geniaux and Napoléone, 2008). The aggregated scale of the transaction dataset did

[4]These maps should however be analysed with caution, since the datasets for public and car transport were not computed at the same dates (between 2010 and 2012) and do not necessarily translate the reality in 2007. It is assumed that the evolution of road infrastructure is slower than for public transport network, as frequencies of public transport can easily be increased, while building new roads is a long process (i.e. *Nordstroos*).

however not allow the generation such fine scale accessibility measures.

Variable	Unit	Description	Impact
$tLUX_{ci}$	min	time to Luxembourg by car	-
$tLUX_{pi}$	min	time to Luxembourg by public transport	-

Table 2.3: Expected impact of time-distance to Luxembourg-city

2.2 Diversity of local urban amenities

The benefits individuals obtain from local urban amenities have been discussed in theoretical urban economic literature for some time (Tiebout, 1956; Brueckner et al., 1999; Glaeser et al., 2001), showing that a wide offer and variety of local public goods promote economic and urban growth[5]. Brueckner et al. (1999) show that the spatial pattern of urban amenities in a city impacts on the location of different income groups and that the valuation of these amenities is rising rapidly with income. These local urban amenities are highlighted by Cavailhès et al. (2004) to promote the rise of the periurban belt. While, on the one hand, large urban markets may increase the welfare of consumers by scale economies of some goods (Glaeser et al., 2001); on the other hand consumers have preferences for a wide offer of basic products (e.g.: groceries) and public services (e.g.: schools, sports facilities) at local scale. Glaeser et al. (2001) highlight four particular critical urban amenities in the context of explaining urban growth and the attractiveness of cities. First, the importance of a rich variety of services and consumer goods is put forward. Second, the aesthetics (e.g.: architecture) and physical characteristics (e.g.: climate) of a city may add to its attractiveness. Third, the quality of public services (e.g.: schools, parks) is pointed out as important to attract new (educated) residents. Eventually, Glaeser et al. (2001) add speed of access to the above mentioned amenities as an urban amenity itself. Related to the general benefits consumers obtain from shorter commutes, Fujita and Thisse (2013) claim that consumers prefer as well shorter trips to shopping opportunities. In general, urban amenities are considered to increase individuals' quality-of-life and hence the presence and proximity to such goods is assumed to be considered when purchasing developable land. This has been discussed in the hedonic context mostly by Cavailhès (2005); Des Rosiers and Thériault (2006) and more recently

[5]As Brueckner et al. (1999), we do not account for the potential reverse causation that might exist between the location choice of consumers and the offer of urban amenities, the so called *Tiebout bias*. This bias arises because of the endogeneity of the household's location when estimating the demand for local public goods (Hoyt and Rosenthal, 1997, p.160). They point out the fact that either consumers are attracted by a location in space because of local urban amenities, the presence of such amenities might be as well the consequence of the arrival of new residents, locating where they expect to reach the most and well-off consumers.

by Öner (2013). Thériault et al. (2005) showed that access to schools and to health-care as well as to restaurants exerts a strong positive impact on prices. In the specific context of Quebec-city they found that good access and choice of these urban amenities is dearer to households than proximity to the CBD (Thériault et al., 2005). Further, diversity of retail and service opportunities have been found to have a positive impact on individuals' satisfaction (Youssoufi, 2011) and could be considered as an indicator for urban centrality (Öner, 2013).

Considering diversity of local urban amenities explicitly is less frequent in the hedonic pricing context. A positive marginal impact on residential land prices is expected for an increased diverse offer of local retail and (public) service opportunities. In this perspective, we seek to identify and quantify the different retail and (public) service opportunities in section 2.2.1 and a diversity index of these is presented in section 2.2.2.

2.2.1 Retail and (public) services in Luxembourg

An extensive database regrouping a selection of basic retail (bakeries, butchers, supermarkets, book-stores and press and tobacco retailers) and service opportunities (bars, restaurants, hairdressers and service stations for car repair) has been generated (table 2.4). Some reflections on retail and settlement development in Luxembourg has been provided by Affolderbach and Becker (2011), pointing out that in 2007 the average sales area per resident was of 1.9m^2, largely above the averages observed in France and Germany (Affolderbach and Becker, 2011). Decoville (2012) identified different retail structures in Luxembourg, with the offer and types of retail varying between the different regions of the country. With the urban centres showing a rather divers offer of small retail, while in the periurban area, located close to the major transport network, larger sales areas in the form of malls. A third type of retail implantation was identified, characterised by a rather remote location in the rural areas, identified as specific to Luxembourg as mainly targeting international costumers (Decoville, 2012). Further, health infrastructure has been represented by three main categories: general practitioners, pharmacies and hospitals (see map B fig. C.2 in appendix C.2). As was shown in several hedonic studies school quality plays an important role in residential location choice (Thériault et al., 1999; Uyar and Brown, 2007; Clapp et al., 2008; Kiel and Zabel, 2008). Especially with regard to the socio-economic composition of periurban households, often associated to the presence of children, school quality is expected to have a strong impact on income sorting of households. Since school quality information was not available, only their location within sections could be considered. Education infrastructure is mainly organised by the Ministry of Education, its distribution is quite homogeneous throughout the country as it is fixed with regard to the respective needs. The distribution of education infrastructure is guided by two criteria: an optimal coverage of primary schools and the concentration of higher level education in urban regional centres (Decoville, 2012). In addition, day-care services for early childhood[6] and for children in primary school ("crèches" and "maison relais") have been considered. Additional public services have been taken into account such as the location of town-houses and postal offices[7]. Especially ,the town-house emphasises the hierarchical structure among sections within a same municipality, which might be considered as a premium for consumers as it implies shorter ways for administrative procedures.

A limitation related to this dataset is that the different observations based on the online phone register (editus.lu) do not describe the situation in 2007, but rather the reality of 2011. It is assumed that during the reference period the amount of amenities did not substantially change and that the variations occurred homogeneously through space.

To approximate the general offer in local urban amenities, the total number of amenities (*SSopp*) was considered within the hedonic model. *SSopp* thus presents the count of all different types of amenities within one section, the marginal effect on price is assumed to be positive, consumers should obtain utility by an increased offer in retail and service opportunities. As illustrated in figure 2.4, the distribution of the 4,777 opportunities throughout the country is not homogeneous. Concentrations can be observed within the

[6]A discussion on the spatial distribution of the early child day-care infrastructure can be found in Leduc (2009).

[7]Decoville (2012) also include police stations, employment agencies, to measure the centrality index of public services at municipal scale.

	Rural	Distant periurban	Close periurban	Agglo South	Agglo Lux.	Total
Bakery	35	42	31	35	56	199
Butcher	27	30	31	20	33	141
Hairdresser	62	101	77	76	144	460
Restaurant	182	243	242	116	479	1,262
Pub & Bar	127	149	113	98	155	642
Service station	56	71	80	36	87	330
Book store	10	8	12	13	35	78
Press & Tobacco	9	10	10	6	31	66
Supermarket	45	55	58	26	71	255
General practitioner	67	88	79	72	200	506
Pharmacy	15	19	16	15	26	91
Town-house	38	29	38	4	7	116
Postal Office	25	26	30	8	24	113
Recycling centre	3	5	8	2	4	22
High-school	11	6	1	7	25	50
Primary school	28	35	44	18	28	153
Daycare	22	36	91	28	116	293
SSopp	762	953	961	580	1,521	4,777
Diversity index	0.42	0.53	0.54	0.84	0.77	0.53

Table 2.4: Number of retail and service opportunities by region

sections of the agglomeration of Luxembourg as well as the Agglomerated South, in the rest of the country, increased amounts of opportunities are mainly located within the regional urban centres (e.g.: Nordstad, Echternach, Wiltz,...).

All these variables, aiming at quantifying the offer of local urban amenities, have been geo-referenced and tested within the hedonic model. However, they present high correlations among themselves[8]. Since accessibility measures to different amenities from residential land transactions are impossible to compute and to get around the problem of multicollinearity, a diversity index was generated, in order to account for individuals' preferences for a diverse offer of urban amenities.

2.2.2 Diversity index for local urban amenities

In order to capture the marginal effect on land values of an increased diversity in retail and (public) services, a diversity index (*DI*) was generated based on the work of

[8]Illustrated in table C.7 and in space as observed by comparing the different maps in figures C.2 and C.3 in appendix C.2.

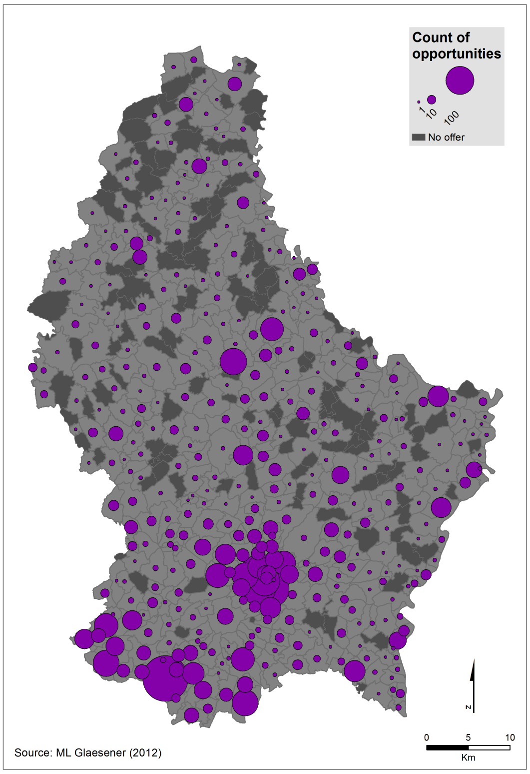


Figure 2.4: Count of urban amenities

Youssoufi (2011), taking the following form:

$$\mathrm{DI} = 1 - \sum_c \left(\frac{n_c}{n}\right)^2 \tag{2.1}$$

Where n_c designates the quantity of observations for the c^{th} category and n the total amount of shopping and service amenities per section. In a second step, different weights were attributes to some of these opportunities according to the frequency of their use. Distinction was made between amenities considered of daily/weekly use (bakery, butcher, supermarket, press and tobacco, high-school and primary school, childcare) and monthly/multi-year use (supermarket, hairdresser, restaurant, pub or bar, garage, postal office, book store, general practitioner, pharmacy, town-house and recycling centre). Supermarkets were included twice, since the frequency is related to weekly provisions and there might be additional, less regular, visits for some other products (Youssoufi, 2011).

Section ID	Section	*n*	*n* weigthed *(nw)*	Total populaton in 2007	CS/1000hab	Diversity Index *(DI)*	Weighted Diversity Index *(WDI)*	Bakery	Butcher	Supermarket	Hairdresser	Restaurant	Pub or Bar	Garage	Postal office	Bookstore	Press & Tobacco	General practitionner	Pharmarcy	Highschool	Primary School	Creche / Maison Relais	Townhouse	Recycling Center
11010202	DES MOULINS	1	2	526	1.90	0.000	0.000	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13121004	WELLENSTEIN	1	1	523	1.91	0.000	0.000	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
13110804	EHNEN	6	6	431	13.92	0.278	0.278	0	5	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12050801	ASSELBORN	4	4	381	10.50	0.375	0.375	0	3	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11010102	HAUTCHARAGE	4	6	1019	3.93	0.500	0.444	0	0	0	0	0	2	0	0	0	0	0	0	0	0	2	0	0
11010203	FINGIG	2	3	203	9.85	0.500	0.444	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0
13120901	WALDBREDIMUS	2	3	387	5.17	0.500	0.444	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
13120701	WINTRANGE	2	4	401	4.99	0.500	0.500	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0
13121001	KLEINMACHER	2	2	121	16.53	0.500	0.500	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0
13120801	STADTBREDIMUS	5	6	789	6.34	0.560	0.611	0	0	0	0	3	0	0	0	0	0	0	0	0	0	1	1	0
13120704	SCHENGEN	4	4	539	7.42	0.625	0.625	0	0	0	0	2	1	1	0	0	0	0	0	0	0	0	0	0
11010304	SCHOUWEILER	13	17	1074	12.10	0.710	0.727	0	0	0	1	6	0	0	0	0	0	1	0	0	1	3	1	0
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
11010302	BETTANGE	4	6	983	4.07	0.750	0.778	0	0	1	0	0	1	1	0	0	0	0	1	0	0	0	0	0
11030102	VILLE HAUTE	225	259	2227	101.03	0.753	0.800	26	102	22	2	2	8	6	14	5	6	23	5	1	0	2	1	0
13120902	TRINTINGERTHAL	6	10	410	14.63	0.778	0.820	0	0	1	0	2	1	0	0	0	0	0	0	0	1	1	0	0
13120602	REMICH	39	44	1738	22.44	0.798	0.827	2	1	0	5	15	4	2	1	0	0	5	1	0	0	2	1	0
13120703	FLOUER	9	13	285	31.58	0.790	0.840	0	1	1	0	3	0	2	0	0	0	0	0	0	0	1	1	0
11010301	DIPPACH	8	11	935	8.56	0.844	0.860	0	0	1	1	2	0	1	1	0	0	1	0	0	0	1	0	0
11010201	CLEMENCY	10	12	1306	7.66	0.860	0.875	0	0	1	0	2	1	1	1	1	0	2	0	0	0	0	1	0
11010103	BASCHARAGE	49	69	5513	8.89	0.865	0.887	2	4	6	5	13	3	3	1	0	1	7	1	0	1	1	1	0
11020401	ESCH-NORD	266	339	23539	11.30	0.878	0.904	40	58	42	12	2	6	6	16	7	12	31	7	5	12	9	1	0
13110502	JUNGLINSTER	38	56	2182	17.42	0.902	0.906	3	5	1	7	2	1	0	2	3	4	2	1	0	1	4	1	1
11030601	SANDWEILER	28	40	2776	10.09	0.903	0.909	2	4	4	0	1	1	1	1	1	3	4	1	0	2	2	1	0
12060401	DIEKIRCH	68	89	4817	14.12	0.891	0.912	6	14	7	6	1	1	2	3	3	5	11	2	2	1	2	1	1
National scale		*4777*	*6123*	*471062*	*28.05*	*0.519*	*0.528*	*460*	*1262*	*642*	*330*	*113*	*78*	*66*	*199*	*141*	*255*	*506*	*91*	*50*	*153*	*293*	*116*	*22*

Figure 2.5: Distribution of amenities per section

Table 2.5 illustrates the distribution of the considered amenities for some sections. Map A (fig. 2.6) illustrates the weighted diversity index at section scale. The distribution of diversity is not homogeneous throughout the country, many sections do not register any opportunities, mainly observed in the northern part of the country. The distribution of the high values is, as expected, in accordance with the urban structure in Luxembourg. With regard to map B (fig. 2.6), two major clusters of high-high values are found around the capital as well as in the south-west. While in the north of the country, except for some punctual high values, corresponding to regional urban centres, mostly low values are observed. The diversity index is close to the national average in periurban sections (table 2.4). These observations are confirmed by the findings of the synthetic index of urban centrality developed by Decoville (2012).

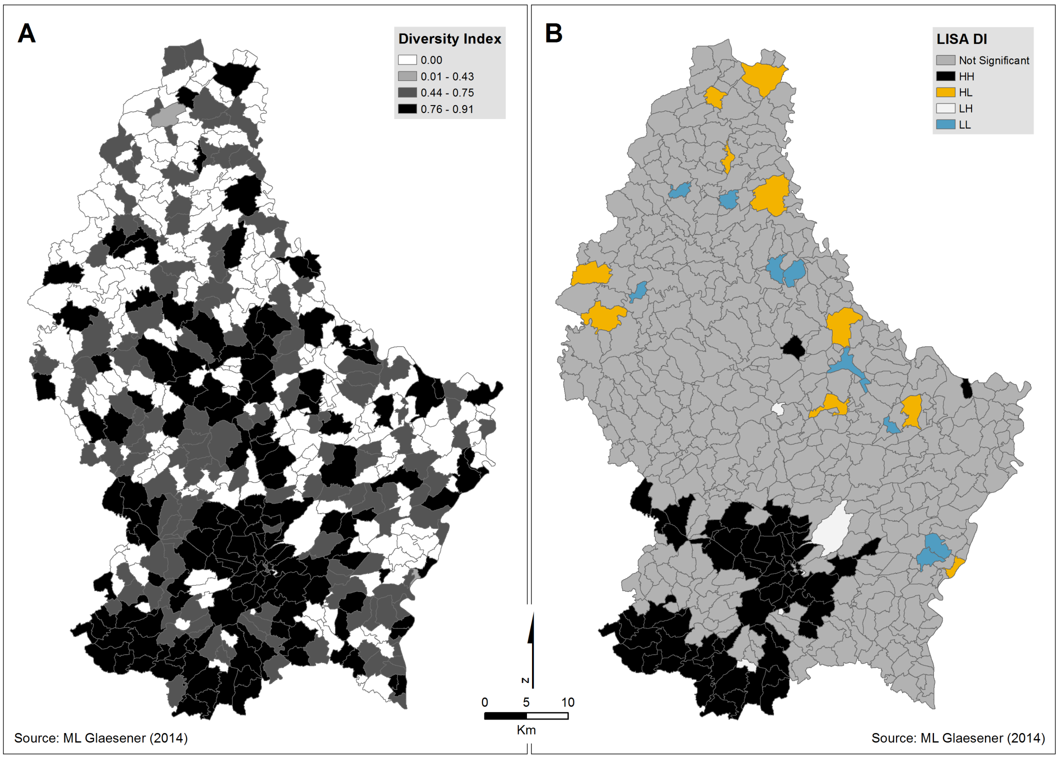


Figure 2.6: Retail and service diversity index in Luxembourg

Expected results

In the impossibility to account for fine scale accessibility measures to local urban amenities, the diversity index of retail and service opportunities (*DI*) was generated to account for a centrality effect within sections. A positive impact on land values is expected from increased diversity. Further, as presented in section 2.2.1, the count of all opportunities (*SSopp*), should capture some urban effect and likewise be valued positively by land consumers.

Variable	Unit	Description	Impact
DI	index	Diversity index of shopping and service opportunity	+
SSopp	count	Shopping and service opportunities	+

Table 2.5: Expected impact of local urban amenities

2.3 Quantifying green amenities

The periurban model shows that the emergence of a periurban belt is linked to the utility land consumers obtain from the presence of rural amenities (Cavailhès et al., 2004). A review of theoretical models from urban economic and geographic literature on the heterogeneity of urban patterns with regard to green amenities[9] can be found in Caruso and Cavailhes (2010), while Bockstael and Irwin (2000) focus on economic modelling of land-use pattern change. Economists are showing thus an increasing interest in land-use and its spatial distribution, as landscape is a decisive factor in people's environment and quality-of-life. However, these amenities are not traded directly in the residential market and thus it is relied on hedonic modelling to estimate their implicit value. The role of local green has been addressed in several hedonic pricing studies; McConnell and Walls (2005) reviewed about 40 studies estimating the value of open-spaces through the hedonic pricing method. More recently, Brander and Koetse (2011) collected over 52 studies addressing the valuation of open-space in their meta-analysis, mainly at micro-scale. In general, the green amenities' impacts on land prices are evaluated positively, but they might as well have negative effects, since they can be "*unsightly, odorous, or insect ridden*" (Geoghegan et al., 2003, p.34). For instance, the role of agriculture is quite mitigated as it generates positive and negative externalities for residential land consumers at the same time (Bockstael and Irwin, 2000). Often distance to different green amenities (Shultz and King, 2001; Des Rosiers et al., 2002; Kestens et al., 2004; Cho and Roberts, 2008) and/or their proportion or presence in varying buffer zones (Cavailhès et al., 2007; Sander et al., 2010; Youssoufi, 2011; Kadish and Netusil, 2012; Melchiar and Kaprová, 2013) are considered, while others estimate the value of viewed landscapes (Cavailhès et al., 2006; Brossard et al., 2008; Cavailhès et al., 2009; Youssoufi, 2011). Distinction has to be made between private and public open-space or between undeveloped and protected and open-space available for development (Brueckner, 1983; Yang and Fujita, 1983; Wu and Plantinga, 2003; Irwin and Bockstael, 2004; Turner, 2005). Whereas in general positive impacts on property prices are found if open-space

[9] In the following chapters green, open-space, "natural" and rural amenities are used as synonyms.

is protected, the proximity to open-spaces available for construction generally has a negative or insignificant impact (Irwin and Bockstael, 2001; Geoghegan, 2002; Geoghegan et al., 2003).

Not only the presence of different types of land-use might determine residential land price, a landscape is always more complex than the sum of its parts and each part has different characteristics on the basis of how it interacts with its surroundings (Finotto, 2011, p.48). The quantification of spatial land-use patterns has been promoted by the works of landscape ecologists (Kong et al., 2007), and in this perspective some diversity measures will be presented in section 2.3.2.

Several hedonic studies consider in this perspective vegetation indices (Kestens et al., 2004; Kadish and Netusil, 2012) including the standard deviation of the normalised difference vegetation index (NDVI). That expresses the diversity of land-use in terms of vegetation cover (to approximate vegetation density). Kadish and Netusil (2012) find for instance that an increased NDVI in 40m buffer around transactions has a positive impact on prices. Geoghegan et al. (1997) used different landscape pattern indices to identify how landscape diversity and fragmentation around parcels are valued and how their marginal impacts vary in spatially. They find that increasing diversity affects property values in two ways: negatively as they introduce higher chances of negative visual and noise externalities, but also positively as diversity may implicitly signify the proximity of important local urban amenities. Their results suggest that an increase in diversity is valued differently by consumers with distance to CBD and that they are generally not a desirable feature in the suburban area (Geoghegan et al., 1997).

2.3.1 Spatial distribution of land-uses

The proportion of different land-uses in Luxembourg is illustrated in figure 2.7, based on the data available from figure C.4 (in appendix C.3 p.221). The share of the country occupied by artificial land cover is quite small compared to the area occupied by the so-called "natural" land cover, which covers more than 85% of the country (fig.2.7). These "natural" land-uses are sub-divided into eight major categories and distinction was made between agricultural land (pastures and crop-land), forestry (forest and brushwood), horticulture (vineyards and gardens) and eventually hydrology (rivers and watersheds).

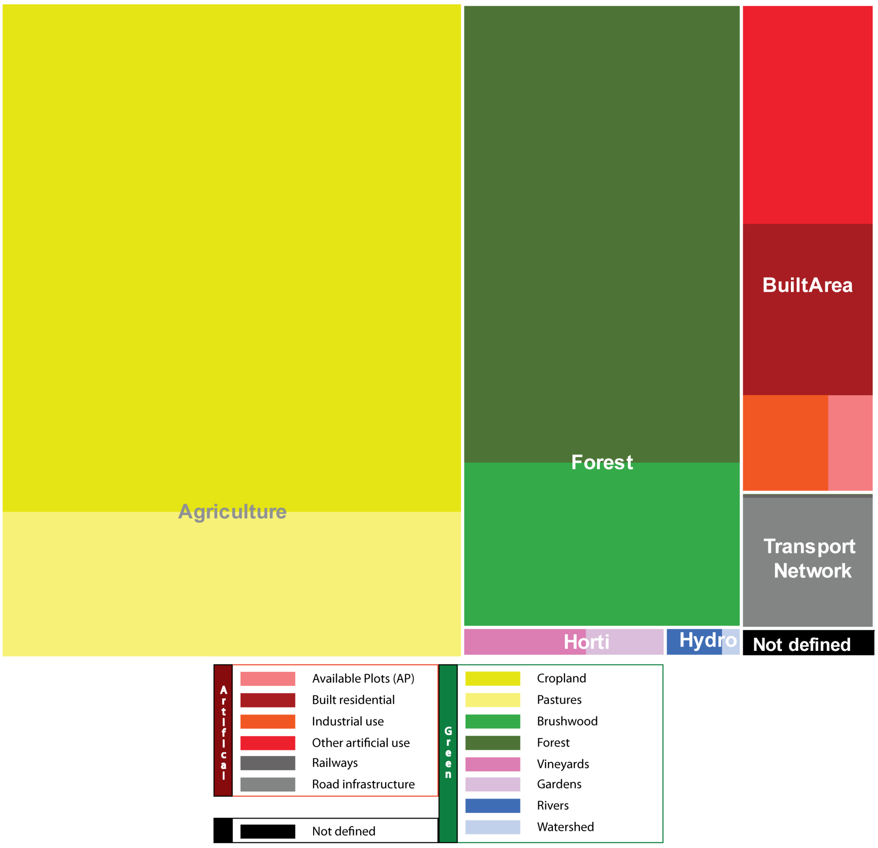


Figure 2.7: Main land-uses in Luxembourg

The valuation of land-use characteristics is generally considered at micro-scale, including proximity measures or considering different land-uses in different buffer zones. However, the computation of such sophisticated neighbourhood statistics around the residential land transactions was not possible. In sections 2.3.1.1 and 2.3.1.2 means to approximate the neighbourhood characteristics of the residential land transactions will be presented, first identifying the theoretical supply of developable land and second the proportions of neighbourhood occupied by different land-uses.

2.3.1.1 Artificial land-uses

Map A figure 2.8 illustrates the part of area per section covered by artificial land-use[10], represented on the right side of figure 2.7. In the northern part the proportion is rather low, while in the southern part sections tend to present high percentages of artificial land. The highest average proportions of artificial land-use are hence observed in the Agglomerated South, followed by the agglomeration of Luxembourg (table 2.6). Within periurban sections the proportion of artificial land-uses is around the national average, suggesting low density urban structure.

The part of artificial land occupied by residential build-up area (BR) (map B fig.2.8) is on average about 41%, except for some sections in the east and north, a large part of the area considered as artificial is occupied by residential use. In the sections of the Agglomerated South, however 63% of the artificial land is covered in other uses than residential which most probably reflects for a large part the former industrial sites (e.g.: brownfields). The sections of the agglomeration of Luxembourg and the close periurban area present above national average proportions of residential land-use. The rural and distant periurban sections present similar proportions than the Agglomerated South, here it is however assumed that the space covered in infrastructure (i.e. road network) is proportionally larger compared to the generally low proportion of artificial land-use.

Particular effort was put in the identification of the parcels available for construction (AP)[11] within the boundaries defined by the "Plan d'Aménagement Général (PAG)" . The AP, classified in 2008, have been considered to gain insights into the theoretical supply in land at section level. Despite the one year difference to the AED dataset, we assume that most of the 2,139 ha available in 2008, were also available in 2007.

The area considered as covered by available plots is a means to approximate the location of our transactions within the sections. The aim was not to identify the actual transactions and their micro-scale location in space, but to be able to account for the land-uses in the neighbourhood of land available for construction and hence approximate

[10] Regrouping: Transport network as well as different residential and industrial land-uses.
[11] A detailed presentation can be found in appendix C.3.

the neighbourhood of the actual transactions more accurately. AED transactions are assumed to belong to the AP. However, the theoretically available residential land is not necessarily available on the market as landowners are not necessarily willing to sell or develop the parcels immediately but might follow some speculative logics and thus wait out before selling. The proportion of AP compared to the total artificial area illustrates the potential of future development of the different sections (map C figure 2.8). The information of AP is only generated for further use in indices and ratios and for the description of the potential supply.

While the total artificial area also covers non residential land use, a ratio of AP compared to all residential land (BR) was generated to identify the part of vacant but available land within a section. On average 19% of the residential area is classified as available for construction in Luxembourg. The distribution of the vacancy rate is rather homogeneous throughout the country (map D fig. 2.8). However, some variation within the municipalities can be observed, which might translate varying municipal planning priorities at section level. In the different regions, only rural sections register a below average rAP rate. Sections of the agglomeration of Luxembourg and the Agglomerated South have on average above mean vacancy rates and the close periurban sections are even slightly above the national average.

The proximity of developable open-space around properties has been found to have in general a negative impact on property prices (Irwin and Bockstael, 2001; Geoghegan, 2002; Geoghegan et al., 2003; Anderson and West, 2006). These open-spaces might reflect positive amenities in the first place (e.g.: undeveloped land) however, consumers might as well anticipate future land conversion to artificial uses and considered as a negative amenity (e.g.: increased density). However, as shown in Singapore housing market by Ooi and Le (2013), infill developments might also have a positive impact on local property values. Land consumers might expect positive effects from future developments in the same section (e.g.: increase in local urban amenities) and at the same time anticipate an increase of the value of their own property.

A similar vacancy rate has been considered by Kiel and Zabel (2008); Liu and Hite

(2013), revealing a negative impact on housing prices. This vacancy rate will be included in the hedonic model to test whether vacant land is perceived rather positively or negatively by developable land consumers in Luxembourg.

	Artificial use	BR	AP	rAP
Rural	8.40%	36.80%	8.26%	17.80%
Distant periurban	13.68%	36.73%	9.27%	19.69%
Close periurban	12.70%	45.61%	11.49%	20.14%
Agglo South	41.75%	36.62%	8.63%	19.33%
Agglo Luxembourg	36.67%	55.19%	12.97%	19.11%
Total	*13.63%*	*40.69%*	*9.82%*	*19.10%*

Table 2.6: Artificial land-use by region

2.3.1.2 "Natural" land-uses and neighbourhood of AP

A main focus of this thesis is to quantify and further estimate the marginal value consumers are willing to pay for "natural"[12] amenities, since individuals are assumed to obtain benefits from green open-space[13] (Cavailhès et al., 2004). First the proportions of land within a section have been considered. However, as was shown in several hedonic studies, open-space is generally perceived at a local scale (Geoghegan et al., 1997; Kestens et al., 2004, 2006; Kong et al., 2007; Goffette-Nagot et al., 2011; Kadish and Netusil, 2012). Although consumers account for these amenities in their location decision, they do most likely not consider an overall availability of certain land-uses at the level of an administrative unit, but in the direct environment of the real estate, especially in the first metres (Cavailhès et al., 2006, 2009).

At section scale, the means of different land-uses around AP were generated to approximate fine scale measures that have been used in the context of address based transaction data in similar studies. Neighbourhood statistics were computed for the cells

[12]It should be highlighted that these "natural" land-uses are mostly not in their natural state, since largely transformed and modelled by agricultural activities.

[13]Maruani and Amit-Cohen (2007) consider open spaces as dominated by natural environment (composed of abiotic and biotic elements) and consider their functions as twofold: open spaces provide recreation and other services to society and conservation of natural values. In addition they mainly represent the ground of agricultural activities, exploiting them as farmland.

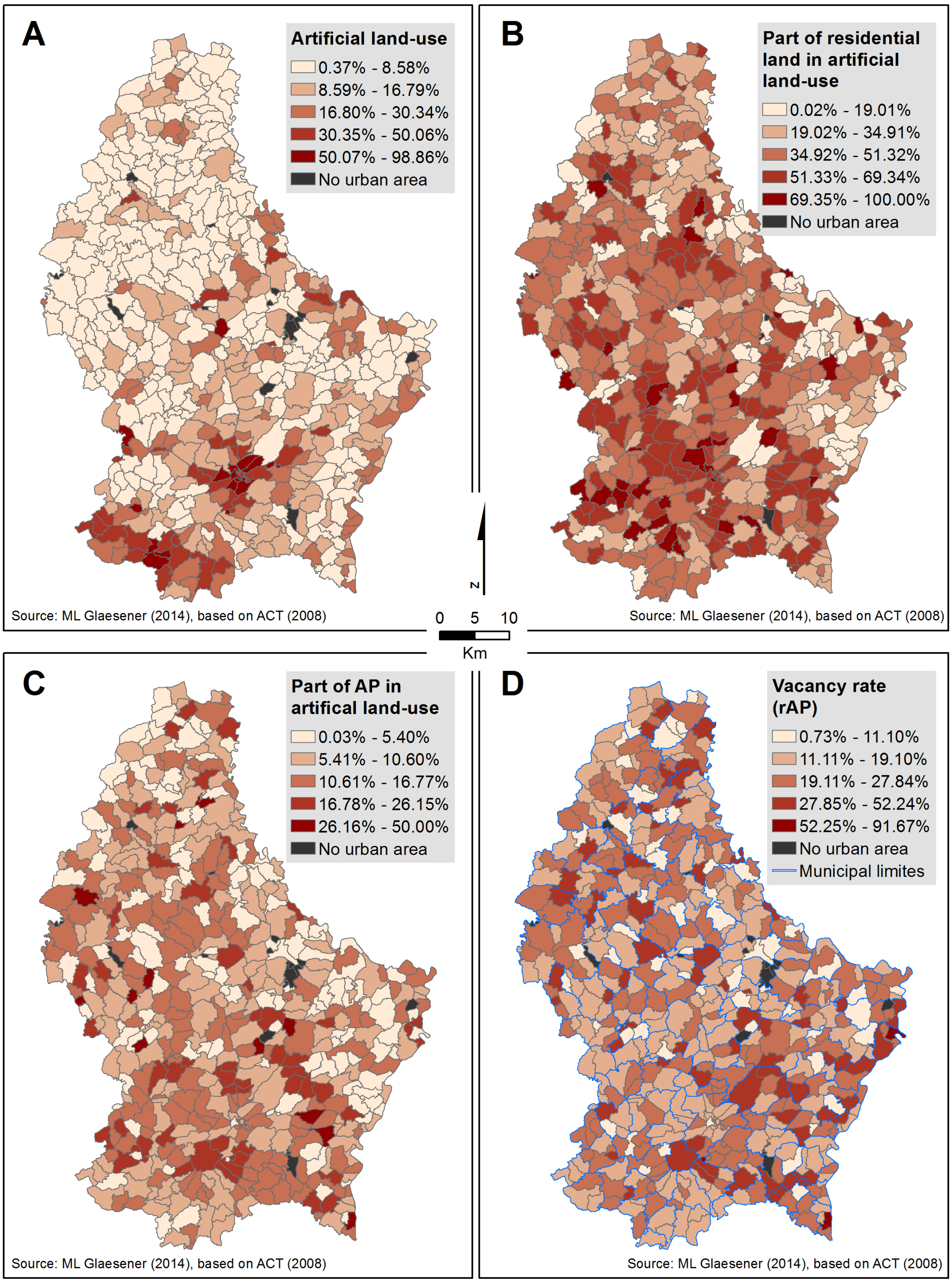


Figure 2.8: Part of artificial land-use

identified in the previous section as available for construction. At different extents focal statistics were generated and the sum of cells covered in a specific land-use around AP were computed[14]. Eventually, the part of the AP neighbourhood (NH) in 100m of land covered by brushwood (*rsap100bw*) and forest (*rsap100for*) (figure 2.9) was considered. In literature mixed results for consumers' preferences with regard to tree cover can be found, depending on the type and proximity of trees, as shown among others by Kadish and Netusil (2012). For instance, Kestens et al. (2004) found negative marginal effects for low tree density and woodlands in a 500 metre radius while the presence of mature trees is found to be valued positively in both, the 100m and 500m extent. Further, the part covered in water flows and horticulture have been considered (fig. 2.10)[15]. A positive effect is expected from the presence of water within close proximity of a property (Kestens et al., 2004).

In figures 2.9 and 2.10 first the ratio of a certain land-use in the 100m neighbourhood of all available plots has been presented in maps A and C, while in maps B and D the total part of the same land-use compared to section area has been presented. As expected, by comparing maps A and B as well as C and D, substantial differences can be noticed. While in some sections the part of a certain land-use is quite important, in a 100m neighbourhood around AP this is not necessarily the case.

The part of area around AP covered in brushwood shows that high values are especially observed in the north of the country, the so-called "Oesling", and especially along the "Sauer" valley, wriggling through the Luxembourgish part of the "Ardennes" mountains (see map H.2 in appendix H). For the rest of the country, in general proportions below 1.4% are observed around AP, confirmed in table 2.7. The proportion of cells covered in forest around AP (C fig. 2.9) is distributed more homogeneously throughout

[14] It was relied on raster calculation tools provided by the toolbox "Spatial Analyst" in ArcGIS10 (2010). Based on the AP raster the area occupied by different land-uses in a defined neighbourhood extent of AP cells were identified. Via "Zonal statistics"this information was further aggregated at section scale; the different are presented in appendix C.3.

[15] Agricultural land-uses were also considered, first making the distinction between pastures and farmland, later they were considered as one land-use (agricultural) since it was assumed that this distinction by the land registry might not be constant in reality as farmers may have changed the type of agricultural use to preserve their fertility.

the country[16]. On average, 4.13% of the 100m neighbourhood around AP are registered as covered in forest, whereas the sections of the agglomerations in general register lower percentages than the distant and close periurban. The lowest values are registered within the rural sections probably due to the more hilly landscapes with the villages rather located on the tops (agricultural organisation of type *open-field*) and hence the available plots as well.

Further, the proportion of the 100m NH of AP was considered in maps A and B in figure 2.10; as expected the distribution of sections representing these amenities follows the main rivers observed in Luxembourg. Hence, very high values are observed in the rural area as well as within the agglomeration of Luxembourg, while the available plots in the periurban and Agglomerated South are much less likely to be located in proximity of water flows.

Eventually, the part of AP neighbourhoods covered in horticultural land-use are considered, where substantial differences can be observed locally. Since this variable considered both vineyards and gardens, almost all sections are concerned, but it is mostly the "Moselle" region that stands out and hence explains the high average ratio for the distant periurban area. Land used for horticultural activities (vineyards and gardens) are rather perceived as positive amenities by land consumers. We assume that less negative side-effects are associated with this form of working the land compared to agricultural activities.

	rap0100bw	rap0100for	rap0100riv	rap100horti
Rural	3.05%	3.28%	25.05%	1.70%
Distant periurban	1.16%	5.41%	6.93%	4.62%
Close periurban	0.06%	4.15%	2.25%	2.14%
Agglo South	0.09%	3.31%	0.00%	1.05%
Agglo Luxembourg	0.11%	3.82%	21.36%	2.54%
Total	1.40%	4.13%	12.76%	2.62%

Table 2.7: Proxies for the 100m NH by region (NH100)

A drawback of this research is that it is not able to account for the land-uses in

[16] Very few significant clusters of spatial concentration of high values have been identified.

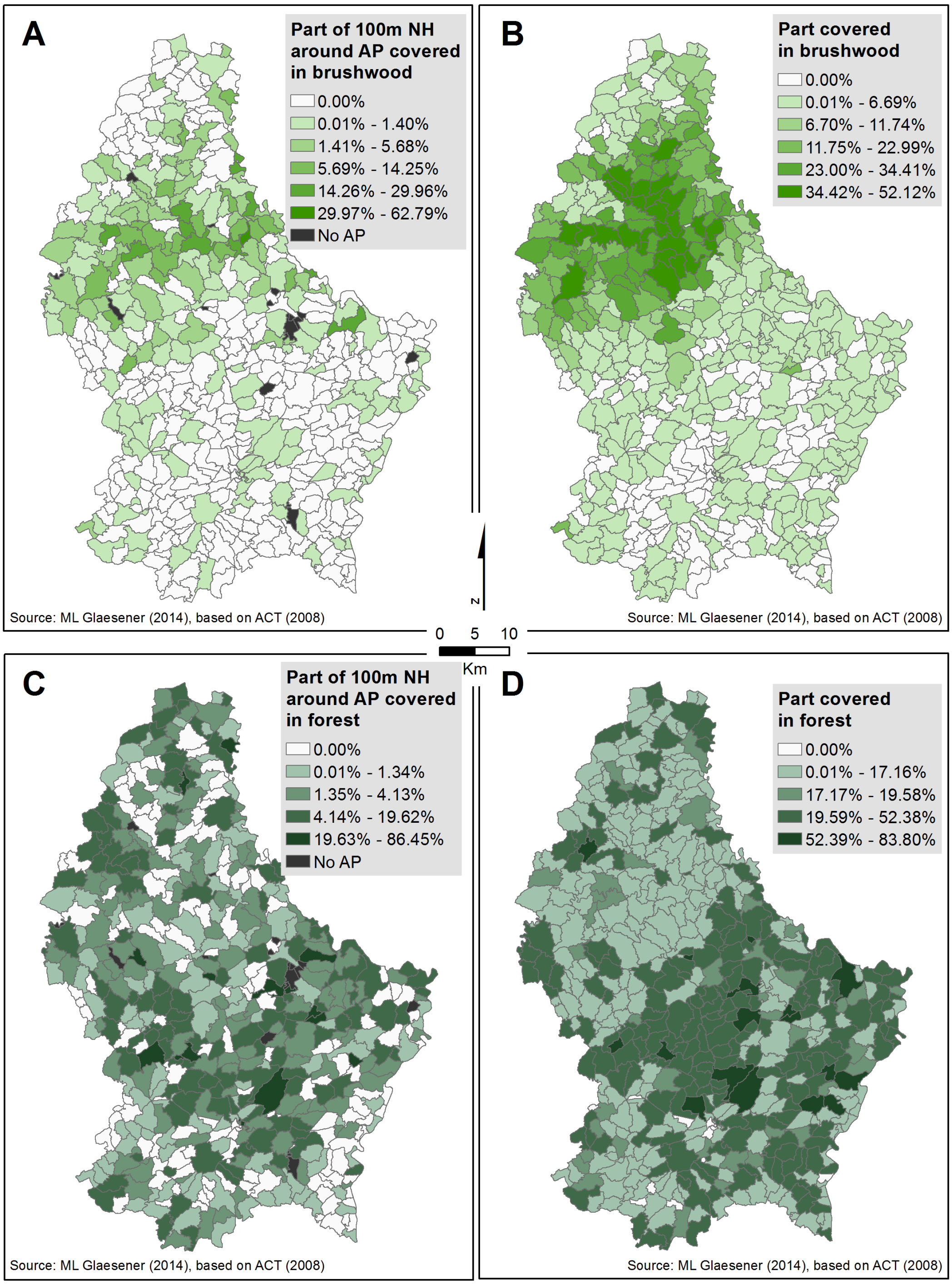


Figure 2.9: Share of forest

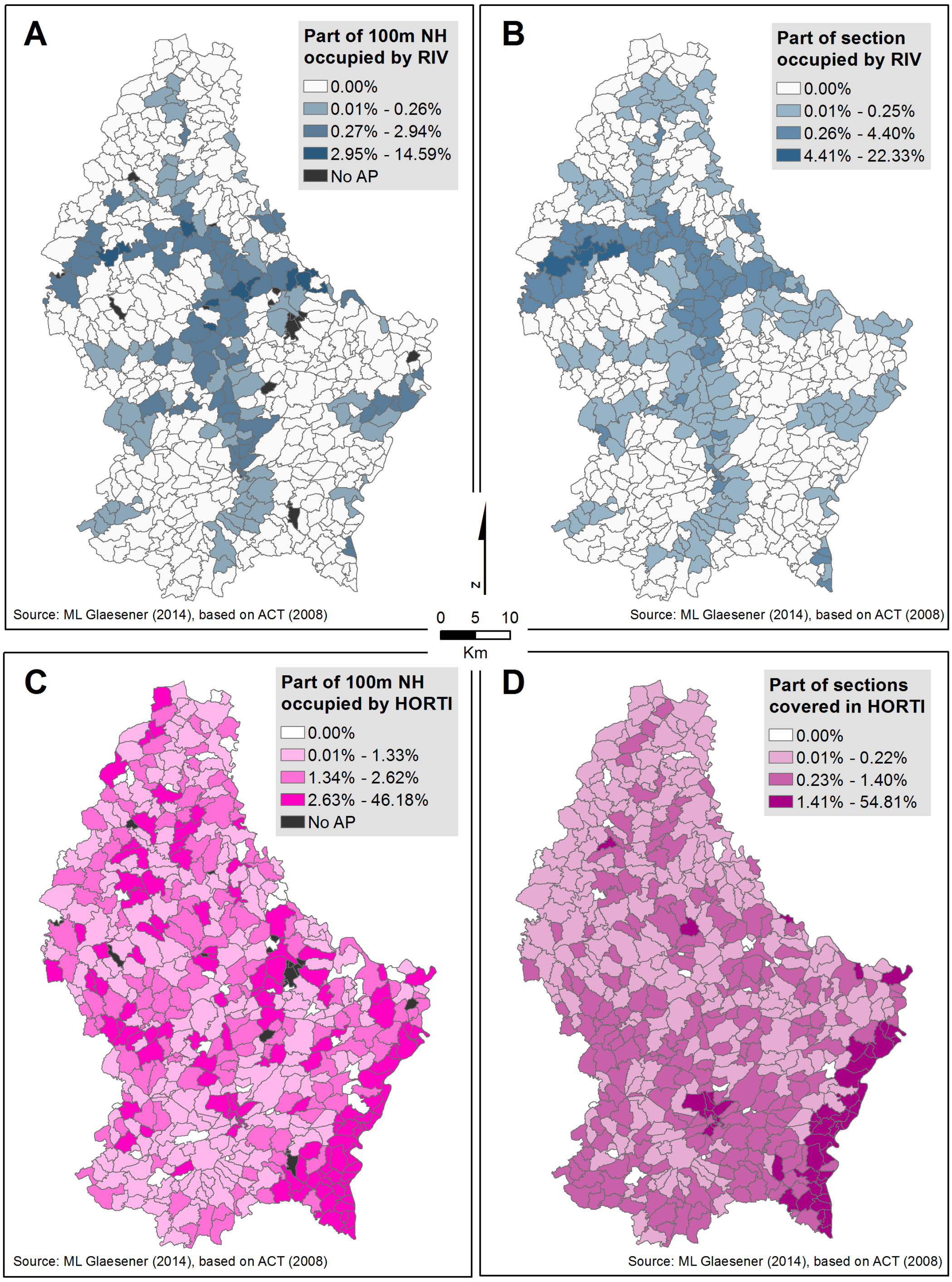


Figure 2.10: Share of rivers and horticulture

the neighbourhood of our transactions; in this section an attempt to approximate this neighbourhood was presented at aggregated section scale. Although the *AP* do not correspond to the actual transactions and despite the fact that only a general section mean could be considered, we believe that these measures might be able to capture some of the effect neighbouring land-uses have on developable land prices in Luxembourg and they are therefore considered in our hedonic model in chapter 3.

2.3.2 Land-use diversity

Besides the presence of different land-uses, it has been shown that land-use diversity and fragmentation are considered by land consumers, but to date these are rarely considered in hedonic pricing studies. Some indicators to analyse composition and configuration of different land-uses have been developed and considered in the hedonic context (Geoghegan et al., 1997; Kestens et al., 2004; Baranzini et al., 2008; Treg, 2010).

To quantify land-use diversity, the Shannon diversity index in different buffer zones around available plots (2.3.2.1) as well as a general green diversity index (2.3.2.2) at section scale will be considered.

2.3.2.1 The Shannon diversity index

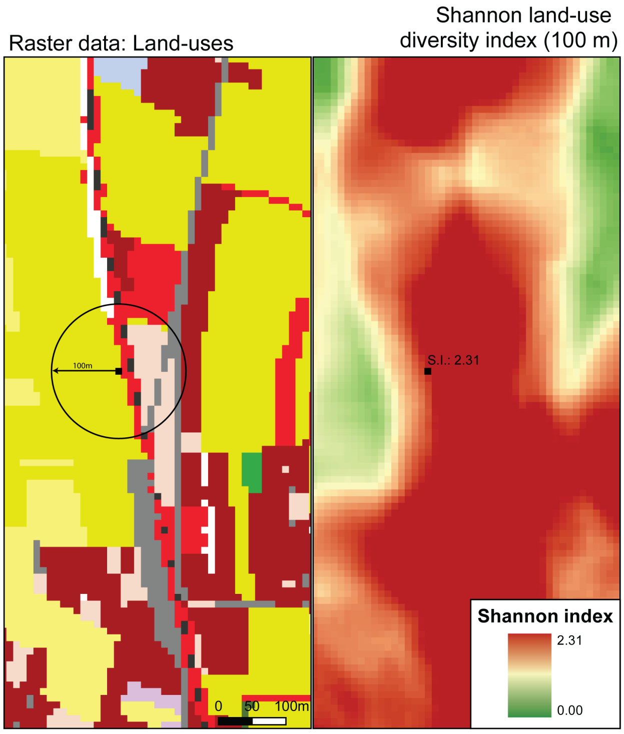
The Shannon Diversity Index is defined as "*a measure of diversity and evenness that reflects both the number and the balance of unique values within an area*" (Jenness et al., 2013), formalised in equation 2.2. This index measures the average "*degree of uncertainty in the prediction that an object, chosen as random from a group, will belong to a certain category*" (Finotto, 2011, p. 56), the uncertainty is supposed to increases with the amount of categories and the even distribution of the same. In other words, the index increases with both the number of classes observed and with how evenly distributed they are. The index varies from 0 to infinity, the greater it is, the greater the diversity of land-uses around a cell[17].

[17] Shannon Diversity Index was computed using an open-source add-in for ArcGIS by Jenness et al. (2013), that provides a diversity measure for every cell considering its neighbouring cells in a defined extent.

The index is calculated as follows:

$$\text{Shannon's Index H} = -\sum_{i=1p_i lnp_i}^{s} \tag{2.2}$$

Where $p_i = \frac{n_i}{N}$ is the proportion of observations in land-use i; n_i the number of observations in land-use i; N the total number of observations and s the number of land facets (Jenness et al., 2013). The land-use dataset considers 14 different land-use classes (fig. C.4 in appendix C.3). The Shannon diversity index was hence measured for every cell considering a fixed neighbourhood (fig. 2.11).



Source: ML Glaesener (2014) based on ACT (2008)

Figure 2.11: From raster to Shannon diversity index

Two extents, 100m and 1,000m have been considered as buffer zones around AP, following Geoghegan et al. (1997). The 100m radius should capture the diversity in the immediate neighbourhood of a transaction; while the 1,000m radius should capture diversity in walking distance. By raster calculation we identified the Shannon diversity index for all AP cells (similar to the previous section). By zonal statistics, the mean Shannon land-use diversity index around AP per section (*mAPsh100* and *mAPsh1000*) was generated. The procedure and steps chosen for the computation of the Shannon diversity index has been further illustrated in figure 2.12.

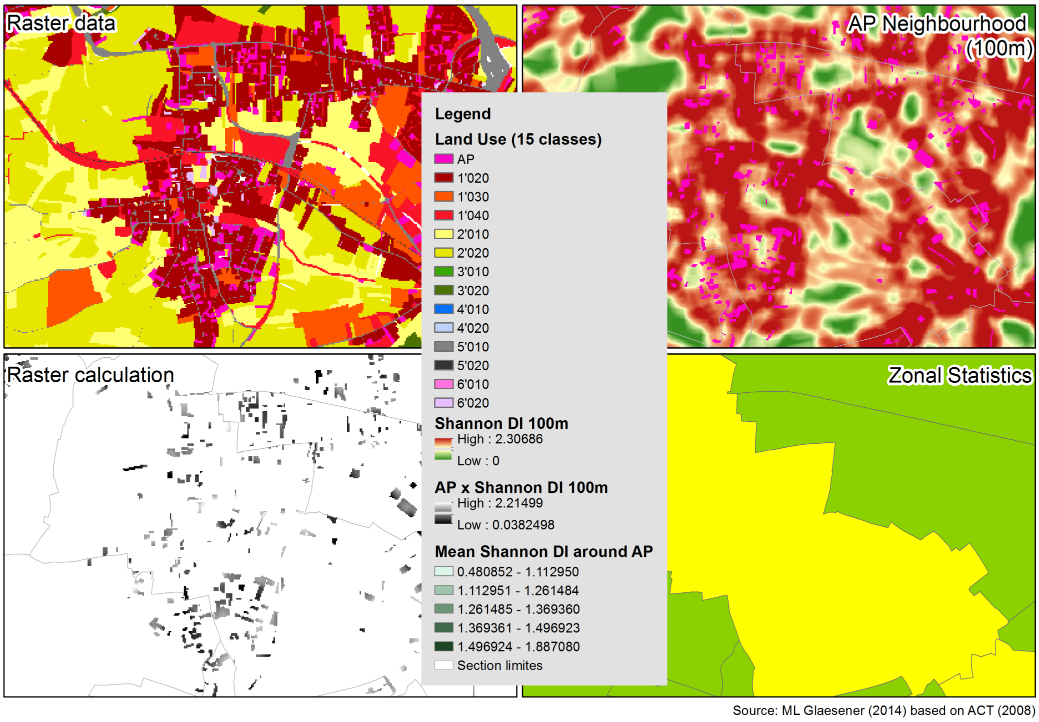


Figure 2.12: Aggregated Shannon diversity index

The average section values of the Shannon diversity index in the 100m neighbourhood around AP show some spatial variations (Map A fig. 2.13) with close land-use diversity lowest in the former industrial south (table 2.8), as well as the agglomeration of Luxembourg and the close periurban. Sections with high values are mainly observed in the east ("Moselle" Valley) and occasionally in the northern part of the country.

Considering the distribution of the average Shannon index in 1,000m around AP (Map B fig. 2.13), diversity is highest in the more urbanised areas (table 2.8) while it is found to be on average lowest in the close periurban, followed by the rural and distant periurban areas. Hence in walking distance around AP, land uses are more diverse in urban areas, and the low values in the rural and periurban area are probably due to the generally larger agricultural areas.

Caution should be put in the interpretation of these maps, since they suggest some similarities in land-use diversity between sections located within very opposite spatial contexts. The land-use diversity observed in the north is however most probably very different to the one observed in the more urban sections, this spatial heterogeneity has thus to be accounted for. Hence consumers might value diversity quite differently depending on the location of their parcel within Luxembourg. Geoghegan et al. (1997) identified non-stationary valuation of land-use diversity through space, a point relating to our third research question which will be further detailed in chapter 5.

	mAPsh		
	100m	1,000m	DIGreen
Rural	1.36	1.49	0.60
Distant periurban	1.35	1.53	0.57
Close periurban	1.31	1.47	0.54
Agglo. South	1.26	1.76	0.53
Agglo. Luxembourg	1.30	1.83	0.51
Total	1.33	1.53	0.57

Table 2.8: Average Shannon diversity indices by region

2.3.2.2 Green diversity index

A diversity index for all "natural" land-uses (DIGreen) based on areas occupied by different green land-uses (forest, brushwood, watershed, rivers, vineyards, gardens, pastures and crop-land) implements the same procedure as for urban amenity diversity index (presented in section 2.2.2 page 55) and formalised in equation (2.3).

The total area of the different "green" land-uses was considered per section and the index indicates on the diversity of different green land-uses[18] within a section; with n_c representing the amount of a specific land-use and n the total area covered in "natural"

[18] with n = sagg+ sagb+ sbw +sfor +sriv+ sws+ svy +sgd; for abbreviations see table C.8 p.227.

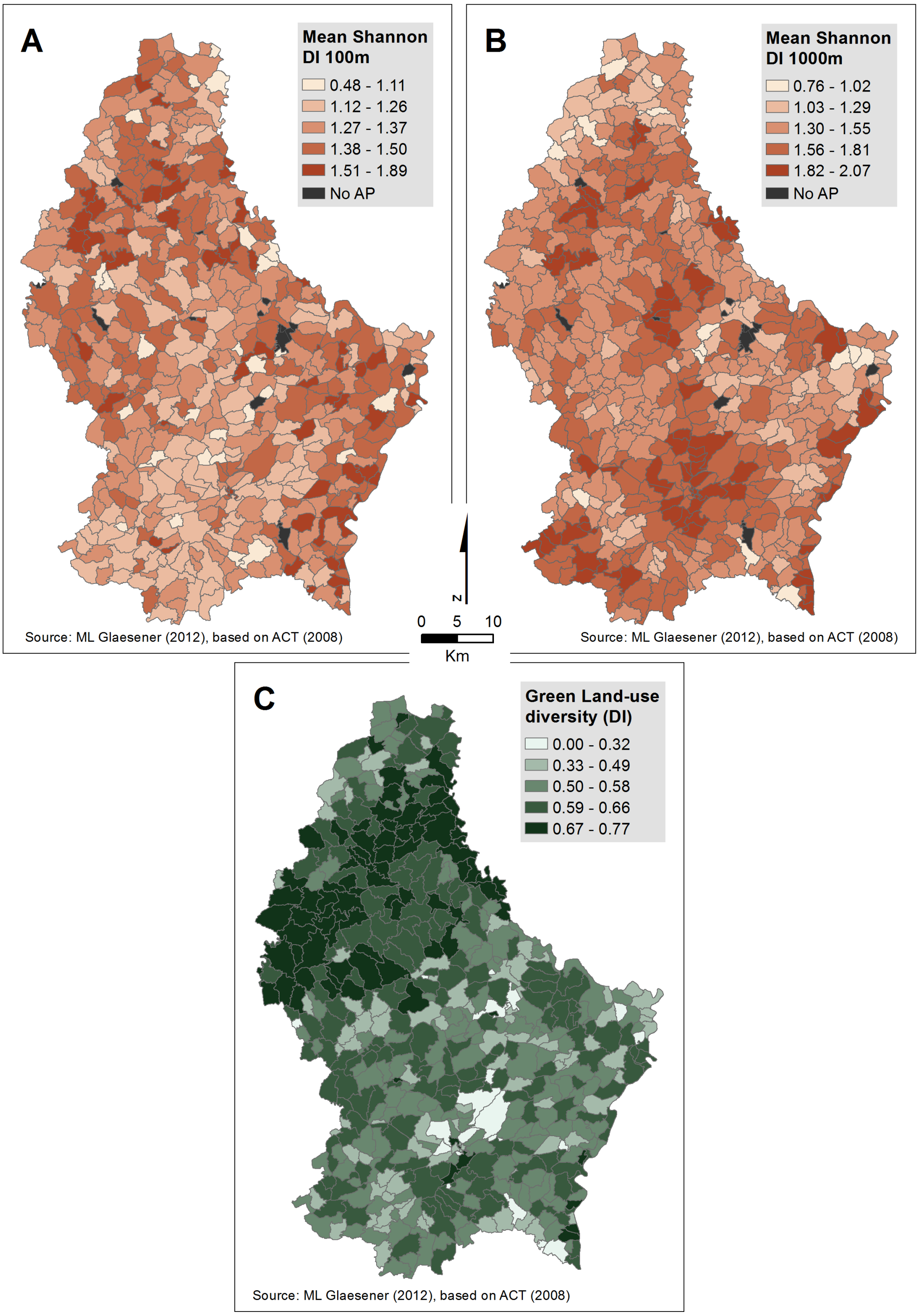


Figure 2.13: Shannon and green diversity indices

land-uses.

$$DIGreen = 1 - \sum_c (\frac{n_c}{n})^2 \quad (2.3)$$

The index varies between 0 and 1, a section with a *DIGreen* of 0 indicates one single land-use. Spatial variations of the index throughout the Grand Duchy are illustrated in map C (fig. 2.13). High values are especially registered in the northern part of the country, while sections around the capital and in the eastern part of the country register clusters of lower green diversity. This map translates the two main geological regions into which the country is commonly divided: in the "Oesling" in the north and the "Gutland" in the south, illustrated in map H.2 (p.260). The first being largely characterised by hilly landscapes (mainly covered in brushwood and forest, as already mentioned) alternating with agricultural land-uses on the plateau. While in the southern part of the country, the landscapes are largely characterised and dominated by agricultural uses (pastures and crop-land) and hence present a more homogeneous land cover. The aim of including this general green diversity index was to identify the marginal effect of an increase in the general green diversity of a section on developable land price, hence no expectations were raised.

Expected results

Table 2.9 summarises the different variables and the marginal impact expected on developable land prices in Luxembourg.

Variable	Unit	Description	Impact
rAP	%	% of *AP* in total residential and available area	+/-
rsAP100bw	%	% of area covered in brushwood in 100m NH around *AP*	+/-
rsAP100for	%	% of area covered in forest in 100m NH around *AP*	+/-
rsAP100riv	%	% of area covered in rivers in 100m NH around *AP*	+
rsAP100horti	%	% of area covered in horticulture in 100m NH around *AP*	+
DIGreen	index	Total green diversity index	+/-
mapsh100	index	Mean Shannon diversity index 100m NH around *AP*	-
mapsh1000	index	Mean Shannon diversity index 1,000m NH around *AP*	+

Table 2.9: Expected impact of land-use variables

2.4 Sections' socio-economic composition

Real estate consumers account for the social and economic composition of their neighbourhood and have preferences for more similar neighbours and hence for more homogeneous socio-economic neighbourhoods (Mieszkowski and Mills, 1993; Glaeser, 2008) and in particular for richer neighbourhoods as they are associated with increases in offer and quality of infrastructure. This has been investigated in several hedonic studies (Orford, 2000; Brander and Koetse, 2011), which underline that the neighbourhood context should be controlled for in the hedonic models to approximate spatial heterogeneity.

In the framework of this thesis we relied on two distinct databases, on the one hand the data available at municipal scale provided by the National Statistics Agency (STATEC), due to the aggregated scale of the census data and the importance of neighbourhood characteristics at local scale, an alternative dataset was generated. The "Inspection Générale de la Sécurité Sociale" (IGSS) provided more detailed information on residents at a more local scale[19], further discussed in appendix C.4. The IGSS dataset contains information on all residents of the Grand Duchy belonging to the national health system, the only exception being the civil servants of the European institutions. We considered the data available for 2007, which is the reference year of our residential land transaction dataset[20]. The dataset is based on the EUROMOD database, generated from different datasets managed by the IGSS.

Initially the anonymised data was provided at postal code level; however since the land transaction dataset was not available at this fine scale, the IGSS data had to be aggregated to section scale, which represented a substantial workload. 29 sections registered no residents in 2007, while 9 sections registered fewer than 11 inhabitants[21]. According to the IGSS data 471,062 residents were registered in Luxembourg in 2007, 476,187 according to STATEC, revealing some differences between the two datasets.

[19] A request at the IGSS was made to access the data they manage on the residents of Luxembourg, after approval of the National data privacy commission (CNPD).

[20] Due to time constraints but also because it is the year of reference of the residential land transaction dataset presented in section 1.

[21] In the following data description and the regression analysis, 491 sections are considered, whereas due to data privacy concerns, illustrations will only represent 482 sections.

Based on these two datasets different measures of the socio-economic composition will be presented in the following subsections. In section 2.4.1 the distribution of population in space will be presented, in 2.4.2 the evolution of population between 2001 and 2007 is illustrated, 2.4.3 discussing measures of the age structure of residents and eventually 2.4.4 the distribution of wealth and unemployment in Luxembourg.

2.4.1 Population distribution

The total population allows for conclusions on the relative importance of a section in terms of residents, while population density should grant insights into the distribution of population within the built area. Although gross population density is generally considered in empirical work, we follow among others Anas et al. (1998) claiming that gross density may overstate the size of density gradients because of the higher proportions of undeveloped land in the more rural areas. While population density, as well as the rate of population variation presented in section 2.4.2, may be seen as a measure for demand and relative scarcity of residential land (Treg, 2010; Brander and Koetse, 2011; Goffette-Nagot et al., 2011) and account for some urban effect, as high densities foster competition for land, it is considered a synonym for high land prices by Goffette-Nagot et al. (2011). On the other hand, Cho and Roberts (2008) claim that although population density captures housing demand at the neighbourhood level, it represents congestion of inhabitants, confirmed by Wu and Dong (2013). Cho and Roberts (2008) estimate a negative impact on property prices from an increase in population density. Population density can have opposite effects on housing prices (Geoghegan et al., 2003), hence either effect in the Luxembourgish case is possible.

Total population is only considered here for the means of illustration (map A fig. 2.14). With regard to the literature presented above, net population density, the part of residents per km^2 of area occupied by build residential land-use (sBR)[22] (Map B figure 2.14) has been considered in the hedonic model. The importance of the south-west as a population centre and in particular the agglomeration of Luxembourg is confirmed by map A (fig. 2.14). With regard to the typology in appendix B, the periurban areas

[22] Explanations on the land-use datasets can be found in section 2.3.1.1.

(close and distant) register by far the biggest part of the population (table 2.10). About 46% of Luxembourg's population lives in periurban areas while 39% resides in the areas classified as agglomerated. Overall, the agglomeration of Luxembourg registered 109,643 residents in 2007, while the sections of the rural area only contain about 15% of the total population.

As expected, a concentration of high net population density is observed in the southern part of the country, mainly around the agglomeration of Luxembourg and the south-west. The average net density in sections located in the distant periurban sections is of 964 inhabitants per square kilometre, and as expected sections located in the rural area are least densely populated.

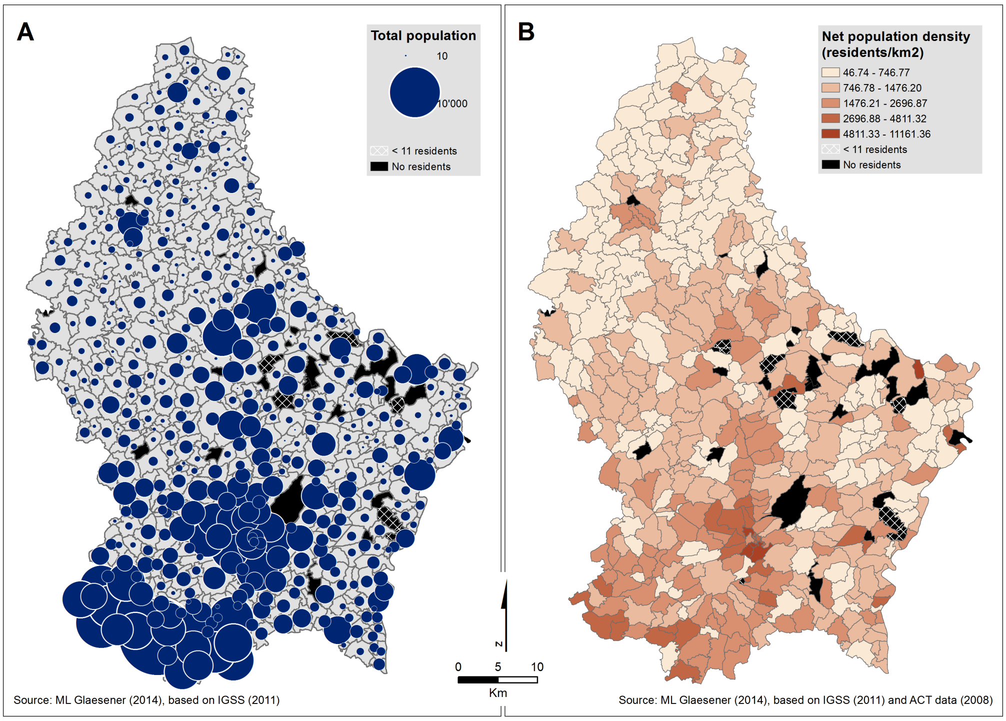


Figure 2.14: Total population and net density (2007)

2.4.2 Population variation between 2001 and 2007

The periurban area is generally associated with increasing population and thus with demographic growth. Luxembourg has known an important population growth in the

Typology	Total Population	Net density
Rural	71,224	753.74
Distant periurban	94,284	964.02
Close periurban	123,299	1,199.99
Agglo Esch	72,612	2,535.61
Agglo Luxembourg	109,643	2,892.32
Total	471 062	1129.91

Table 2.10: Population and net density by region (2007)

last decades (as presented in the introduction), however the distribution of this growth is not homogeneous throughout the country (map A fig. 2.15). Based on the IGSS data the rate of population variation at section scale could be generated for the period previous to the land transactions provided by the AED (2007). This variable should capture the effect of demand for housing at section scale.

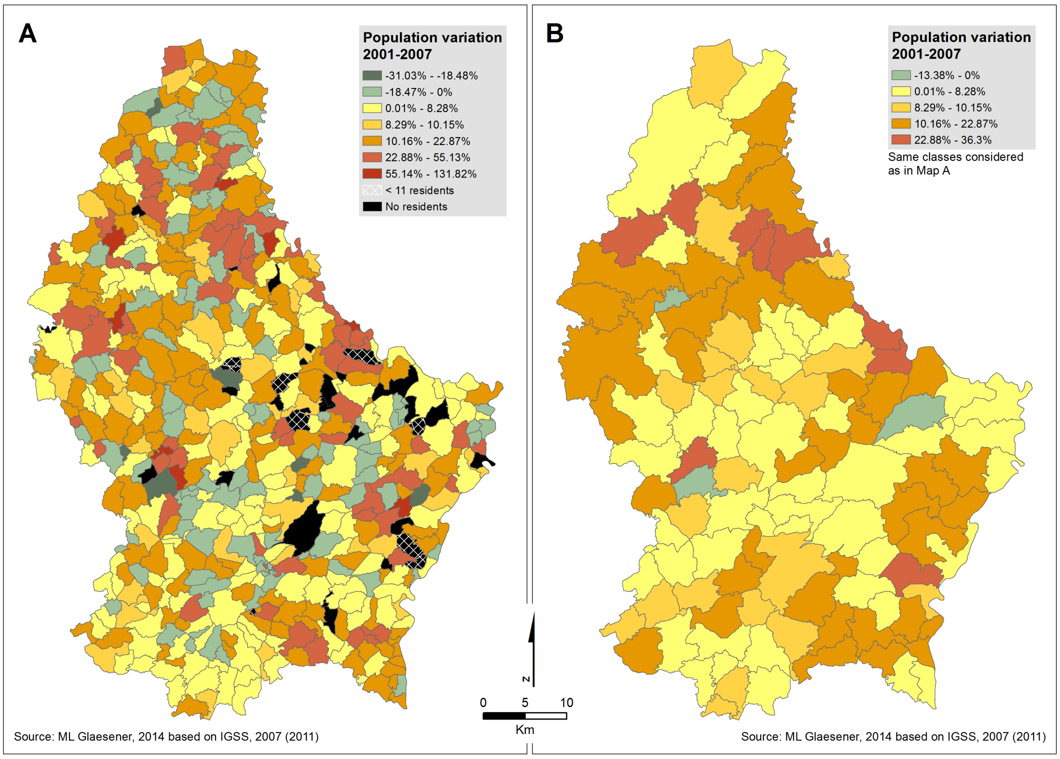


Figure 2.15: Population variation at two scales

To illustrate the differences between IGSS and STATEC data, maps A and B (fig.

2.15) represent the total population variation between 2001 and 2007[23] at section and municipal level. In general, both maps illustrate higher population growth in the northern part of the country, however some within municipal variations of population growth can be observed. Especially in the distant periurban and rural areas, many municipalities, despite the fact that they show a positive variation on average, register sections with negative population variation. Thus within municipalities, some sections appear more attractive than others. The intra-municipal variations confirm the importance of considering the finer scale to control for the effect of past demand on land values. Different assumptions can be made with regard to the impact of population variation and how it affects residential land price. Villages (sections) within a municipality might not be equally attractive (e.g.: due to missing public amenities); the main village (including the town-house) might offer a wider range of local public services or shopping opportunities. This pattern is most likely as well related to the supply of land and housing, in turn possibly related to spatial planning policies.

In general, the rate of population variation is on the one hand increasing with distance to Luxembourg-city (map A fig. 2.16); on the other hand, a negative relationship is observed between *PopVar* and net population density (map B fig. 2.16), as expected. By comparing the different regions (table 2.11), the rural area has known the strongest population growth, followed by the distant periurban area. The agglomeration of Luxembourg has registered the lowest growth.

Typology	*PopVar*
Rural	13.08 %
Distant periurban	9.02 %
Close periurban	8.94 %
Agglo. South	8.00 %
Agglo. Luxembourg	5.61 %

Table 2.11: Average population growth by region (2001-2007)

High demand for housing within a section or municipality can have a positive impact

---

[23]
$$\text{Variation rate} = \frac{(\text{total07} - \text{total01})}{\text{total07}}$$

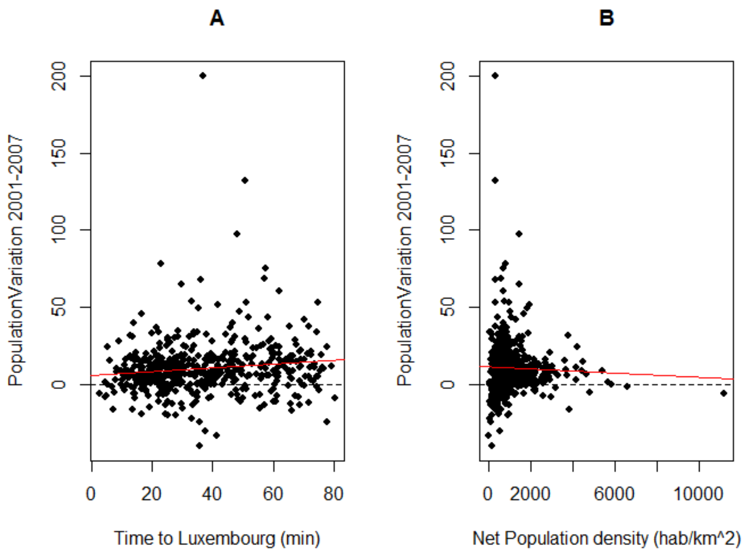


Figure 2.16: Population variation vs. distance and net density

on land values, since land consumers expect them to be related to future infrastructure improvement and additional urban amenities. However, the effect can as well be negative, since they imply mostly modifications in "natural" land-use, that could be anticipated by land consumers and considered in their choice.

2.4.3 Residents' age structure

Age or education level are usually considered as a proxy for social composition or structural weakness of a neighbourhood (Beron et al., 1999; Treg, 2010; Brunauer, 2013; Chasco and Le Gallo, 2013). Hence, to account for the residents' social composition within a section or municipality, the ratio of residents belonging to different age classes has been considered. Residents below 15 years (*rPopbelow15y*) are assumed to live with their parents and this ratio should allow to approximate the presence of households with children in a section.

In map A (fig. 2.17) the section averages per section are represented and the lowest share of young residents is observed in and around the agglomeration of Luxembourg, registering an average of 16%. No substantial differences can be observed between the

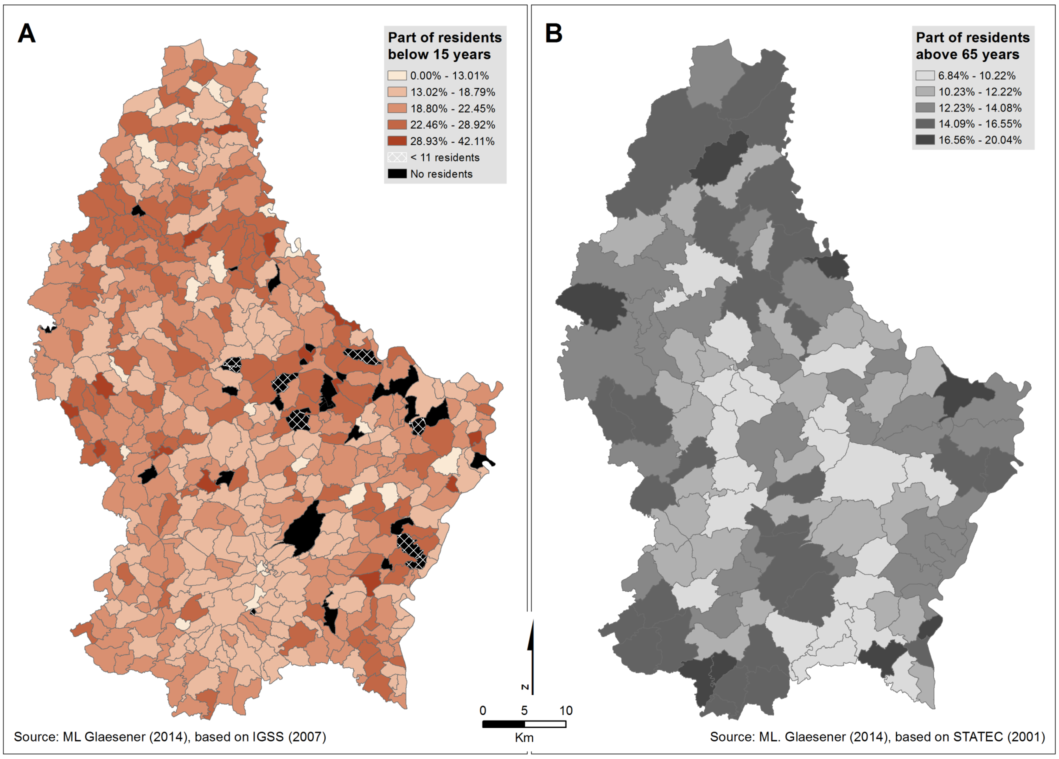


Figure 2.17: Residents below 15 years and above 65 years

other regions (table 2.12). The distant periurban and rural area being slightly above average might be interpreted as confirmation on the attractiveness of these regions for families with children. Additionally including average household size per section did not provide significant results. From map A (fig. 2.17) it can be observed that the sections belonging to the agglomerations and the periurban area register lower parts of young residents, than the more rural sections of the north. This finding is as well confirmed by Bousch and Decoville (2012), showing that the urban centres and the periurban belt register in general older populations and that the rural municipalities register a relatively younger age structure.

The part of residents above 65 years has been considered at municipal scale[24], and it can be observed that in the more urbanised areas, the share of retired individuals is

[24]In table 2.12 the section means from have been considered.

above the national mean compared to the observations made in the periurban and rural area, as illustrated in map B (fig. 2.17). This might on the one hand be determined by the distribution of rest homes and on the other hand due to the recent population growth in these areas (as shown in section 2.4.2). As was shown by Bousch and Decoville (2012), caution should be taken with regard to the part of residents above 65 years, as this ratio might be biased due to the existence of retirement homes within sections. For this reason, this ratio has only been considered at municipal scale, which might still be to disaggregated (Bousch and Decoville, 2012).

	$\prec$ 15 years	$\succ$ 65 years
Rural	21%	14%
Distant periurban	21%	14%
Close periurban	20%	13%
Agglo. South	20%	16%
Agglo. Luxembourg	16%	17%
Total	20%	14%

Table 2.12: Part of residents by age class by region

2.4.4 Wealth distribution and unemployment

To approximate the economic status, unemployment rate and median income are frequently considered (Treg, 2010; Chasco and Le Gallo, 2013). The age or education level are usually taken into consideration as a proxy for social composition or structural weakness of a neighbourhood (Beron et al., 1999; Treg, 2010; Brunauer, 2013; Chasco and Le Gallo, 2013)[25].

Information on residents' income classes was provided in the IGSS dataset[26]. These classes are based on the minimum social wage (SSM), which was fixed at 18,843 € per year in 2007. The IGSS considers all earnings residents obtain from: a regular job, pensions, apprenticeship, from sick, maternity or parental allowance, accident compen-

[25]Variables on income or the level of education, as a proxy for disposable income, were not available neither from the IGSS nor from STATEC dataset.
[26]The data available from STATEC does not provide an alternative to account for residents' income.

sations[27] as income. They range from below one SSM ($\prec$ 18,843€[28]) to more than seven SSM ($\succ$ 131,901€) per year. In this perspective we considered the ratio of residents belonging to the class of income above five SSM ($\succ$ 94,215€ per year), similar to what was done by Liu and Hite (2013), to identify the part of residents that could be considered as "rich". The distribution of the ratio of *HighIncome* residents is illustrated in map B figure 2.18. Although this does not inform directly on the wealth of households, it identifies sections with larger percentages of "rich" residents and might be comparable to approximated income variables (e.g.: median household income) considered in other hedonic studies (Anderson and West, 2006; Liu and Hite, 2013).

Map B (fig. 2.18) confirms significant clusters of high-high percentages in and around the capital, although there is a trench in the south east of the capital, suggesting a random distribution within this part of the agglomeration of Luxembourg. This confirms that the high income residents tend to cluster in the southern part of the country in proximity to the agglomeration of Luxembourg. On the other hand, significant clusters of sections with low shares of "rich" residents have been observed in the north and southwest of the country. Table 2.13 confirms that the share of residents earning more than five SSM tends to be higher in the agglomeration of Luxembourg and the close periurban area, while the Agglomerated South and the rural area register the lowest part of "rich" residents. Further, as illustrated in fig. 2.19, the part of high income residents decreases with distance to the centre, especially if travel time to Luxembourg-city exceeds 20 minutes.

Controlling for the share of high income residents in the hedonic model, should allow to control for the economic composition at section level and identify whether the theory of individuals aiming at being located close to similar neighbours can be confirmed. Especially since residential land consumers themselves are most probably part of a higher income class. Hence a positive effect on land prices is expected from an increase of the part of rich residents within the same section.

[27] A drawback of the IGSS dataset is that the first class, residents earning less than one SSM, includes all residents having no income (e.g.: children, housewives) as well as those earning the minimum SSM.
[28] Defined by the Ministry of Social Security.

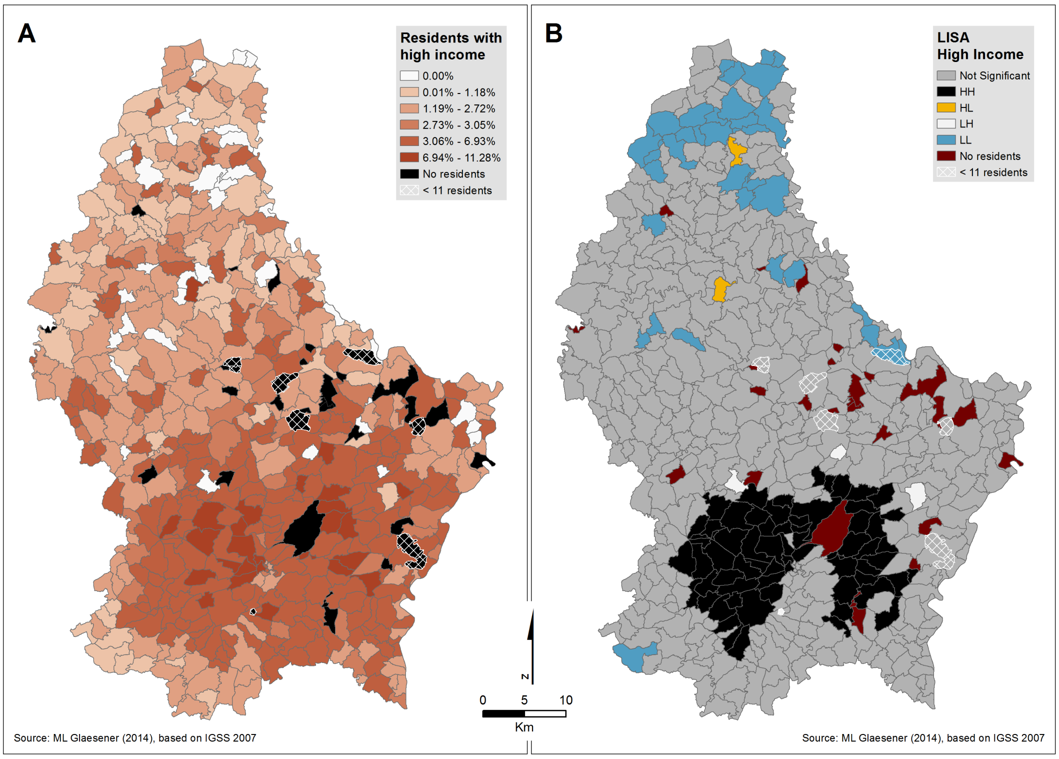


Figure 2.18: Part of "rich" residents

	rHighIncome
Rural	1.64%
Distant periurban	2.24%
Close periurban	4.33%
Agglo South	1.62%
Agglo Luxembourg	5.69%
Total	2.88%

Table 2.13: Part of high income residents by region

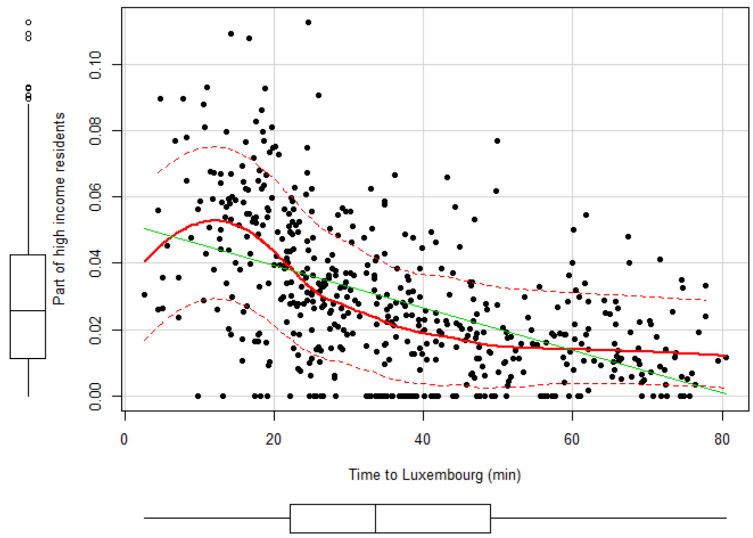


Figure 2.19: High income vs. distance to Luxembourg-city

To further approximate the economic status, unemployment rate is frequently considered in the hedonic pricing context (Kiel and Zabel, 2008; Treg, 2010; Chasco and Le Gallo, 2013). In this thesis, the unemployment rate at municipal scale, provided by STATEC (2008)[29] was hence considered to account for the marginal effect of high rates of unemployment.

Map A (fig. 2.20) represents the proportion of residents considered as unemployed within the total population, while map B illustrates the distribution of unemployment as determined by STATEC at municipal level. On average, the unemployment rate in Luxembourg was quite low in 2009 (4.52%). High rates are noted in particular in the municipalities of the former industrial south as well as in the north west (around Wiltz) and the east (around Berdorf), and eventually in the municipality of Luxembourg.

Land consumers generally avoid areas with high unemployment (Kiel and Zabel, 2008; Treg, 2010; Chasco and Le Gallo, 2013), since they are in general associated with higher criminality rates and less neat neighbourhoods, and this effect should be ac-

[29]As we were not able to generate this information from the IGSS dataset in accordance with the definition of the active population generally considered in Luxembourg.

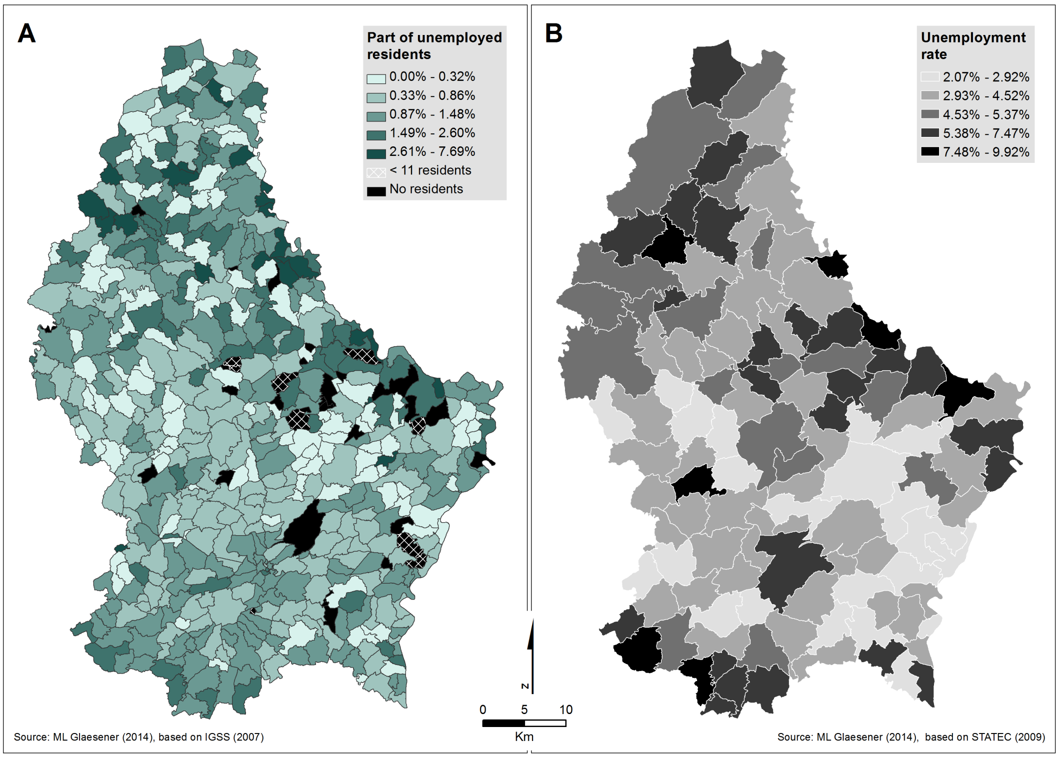


Figure 2.20: Unemployment rate

counted for.

Expected results

These socio-economic variables were considered in the hedonic analysis to control for further contextual effects at section and/or municipal scale, depending on data availability. The interpretation of the coefficients of these results is not the main focus of this thesis, however some expectations towards these variables have been recalled in table 2.14.

Variable	Unit	Description	Impact
NetPopDens*	hab/km^2	Net population density	+/-
rBelow15*	%	Population younger than 15 years	+/-
rAbove65**	%	Population older than 65 years	+/-
rHighIncome*	%	Population above 5 social minimum wages	+
VarPop0107***	%	Population variation between 2001 and 2007	+/-
rUnemployment**	%	Unemployment rate	-

Note: Considered at: * section scale, ** municipal scale and *** both levels

Table 2.14: Measuring socio-economic context

Concluding remarks

This chapter aimed at presenting the variables considered in the empirical part of this research to capture the marginal effect of the local urban and rural characteristics. The focus is turned to measures of accessibility, the diversity of local urban amenities as well as land-use diversity. Eventually a variety of socio-economic variables have been presented to control for the effect of the neighbours' characteristics on land prices.

Conclusion to part I

In this part the different types of attributes considered to impact on the consumers' willingness to pay for a piece of land have been presented. The limitations are mostly related to the aggregated level of the transaction dataset and the little information available on the structural attributes of the developable land transactions. The consideration of different transport modes, the attempts to approximate the local neighbourhood at aggregated scale and the land-use diversity indices are rather uncommon in hedonic literature, and analysing the consumers' preferences for these should contribute hence to the existing literature. Although the expected effects of the explanatory variables (X) have been presented, they are recalled briefly within the different chapters of part II, accounting for the respective research question and model specification.

Further, the local specificities of the periurban areas, determining their attractiveness and the related challenges for spatial planning, have been highlighted and related to the other morphological and functional areas of the Grand Duchy as identified in appendix B.

The second step is now to relate the dependent variable (Y) to the identified explanatory variables (X), this will be done via the hedonic pricing method, presented in part II.

Part II

Measuring consumers' preferences: Spatial econometric modelling

Introduction to part II

As discussed in the introduction, land values mainly depend on the standard trade-off described in the monocentric model, between accessibility to jobs and land consumption (Alonso, 1964; Fujita, 1989). Further, it is well-known since Tiebout (1956); Brueckner et al. (1999) and Glaeser et al. (2001) that the provision of local public goods is an important aspect in residential competition. And recent theoretical advances have shown that the local arrangement of green space impacts on urban form and its scattered or leapfrogging nature (Cavailhès et al., 2004; Caruso et al., 2007, 2011; Turner, 2005).

Cavailhès et al. (2004) demonstrate how the emergence of the periurban belt is linked to the benefits land consumers obtain from a mix of local urban and rural amenities. The spatial distribution and diversity of land-uses as well as neighbourhood retail and services around residential places are of particular interest here. As discussed, land prices in a competitive market in equilibrium are the result of the consumers' demand as well as the supply. The price paid for land translates the consumers' willingness to pay for a specific parcel, which can be considered as a composite of the structural and location-specific attributes it encompasses. The hedonic pricing method[30] allows to assess the amenity value of these characteristics of land.

This part aims at estimating how the basic trade-off and the periurban mix of rural and local urban amenities are valued by land consumers in Luxembourg. It is subdivided into three chapters[31], illustrated in figure 2.21, dealing with the research questions raised in the introduction:

- What are the geographical determinants of residential land prices in Luxembourg?
- Are these determinants valued differently by different consumers?

[30]Further discussed in chapter 3.

[31]These three chapters are not self-contained, since the aim was to avoid unnecessary repetitions hindering the reading flow and to prevent too much back and forth between the parts at the same time. Hence the general theoretical background and literature have been presented in the introduction and part I, nevertheless each empirical chapter includes, in light of the specific research question, a short review of our expectations toward the considered explanatory variables.

- Are there spatial variations in the valuation of local green amenities and are we confronted to spatial market segmentation?

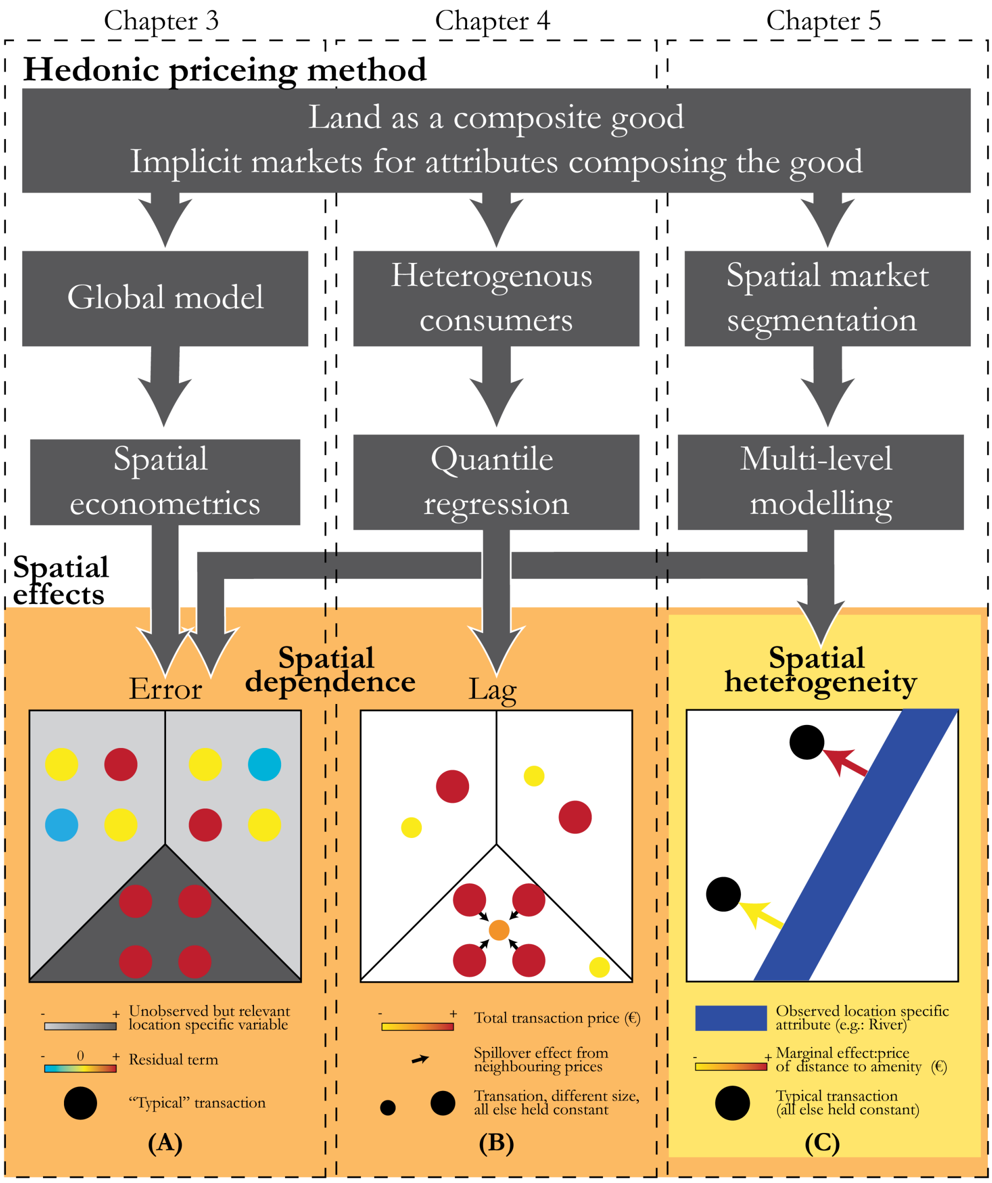


Figure 2.21: Parts' structure and spatial effects

The distribution of property prices is generally not random and homogeneous in space and the decision to purchase a good is the result of individuals' choices in a heterogeneous geographical context. Spatial patterns are expected to arise from a combination of spatial heterogeneity and spatial dependence (Anselin, 1988b). This part addresses issues related to both, spatial dependence and spatial heterogeneity, as illustrated in figure 2.21. Spatial dependence is considered to be the "*existence of a functional relationship between what happens at one point in space and what happens elsewhere*" (Anselin, 1988b, p.11). Distinction is made between two types of spatial dependence: error and lag dependence.

The first is usually related to omitted spatially autocorrelated variables (fig. 2.21: grey shades), shared by all the transactions of a particular neighbourhood but absent in the model specification (Ward and Gleditsch, 2008). The omission of a relevant spatial variable will bias the estimation results as the error term will be spatially autocorrelated and the assumption of independently and identically distributed errors is not met (fig. 2.21 (**A**)). In chapter 3, test results confirmed the existence of error dependence. While this chapter aims at providing first and general insights on consumers' preferences for structural and location-specific attributes in Luxembourg, the appropriate spatial econometric techniques have to be applied to obtain unbiased estimates for this global model.

The global hedonic model assumes homogeneous consumers with homogeneous preferences. As already discussed, consumers' preferences are likely to vary through life-cycles and their socio-economic background. To better understand Luxembourgish land consumers' preferences, it is hence important to account for the differences between the individuals purchasing land. This becomes even more important as the observations only allow limited information on the agents involved in a transaction, considering private and professional developers at the same time. Preferences have been shown to vary between private end-users and between the latter and those of professional developers, the first searching to maximise their quality-of-life and the second to maximise their profit. In chapter 4, the focus is turned to understanding the extent to which the consumers' socio-economic background impacts on their preferences for local amenities. Since within the different deciles, prices of typical transactions have been found to be spatially autocorrelated among themselves. Spatial lag dependence had thus to be accounted for via a spatial quantile model.

This relates to the second form of spatial dependence announced and frequently encountered in the hedonic pricing context. Spatial lag dependence can exist among real estate prices, for example if similar valuation practices prevail in a neighbourhood. It is a special case of cross-sectional dependence, where the co-variation structure between transactions at different locations is subject to a spatial ordering implying a simultaneous feedback between transactions (Anselin and Lozano-Gracia, 2009). This is for instance observed when landowners fix their prices according to those paid for other transactions in the same neighbourhood. Such spatial spillover effects induce for instance that transactions are even more expensive, all else constant, if other observations in the neighbourhood are more expensive. In other words, the same land parcel would be sold at different prices in different neighbourhoods, only because of the prices of neighbouring prices (fig. 2.21 (**B**)).

Further, market segmentation might arise if the prices of land and their characteristics vary substantially with their location in space (Le Gallo, 2004). This implies that OLS estimations, that impose spatial homogeneity, will be misspecified. In general, spatial heterogeneity can be referred to as a "*a special instance of structural instability*" (Anselin and Lozano-Gracia, 2009, p.1214). In the housing price context the term refers to a non-constant price-attribute relationship over space, hence that the marginal price of an attribute varies with regard to the transactions' location (fig. 2.21 (**C**)). Potential causes of spatial heterogeneity might be spatially differentiated characteristics of demand or supply or to local differences in land-use policies which lead to a systematic spatial variation in the location decision.

Although spatial heterogeneity and dependence have been presented separately here, spatial patterns in real estate markets are expected to be the result of a combination of spatial dependence and heterogeneity (Anselin and Lozano-Gracia, 2009), discussed and illustrated by De Graaff et al. (2001)[32] and Le Gallo (2004). Spatial heterogeneity and dependence are further addressed in chapter 5 via the multilevel modelling approach

[32]Presenting different reasons to consider them jointly: First, from an observational viewpoint there might be no difference between both spatial effects. Second, spatial dependence might lead to a specific form of heteroskedasticity and third tests for both effects might be incapable of making distinction between the two different processes sufficiently (De Graaff et al., 2001).

which allows to account for issues related to spatial error dependence and spatial heterogeneity at the same time.

Chapter 3

Geographical determinants of developable land prices in Luxembourg: Model specification and spatial effects

3.1 Introduction

As presented in the general introduction, considering the local context in which transactions take place is important to understand what determines the price consumers are willing to pay for land. The identification of geographical determinants is based on the findings of different urban economic models, highlighting on the one hand the importance of access to the CBD and parcel size (Alonso, 1964; Fujita, 1989) and on the other hand consumers' preferences for a mix of local urban and green amenities (Cavailhès et al., 2004). As also discussed, understanding what determines a consumers' choice to buy land is important to design effective and acceptable planning policies at different scales.

The hedonic pricing method developed by Rosen (1974) allows to estimate the marginal price consumers are willing to pay for the different attributes composing the good. It is a revealed preference analysis, based on the total price consumers have paid for a particular piece of land. Although at first housing prices and their structural determinants were at the centre of attention of hedonic modelling, Cheshire and Sheppard (1995) emphasised the need to simultaneously consider a broad range of location-specific attributes.

There have been numerous studies applying the hedonic pricing method with very different centres of interest. Literature reviews of hedonic studies focussing on the impact of local green on land prices have been presented by McConnell and Walls (2005) and Brander and Koetse (2011). The latter having collected in their meta-analysis over 52 hedonic studies addressing the valuation of open-space amenities. Brander and Koetse (2011) show that in general, neighbourhood land-use patterns and access to green amenities are valued by land consumers. The proximity to different land-use and rural amenities has been considered in a large body of hedonic literature, with a main focus on the valuation of open-space amenities (Bolitzer and Netusil, 2000; Acharya and Bennett, 2001; Shultz and King, 2001; Lutzenhiser and Netusil, 2001; Irwin and Bockstael, 2001; Smith et al., 2002; Geoghegan, 2002; Geoghegan et al., 2003; Anderson and West, 2006; Treg, 2010). Further the specific configuration of land-use patterns (i.e. fragmentation, diversity) have also been considered (Geoghegan et al., 1997; Cho and Roberts, 2008). A landscape is always more complex than the sum of its parts and each part has different characteristics on the basis of how it interacts with its surroundings (Finotto, 2011, p.48).

Local urban amenities have been accounted for in fewer hedonic pricing research. Some considered measures of school quality (Thériault et al., 1999; Uyar and Brown, 2007; Clapp et al., 2008; Kiel and Zabel, 2008), distance to public open-space (Espey and Owusu-Edusei, 2001; Mahmoudi et al., 2013; Wu and Dong, 2013) or the proximity of retail and services (Thériault et al., 1999, 2005; Cavailhès, 2005; Des Rosiers and Thériault, 2006; Des Rosiers et al., 2008; Youssoufi, 2011; Öner, 2013). Besides proximity to different local urban amenities, a rich variety in their supply has been shown to have a positive marginal effect on individuals' utility (Brueckner et al., 1999; Youssoufi, 2011; Öner, 2013).

The aim of this chapter is thus to account for these local amenities simultaneously in a hedonic modelling framework. In addition, a selection of structural and socio-economic control variables is considered to account for specificities related to the transaction and the neighbourhoods' socio-economic context. As real estate consumers account for the social and economic composition of their neighbourhood and have preferences for more homogeneous socio-economic neighbourhoods (Mieszkowski and Mills, 1993).

The structural and location-specific attributes are considered to compose the developable land. They are non-market goods and their marginal value can be estimated via regression analysis, based on the observed prices paid for land. With regard to the limitations imposed by the aggregated location of the individual transactions, different land-use measures focussing on the neighbourhood of all available parcels as well as land-use diversity indices have been considered at section scale. Under the assumption that these measures may approximate micro-scale measures of the transactions' neighbourhood, their impact on prices will be tested and revealed in this chapter.

In the last decades, special interest has been turned to accounting for spatial effects, mainly the issues related to spatial dependence, in spatial econometric modelling[1]. As pointed out in this part's introduction, spatial dependence is considered to be related to the presence of functional relationship between what is happening at different points in space (Anselin, 1988b) and coping with estimation bias related to spatial dependence is a crucial precondition for hedonic price studies today (Wilhelmsson, 2002)[2]. It is now commonly accepted that hedonic pricing models need to account for different forms of spatial dependence (Krause and Bitter, 2012) and many empirical examples can be found (e.g.: Möller, 2008; Brady and Irwin, 2011; Furtado, 2011b; Abelairas-Etxebarria and Inma, 2012; Kadish and Netusil, 2012; Seya et al., 2013, ...). As illustrated in figure 2.21 (p.94), spatial dependence can arise either between transactions or among the regression residuals.

The objectives of this chapter are twofold. First, we aim at specifying a global hedonic model to identify the significant determinants of land prices in Luxembourg. The hedonic model proposed aims at going beyond the traditional trade-off between land consumption and commuting cost. Further different proxy measures of local green

[1]The field of spatial econometrics deals with spatial interaction or/and spatial structure in regression models, it can be defined as "*a subset of econometric methods that is concerned with spatial aspects present in cross-sectional and space-time observations*" (Anselin, 2010, p. 5-6).

[2]Distinction has to be made between spatial and time series dependence, which is unidimensional as only the past may influence the future (Le Gallo, 2002), as it is at the same time two-dimensional (space and time) as well as multi-directional (Anselin and Lozano-Gracia, 2009). In this thesis, we only concentrate on the spatial aspects of dependence, which should be sufficient with regard to the short period of time considered.

neighbourhood at aggregated scale are tested as well as diversity indices and local urban amenities to explain developable land prices. Second, with regard to the critiques raised towards the general OLS estimation of the model, we aim at testing and eventually accounting for issues related to spatial dependence.

In section 3.2, we will further discuss the hedonic pricing method and different spatial econometric models considered here and partly in the following chapter[3]. The different explanatory variables will be briefly recalled in section 3.3, before eventually presenting the estimation results and findings in section 3.4. Eventually, section 3.5 draws a conclusion.

3.2 The hedonic pricing method

3.2.1 The hedonic price function

The hedonic pricing method by Rosen (1974) is based on the consumer theory by Lancaster (1966). It allows to estimate the marginal price consumers are willing to pay for the attributes composing a good[4]. Developable land should be seen as a composite good rather than a "*generalized housing commodity*" (Brueckner, 2011, p.117)[5]. The different attributes composing the good are assumed to be capitalised in the price consumers pay for real estate. The determinants of residential location choice as identified in urban economic literature are not traded explicitly on a single market.

Four main categories of price determinants from which consumers obtain utility have been identified in part I. First, the structural characteristics of the developable land parcel (S_i) (e.g.: parcel size, slope, orientation). Second, the accessibility measures to Luxembourg-city (A_i)(e.g.: distance to CBD) describe the functional link with the CBD, a major determinant identified in the monocentric urban model. Third, E_i describes the

[3]The underlying assumptions are presented in more detail in appendix D.

[4]For a more detailed presentation and discussion see (Palmquist, 2003; Cavailhès, 2005). Malpezzi (2002); Sirmans and Macpherson (2003) provide a general overview on empirical applications of the hedonic pricing method on real estate prices, and McConnell and Walls (2005); Brander and Koetse (2011) focus on those concentrating on open-space amenities.

[5]This is in opposition to Muth (1969), who considered price of housing as an unidimensional service, an expenditure needed to purchase a standardised quantity of services, produced by stocks of housing capital and land (McDonald and McMillen, 2011).

"natural" neighbourhood attributes (e.g.: open-space, land-use diversity) and the local urban amenities (e.g.: retail and services) surrounding a parcel i, as identified mainly in the periurban model. Eventually, N_i describes the socio-economic composition of the neighbourhood (e.g.: net population density, unemployment rate). The equilibrium price (P_i) can thus be considered as a function of these different attributes composing the good (i)[6]. The price function can hence be written as follows:

$$P_i = f(S_i, A_i, E_i, N_i) \tag{3.1}$$

The utility (U) consumers obtain from purchasing a residential land parcel (i) is a function of these attributes (S_i, A_i, E_i and N_i), added to all other consumed goods, c. The consumers' heterogeneous preferences for alternative packages are translated by the individual taste parameter α. With regard to their preferences, consumers obtain utility from the different attributes composing the good and all other consumed goods ($U(\alpha, c, S_i, A_i, E_i, N_i)$. They aim at maximising their utility under lowest cost[7], under a budget constraint, which depends on their income and the price paid for all other consumed goods (c). The linear hedonic function takes the following form:

$$P_i = \beta_0 + \beta_S X_S + \beta_A X_A + \beta_E X_E + \beta_N X_N + \epsilon \tag{3.2}$$

Where P_i is the total transaction price of residential land and β_0 the intercept, the overall transaction price if all explanatory variables are accounted for[8]. X_S is a vector of structural characteristics, X_E representing the natural environment attributes, X_N socio-economic characteristics and X_A distance to CBD. $\beta_S, \beta_E, \beta_N, \beta_A$ represent the regression estimates. The error term (ϵ) represents all the attributes not explicitly accounted for in the model (e.g.: due to missing data)[9].

[6] We mainly fall back on the presentation of Bowen (2001) and Anselin and Lozano-Gracia (2009).

[7] If two objects offer the same attributes but are sold at different prices; only the less expensive is considered by consumers.

[8] Hereafter and especially in the results section, it will be referred to as the estimated price of a typical transaction.

[9] An overview of the methodological challenges related to the specification of the hedonic model has been presented in appendix D.2.

By first order maximisation conditions of the consumers' utility and under the given budget constraint, this approach allows to estimate the marginal price consumers are willing to pay for the different attributes, all else held constant and normalised to all other consumed goods (c)[10].

This marginal price corresponds to a general willingness-to-pay for this amenities by all the considered consumers, but does not translate the demand for this attribute[11].

The question of the appropriate functional form has been largely discussed in literature and a more detailed review can be found in appendix D.2. The semi-log form is often favoured for hedonic housing price models as the log of the dependent variable promotes its distribution to be normal and allows to account further for residual heteroskedasticity (Anselin, 1988b). In this research, we considered the natural logarithm of the total transaction price deflated to 2007-Euros as a dependent variable, on the right-hand-side (3.2) parcel size, accessibility measures and population density were log transformed.

3.2.2 Spatial econometric models

With regard to the concerns raised regarding the robustness of hedonic models if spatial dependence is not explicitly accounted for, the main spatial models will be briefly described here. A detailed review of the different spatial models concentrating on spatial dependence and how to handle it via spatial econometric techniques can be found moreover in Anselin (2002); Le Gallo (2002); Ward and Gleditsch (2008); Anselin and Lozano-Gracia (2009)[12]. As presented before, two main types of spatial dependence might arise in the hedonic pricing context, spatial autocorrelation among the dependent variables (lag) and among the errors (error). For the sake of readability, of the following

[10] The marginal price p_{E_i} consumers are willing to pay for a change in the environmental variable E_i can be derived as presented by (Cavailhès, 2005, p.93).

$$\frac{\delta u/\delta E_i}{\delta u/\delta c} = \delta P_i/\delta E_i = p_{E_i} \tag{3.3}$$

[11] In chapter 4, the demand estimation through the second step of the hedonic price method is presented in more detail.

[12] A literature review of papers comparing traditional econometric techniques to spatial models are found among others in Bowen (2001); Anselin and Lozano-Gracia (2009).

explanations, a simplified formalisation of the hedonic price function will be presented in equation D.1 (p. 234), where Y is the total transaction price and X represents the sum of structural and location-specific explanatory variables: $X_S + X_A + X_E + X_N$.

Spatial autoregressive model

In presence of spatial lag dependence, the spatial autoregressive model (SAR) is generally put forward. The SAR completes the linear hedonic equation (equation D.1 p.234) by introducing a spatial lag operator (ρ), the spatial weight matrix W and WY the spatially lagged dependent variable, on the right-hand-side of the model. The linear spatial autoregressive model (SAR) takes the following (simplified) form:

$$Y = \rho WY + \beta X + u \tag{3.4}$$

The price of a transaction is thus not exclusively a function of the explanatory variables (X), it is at the same time determined by the prices of other transactions in the neighbourhood. The spatial relationships are defined by the spatial weight matrix (discussed in section 3.3 and appendix E). By introducing the spatial lag operator on the right-hand-side of the equation, an endogeneity problem arises and OLS estimates are no longer consistent, as it implies simultaneous spatial interaction (Fingleton and Le Gallo, 2008). This requires the use of an other estimation technique, accounting for this simultaneity[13]. Interpretation of the SAR estimates is not as straight forward as OLS estimation results, since the estimated coefficient of a SAR is not the marginal effect (Brady and Irwin, 2011). For a correct interpretation of the regression coefficients and their impact on price, the calculation of impacts is needed for spatial lag model because of the spillovers between the terms (fig. 2.21) (Kim et al., 2003; Anselin and Lozano-Gracia, 2009; Bivand, 2010, 2013). Among others, Anselin and Lozano-Gracia

[13]Two main types of estimators for the spatial lag that have been extensively studied, the maximum likelihood (ML) or quasi-maximum likelihood (QML) as put forward among others by Anselin (1988b) and the instrumental variable approach (Kelejian and Prucha, 1998). In this chapter we rely on the ML estimation, but the instrumental variables (IV) estimation will be discussed in more detail in the spatial quantile regression context, in chapter 4.

(2009) and Brady and Irwin (2011) present an overview on how to derive the marginal prices of different explanatory variables in the maximum likelihood context[14].

Spatial error model

The second case of spatial dependence, if there are spatially autocorrelated variables omitted from the model, can be addressed by the spatial error model (SEM). Thus, the SEM is appropriate if some unobserved feature is assumed to lead to a spatially correlated pattern in the error term (Ward and Gleditsch, 2008). It is considered "*the more natural*" way to consider spatial effects in hedonic modelling (Anselin and Lozano-Gracia, 2009, p.1221). In the spatial error model, which is conceptually simpler than the SAR model, "*the only problems are heteroskedasticity and non-linearity*" (Viton, 2010, p.11)[15]. The spatial error model takes the following form:

$$Y = X\beta + \epsilon \tag{3.6}$$

$$\epsilon = \lambda W \epsilon + u \tag{3.7}$$

The spatial error model decomposes the error term (ϵ) from the linear model (3.6) to λ, the spatial autoregressive coefficient and the independent and identically distributed error term (u). With λ indicating the extent to which the spatial component of the errors are correlated based on the predefined spatial weights matrix (W) (Ward and Gleditsch, 2008)[16]. The interpretation of the regression results is identical to OLS estimation results.

[14] The marginal implicit price of an explanatory variable in the SAR or SAC context, can be derived if the "spatial multiplier effect" is accounted for (3.5). Where Y is the price (e.g.: mean or median sample price), an ρ the estimate of the spatial autoregressive coefficient (Anselin and Lozano-Gracia, 2009, p1244).

$$\text{MWTP}_{\beta_i} = \beta_i \text{Y} \times \frac{1}{1-\rho} \tag{3.5}$$

[15] Further discussed in appendix D.2.

[16] For the estimation of the spatial models we relied on the tools provided in R (R Core Team, 2013) in the "*spdep*" package (Bivand, 2013).

Spatial simultaneous autoregressive model

Eventually, in the case where both forms of spatial dependence are detected, the spatial simultaneous autoregressive model (SAC) (Anselin, 1988b) incorporates the spatial structure of the W matrix, to the dependent variables as well as the error term. The dependent variable (Y) is thus, as in the SAR, a function of the independent variables as well as of the Y of the neighbouring sections, measured by ρ. Meanwhile, the spatial autoregressive structure that might exist among the error terms is accounted for as in the SEM by λ[17]. As a combination of SAR and SEM model, it takes the following form:

$$Y = \rho WY + X + \epsilon \tag{3.8}$$

$$\epsilon = \lambda W\epsilon + u \tag{3.9}$$

While accounting for spatial residual autocorrelation the SAC provides information on the existence and importance of spatial spillovers between the residential land prices in the defined neighbourhood as well as spatial autocorrelation of the errors.

Testing for spatial dependence and implementing the appropriate estimation methods is thus necessary to account for potential estimation bias, as a non-spatial model might underestimate the marginal effects of location-specific attributes (Brady and Irwin, 2011). The procedure we follow to identify spatial dependence is essentially based on the works of Anselin (1988b); Le Sage (1999); Anselin (2002) and (Elhorst, 2010). Different tests have been developed (i.e. Moran's I, LM-test) to identify the presence and form of spatial dependence (D.4)[18].

3.3 Precisions on data and the SWM

The expectations towards the explanatory variables presented in part I will be briefly recalled in the following sections.

[17] As for the previous model, it is necessary to rely on maximum likelihood estimation to obtain unbiased results (Wilhelmsson, 2002).
[18] Further presented in appendix D.4.

3.3.1 Transaction-specific attributes

After some clean-up, the developable land transaction dataset from the AED[19], includes 6,367 observations. As presented in section 1.2.1 these prices were deflated to the value of January 2007-Euros. For ease of interpretation, some explanatory variables have been mean centred and eventually the intercept represents the price of a typical transaction of mean size, located at mean distance to the centre and with average net population density in the section. Summary statistics are presented in table 3.1.

The AED dataset distinguishes between different types of transactions, and some assumptions were made to account for possible differences in actors. First, *BigSize* identifies parcels larger than 22.77 are, corresponding to 2% of the transactions, that are assumed to be rather bought by professional developers, with the purpose of future development and/or resale (speculation). All else constant, the effect on price is expected to be positive. Second, a binary variable identifying transactions sold with existing development project (*dVFA*) is considered to be in general sold by a professional developer to an individual consumer. A positive marginal effect is expected as such projects imply some previous investments and some amenities related to the project implementation (carefree package). Third, *dTerrain* informs if a parcel is sold separately from some existing non-residential buildings. The presence of such constructions is expected to have a negative impact on prices, since additional fees are likely related to the destruction of these buildings. Eventually, the interaction term between *dVFA* and *dTerrain* is considered to investigate the impact on the overall transaction price if a *dVFA* transaction was at the same time occupied by a ruin. By including these dummies, we assume that a typical transaction would in general be sold by private landowners to either other private households or real estate developers (specialised in smaller development projects, mainly for residential use; e.g.: apartment buildings, terraced houses).

[19] A detailed presentation of the dataset can be found in part I chapter 1 and appendix C.1.

3.3.2 Location-specific variables

The price of developable land is also expected to be largely determined by local neighbourhood amenities[20]: accessibility to Luxembourg-city, local green and local urban amenities.

Time-distance by car via road network (lntLUX_{ci}) was generated to approximate the distance as perceived by individuals[21]. In addition, travel time by public transport (lntLUX_{pi}) is considered to identify to what extent the different transport means are valued by land consumers. A non linear relationship is expected from the distance variables and so their natural log was considered. Luxembourg-city is at the centre of a monocentric organisation of the country and good access to this employment centre is assumed a premium for developable land consumers and hence a major determinant of land price. Increasing distance to Luxembourg-city is expected to have a negative marginal effect on the price of land.

To control for the marginal price consumers are willing to pay for the presence of local urban amenities, the amount of retail and service opportunities ($SSopp$) will be considered. Further, as suggested by Youssoufi (2011), a diversity index of this supply at section scale (DI) was generated. The interaction between both variables (DI*$SSopp$) should provide additional insights into the impact of an increased amount of opportunities on the valuation of diversity. Both, the diversity and the supply of retail and service opportunities are expected to be valued positively, while their interaction term is expected negative.

To identify and value the impact of different green amenities on residential land prices in Luxembourg, different land-use measures and diversity indexes have been generated at section scale, based on the datasets provided by the Administration du Cadastre et de la Topographie (2008a,b)[22].

[20]The results of Hoyt and Rosenthal (1997, p.175) "*unambiguously refute any theory that would imply that (marginal) benefits from locational amenities and public services are correlated with the structural characteristics of individuals homes at street level*".

[21]Detailed presentation of this data can be found in section 2.1.

[22]A detailed description of the dataset and the related literature was given in part I section 2.3. The results presented here are based on a large number of tests on different land-use variables at different

To account for the marginal effect of developable open-space within a section, the vacancy rate (*rAP*) was considered with regard to the total residential area (Kiel and Zabel, 2008; Liu and Hite, 2013). The impact of an increase in the proportion of vacant land might either have a negative impact (Irwin and Bockstael, 2001; Geoghegan, 2002; Geoghegan et al., 2003) or a positive one (Ooi and Le, 2013).

A main challenge was to capture local effects at aggregated scale, especially since it was shown that consumers perceive green amenities generally at a very local scale (Cavailhès et al., 2006; Goffette-Nagot et al., 2011). The land-use around *AP* was thus considered[23]. The proportion covered in a certain land-use in the neighbourhood of 100 metres around all AP of a section has been generated. The 100m extent should capture the immediate proximity around all AP in a section, intending to approximate the close neighbourhood around the transactions as it was considered in several micro-scale hedonic studies (Geoghegan et al., 1997; Kestens et al., 2004; Kadish and Netusil, 2012). Eventually, the part of the neighbourhood (NH) covered by brushwood (*rsAPbw100*) and forest (*rsAPfor100*), rivers (*rsAPriv100*) and horticulture (*rsAPhorti100*) were considered. On the one hand, we expect increased land covered in brushwood to have a negative impact on residential land transactions, as blocking view and natural barrier. On the other hand, neighbourhoods with access to water flows, forest and horticultural land-use (gardening and vineyards) are rather expected to be valued positively, for instance for their recreational opportunities and picturesque landscapes.

The Shannon land-use diversity index was generated considering two extents: 100m and 1,000m around *AP*. According to Geoghegan et al. (1997), the 100m radius is expected to capture a negative effect of land-use diversity in the immediate neighbourhood and the 1,000m radius a positive effect of increased diversity in walking distance to *AP*. Further, the diversity index for green land-uses at section scale (*DIGreen*) should allow insights into the preferences of developable land consumers for the general diversity of green land-uses at section scale.

extents, considered to approximate local green amenities.

[23] See C.3 for the different steps.

In this chapter, the socio-economic neighbourhood characteristics are based on the data provided by the IGSS, described in chapter 2.4. Net population density (*NetPopDens*) and the population variation rate between 2001 and 2007 (*VarPop0107*) should allow to measure relative scarcity of residential land and past demand within a section. The part of population younger than 15 years (*rPopbelow15*) was added as a proxy for the social composition of a section. The proportion of high income *rHighIncome* residents was considered to approximate the economic status of a sections population. In the same perspective, the unemployment rate (*rUnemployment*) provided by STATEC (2008)[24], was considered at municipal scale.

3.3.3 Spatial weight matrix

Spatial econometric techniques require the definition of a spatial weight matrix to define the spatial relationships between transactions. Its specification has been largely discussed in hedonic literature and details are presented in appendix E.1. Spatial econometricians generally turn to an exogenously determined spatial weights matrix (W) of dimension *nxn*, 6,367 x 6,367 in our case, (De Graaff et al., 2001), which specifies a neighbourhood set for each observation. For each row i a positive weight (w_{ij}) identifies j as a neighbour of i. Observations are not neighbours to themselves ($w_{ii} = 0$) and the diagonal elements are thus 0. Typically the spatial weight matrix is row-standardized (where each element is between 0 and 1), to allow more comparable estimation results and to ease their interpretation (Anselin, 2002).

In this thesis, different spatial weight matrices have been tested to capture potential spill-over effects between transactions at section level. The aggregated location of transactions in space implies that within a section it was impossible to differentiate the spatial relationships among observations. With regard to these limitations a contiguity matrix as well as inverse distance matrices were considered, as presented in more detail in appendix E.2 and table E.1.

[24]As we were not able to generate this information from the IGSS dataset in accordance with the definition of active population generally considered in Luxembourg.

Variable	Description	Min.	Median	Mean	Max.	Unit	Impact
Price	Total transaction price	14,458	255,250	257,816	2,870,509	€	DV
lnSize$^+$	Total size of transaction	-0.46	1.56	1.58	3.92	are	+
BigSize	1 if size $\succ$ 22.77 are	0.00	0.00	0.02	1.00	dummy	+
dVFA	Transaction sold with development project	0.00	0.00	0.19	1.00	dummy	+
dTerrain	Complex transactions with land sold apart	0.00	0.00	0.08	1.00	dummy	-
lntlux$^+_{ci}$	Time to Luxembourg-city by car	1.51	3.16	3.23	4.35	min	-
lntlux$^+_{pi}$	Time to Luxembourg-city by public transport	2.07	3.64	3.63	4.80	min	-
DI	Retail and service diversity	0.00	0.79	0.64	0.91	index	+
SSopp	Retail and service opportunities	0.00	9.00	19.16	266.00	count	+
rAP	Vacancy rate	0.72	10.34	11.41	35.96	%	-/+
rAPbw100	Part of brushwood in 100m	0.00	0.00	0.39	18.46	%	-
rAPfor100	Part of forest in 100m	0.00	1.85	3.19	34.37	%	-/+
rAPriv100	Part of rivers in 100m	0.00	0.00	0.10	4.54	%	+
rAPhorti100	Part of horticulture in 100m	0.00	1.15	2.02	36.96	%	+
DIGreen	Green land-use diversity	0.08	0.58	0.57	0.75	index	-/+
mAPsh100	Mean Shannon diversity 100m	0.99	1.27	1.29	1.69	index	-
mAPsh1000	Mean Shannon diversity 1000m	0.85	1.65	1.63	2.07	index	+
lnNetDens07$^+$	Net population density	0.06	7.12	7.04	8.67	hab/km^2	-/+
rBelow15y	Part of pop. below 15 years	0.00	19.66	20.03	42.11	%	+
rHighIncome	Part of pop. above 5 SSM	0.00	3.41	3.75	11.28	%	+
VarPop	Population variation between 2001 and 2007	-33.33	9.55	11.53	131.80	%	-/+
rUnemployment	Unemployment rate	2.07	4.04	4.33	9.92	%	-
Note:	$^+$ grand-mean centred variables						

Table 3.1: Summary statistics: chapter 3

3.4 Global model results

Eventually the main models and results will be presented, staring with the specification of the hedonic model and further the interpretation of the spatial estimation results. Step by step the different explanatory variables have been added to the hedonic model, OLS estimation results are summarised in table 3.2.

Model (1) includes the structural characteristics, size and legal framework, providing highly significant estimates, confirming our expectations. Both, time to Luxembourg-city by car (lntLUX_{ci}) and by public transport (lntLUX_{pi}), added in **model (2)**, show significant negative marginal effects on the price of a typical transaction. Including the accessibility variables considerably adds explanatory power to the model, increasing the adjusted R^2 from 0.29 to 0.46. The average price of a typical transaction, all else held constant, decreases from 288,292€ to 283,225€[25]. The observed variations in the estimates of the structural variables are related to the grand-mean centring of the distance measures to Luxembourg-city. It is now rather located in the periurban area, and we observe a partly increased marginal effect for an above average sized parcel. Adding the socio-economic context control variables in **model (3)**, modifies the characteristics of a typical transaction, since *lnNetDens* is mean centred. At this stage, a typical transaction considers an average sized parcel located in a section at average time-distance from the city centre and with average net density, all else held constant. Controlling for the socio-economic context at section scale slightly improves the model fit. It considerably lowers the marginal effect of accessibility measures, suggesting that the latter captured some agglomeration effect in the previous model. In **model (4)**, we aimed at capturing the effect of local urban amenities on land prices, by including the retail and service diversity *(DI)* as well as the amount of opportunities (*SSopp*). While the diversity index is of low significance, the count of the different opportunities for shopping and services is very significant and positive. As expected, the interaction term (*DI:SSopp*) has a negative effect, suggesting that with increasing opportunities, diversity is valued less positively. Next, the proportions of different green land-uses in the 100m neighbour-

[25] Applying the method suggested by Verbeek (2008) adding the half variance term to the intercept.

hood around AP were added in **model (5)**. Although the R^2 does not further increase, AIC confirms an improvement of the model fit. All else held constant, the effects of these different proxies for neighbourhood green have the expected signs, especially the positive effects of proximity to water flows and horticulture are confirmed. Eventually, the land-use diversity measures are included in the model specification (**model (6)**). A negative impact is estimated for *DIGreen*, suggesting that more diverse land-uses are valued negatively at section scale. The negative estimates of a section's green diversity capturing most likely some "Oesling"[26] effect, since this area registers particularly high green land-use diversity (map C fig.2.13), which might also translate some additional distance effect. With regard to the findings of Geoghegan et al. (1997), the positive effect of increased land-use diversity within walking distance (*mAPsh1000m*) is confirmed, while *mAPsh100m* is not significant.

The AIC indicates that considering the land-use and green diversity variables improves the model fit. However, the adjusted R^2 reveals that still only about 48% of the price is explained, indicating that there is a substantial part of price variance is not captured by the specified model. Most probably this is related to missing structural information and the aggregated scale of the explanatory variables, considering the neighbourhood of all available plots rather than specific context of individual transactions. The Breusch-Pagan test (*BPtest*) indicates the rejection of the homoskedasticity assumption; therefore all results presented in table 3.2 consider heteroskedasticity robust standard errors according to White[27].

Eventually, spatial dependence was tested for considering the contiguity and inverse distance spatial weight matrices[28]. On the one hand, test results (table 3.3) suggest strong spatial autocorrelation among the residuals. On the other hand spatial lag dependence is rejected by the robust lag LM-test in the first place. The SARMA test for spatial autocorrelation in the dependent variable and the residuals is highly significant, suggesting that there might be spatial lag dependence left after accounting for spatial

[26]The extend of the "Oesling" region being illustrated in map H.2 p.260.
[27]By "*coeftest*" provided in package "*lmtest*" by Zeileis and Hothorn (2002) in R Core Team (2013).
[28]Presented and discussed in E.2.

	Dependent variable: lnPrice					
	(1)	(2)	(3)	(4)	(5)	(6)
Intercept	12.398*** (0.008)	12.423*** (0.007)	12.391*** (0.055)	12.323*** (0.058)	12.210*** (0.063)	12.256*** (0.139)
lnSize+	0.518*** (0.013)	0.579*** (0.011)	0.588*** (0.011)	0.592*** (0.011)	0.593*** (0.011)	0.594*** (0.011)
BigSize	0.315*** (0.058)	0.105** (0.050)	0.096* (0.050)	0.073 (0.050)	0.072 (0.050)	0.066 (0.050)
dVFA	0.448*** (0.022)	0.334*** (0.019)	0.343*** (0.019)	0.351*** (0.019)	0.346*** (0.019)	0.342*** (0.019)
dTerrain	−0.122*** (0.046)	−0.164*** (0.040)	−0.170*** (0.039)	−0.163*** (0.039)	−0.172*** (0.039)	−0.172*** (0.039)
dVFA:dTerrain	−0.474*** (0.059)	−0.377*** (0.051)	−0.376*** (0.051)	−0.392*** (0.051)	−0.378*** (0.051)	−0.376*** (0.051)
$lntlux^{+}_{ci}$		−0.511*** (0.024)	−0.398*** (0.027)	−0.331*** (0.028)	−0.336*** (0.029)	−0.317*** (0.030)
$lntlux^{+}_{pi}$		−0.103*** (0.028)	−0.067** (0.029)	−0.074** (0.029)	−0.094*** (0.030)	−0.096*** (0.030)
lnNetDens+			0.085*** (0.010)	0.072*** (0.011)	0.049*** (0.012)	0.044*** (0.012)
rBelow15			−0.009*** (0.002)	−0.007*** (0.002)	−0.006** (0.002)	−0.004* (0.002)
rHighIncome			0.031*** (0.004)	0.033*** (0.004)	0.028*** (0.005)	0.024*** (0.005)
txVarPop			−0.0003 (0.001)	0.0001 (0.001)	0.0004 (0.001)	0.0004 (0.001)
rUnemployment			0.021*** (0.004)	0.004 (0.005)	0.007 (0.005)	0.0001 (0.006)
DI				0.043* (0.025)	0.054** (0.025)	0.018 (0.027)
SSopp				0.040*** (0.007)	0.042*** (0.007)	0.040*** (0.007)
DI:SSopp				−0.043*** (0.008)	−0.044*** (0.008)	−0.042*** (0.008)
rAP					0.005*** (0.001)	0.005*** (0.001)
rAPbw100					−0.014*** (0.005)	−0.014*** (0.005)
rAPfor100					0.003* (0.002)	0.004** (0.002)
rAPriv100					0.070*** (0.021)	0.058*** (0.021)
rAPhorti100					0.006*** (0.002)	0.006*** (0.002)
DIGreen						−0.146* (0.083)
mAPsh100						−0.113 (0.076)
mAPsh1000						0.137*** (0.041)
Adjusted R^2	0.29	0.46	0.47	0.48	0.48	0.48
AIC	11,347	9,564	9,447	9,389	9,364	9,355
BPtest	403.78***	672.22***	689.38***	690.02***	720.18***	725.18***
Notes:	+ Grand-mean centred variables; (heteroskedasticity robust std. errors) Significance codes: 0 '***'; 0.001 '**'; 0.01 '*'; 0.05 '.'					

Table 3.2: OLS Regression Results

	Model (6)		
	Contiguity	ID3838	ID5000
LMerr	105.91***	145.45***	153.09***
RLMerr	51.82***	67.37***	92.69***
Lag	54.85***	81.43***	61.30***
RLMlag	0.76	3.35	0.9
SARMA	106.67***	148.80***	153.99***

Table 3.3: LM-test for spatial autocorrelation

autocorrelation among the residuals. As presented in section 3.3.3, three spatial weight matrices have been considered and table 3.3 shows that similar conclusions can be derived whatever matrix considered. The SEM and SAC models were estimated with either matrix. The results do not substantially vary and thus we relied on the the AIC suggesting that the spatial contiguity matrix fits the model best. The spatial models thus account for spatial dependence among the observations located within a same section as well as with transactions located in the contiguous neighbouring sections.

Summary of model specification results:

- 48% of price variance explained (omitted variables).
- The expectations derived from urban economic theory are confirmed.
- The hypothesis of spatial lag dependence is rejected, while residual spatial autocorrelation is confirmed.
- The contiguity matrix fits our data best.

The results of spatial dependence tests suggest the implementation of the spatial error model (SEM) and the spatial simultaneous autoregressive model (SAC)[29]. Nevertheless, the SAC estimation results (**model (8)** in table 3.4) confirm insignificant spatial lag dependence even after error dependence is accounted for. Therefore we rely on the SEM results.

Compared to **model (6)** there have been no substantial changes in attribute significance, magnitude or signs of the structural, accessibility, socio-economic context and local urban amenity variables in **model (7)**. The main marginal effects are presented

[29] The tools to implement the spatial models are provided by the "*spdep*" package (Bivand, 2013) in R (R Core Team, 2013).

	Dependent variable: lnprice	
	(SEM) (7)	(SAC) (8)
Intercept	12.306^{***}	12.045^{***}
	(0.161)	(0.161)
lnSize^{+}	0.600^{***}	0.600^{***}
	(0.02)	(0.015)
BigSize	0.016	0.053
	(0.03)	(0.055)
dVFA	0.344^{***}	0.344^{***}
	(0.017)	(0.017)
dTerrain	-0.179^{***}	-0.179^{***}
	(0.027)	(0.027)
dvfa:dterrain	-0.364^{***}	-0.365^{***}
	(0.035)	(0.035)
lntLUX^{+}_{ci}	-0.317^{***}	-0.310^{***}
	(0.044)	(0.044)
lntLUX^{+}_{pi}	-0.106^{**}	-0.104^{**}
	(0.039)	(0.039)
lnNetDens^{+}	0.047^{***}	0.047^{***}
	(0.013)	(0.013)
rBelow15	-0.007^{*}	-0.007^{*}
	(0.003)	(0.003)
rHighIncome	0.028^{***}	0.027^{***}
	(0.006)	(0.006)
txVarPop	0.001	0.001
	(0.001)	(0.001)
rUnemployment	0.014^{*}	0.014^{*}
	(0.007)	(0.007)
DI	0.020	0.020
	(0.029)	(0.029)
Ssopp	0.041^{***}	0.041^{***}
	(0.009)	(0.009)
DI:SSopp	-0.044^{***}	-0.044^{***}
	(0.010)	(0.010)
rAP	0.004^{**}	0.004^{**}
	(0.002)	(0.002)
rAPbw100	−0.005	−0.005
	(0.006)	(0.006)
rAPfor100	0.003.	0.003.
	(0.002)	(0.002)
rAPriv100	0.042	0.041
	(0.026)	(0.026)
rAPhorti100	0.004	0.004
	(0.003)	(0.003)
DIGreen	0.030	0.027
	(0.108)	(0.108)
mAPsh100	−0.144.	−0.144.
	(0.087)	(0.087)
mAPsh1000	0.071	0.071
	(0.051)	(0.051)
ρ	-	0.020
	-	(0.090)
λ	0.488^{***}	0.475^{***}
	′(0.049)	(0.074)
AIC	9,291	9,293
BPtest	726.75^{***}	727.49^{***}
Notes:	$^{+}$ Grand-mean centred variables; (heteroskedasticity robust std. errors); Significance codes: 0 '***'; 0.001 '**'; 0.01 '*'; 0.05 '.'	

Table 3.4: Spatial model results

in table 3.5. The predicted mean price of a transaction of mean size, located at mean distance to the capital and within an average densely populated section, is estimated at 247,822€, not considering spatial error dependence thus underestimated the price of a typical transaction (being of 238,573€ in the non-spatial model). All else held constant, additional 10% parcel size (almost 0.5 are), is estimated to increase the transaction price of about 60%. No significant impact of transactions of very large size could be identified, suggesting that the assumed speculation effects related to these very large parcels does not significantly impact on land values. However, transactions sold with development plans are in general dearer to land consumers, increasing their value by about 41%. The negative impact of additional 2.5 minutes (+10%) of individual transport time to the capital decreases the price by 32% while the same increase in travel-time by public transport (3.8min) has an impact of -11%. On the one hand an increased number of retail and service opportunities (*SSopp*) is valued positively, while on the other hand the diversity index turned non-significant after including the land-use diversity measures to the model specification.

The variables aiming at capturing the socio-economic context show in general the expected signs[30], underlining the importance of the local socio-economic context specificities. However it is worth noting that a 1% increase in net density above average causes an estimated impact of 4.70% on residential land prices[31], further results confirm that sections with higher proportion of high income population have a positive impact on land prices.

The proxies for the green neighbourhood attributes turned insignificant after accounting for spatial error dependence; except for the proportion of area covered by forest in a

[30]The positive effect of unemployment rate does not confirm results presented in hedonic literature, as they generally estimated to be valued negatively (Chasco and Le Gallo, 2013). Unemployment rates are in general quite low in Luxembourg (in 2007 on average 4.67% of the active population was unemployed according to STATEC, 4.33% within municipalities registering residential land transactions), the generally associated negative effects (e.g.: higher criminality) might not be directly perceived by land consumers or as strong as in other case studies.

[31]It might be important to remind the scale of analysis, the average population density at section scale does not allow conclusions on the valuation of population density at micro-level, that might be perceived differently by residential land consumers. At sub-municipal scale it might still capture an urban effect, and stand for the scarcity of developable land and high price (Geoghegan et al., 2003; Goffette-Nagot et al., 2011) in densely urbanised sections.

100m radius. Insignificant results are obtained for green land-use diversity and land-use diversity in walking distance. While the negative impact of mean Shannon diversity index in 100m turned out significant and decreasing the price of a typical transaction about 14%. The marginal effect of a higher vacancy rate (rAP) is valued positively, although the effect on price is rather small.

Main findings:

- Spatial error model is needed to account for spatial autocorrelation.
- Poor performance of proxies for green neighbourhood effects are observed.
- Insignificant estimates for local urban amenity and green land-use diversity are obtained.
- The Shannon land-use diversity measures show the expected signs.

Variable	Marginal effects
lnSize (+10%)	60.00%
dVFA	41.06%
lntLUX$_{ci}$ (+10%)	−31.70%
lntLUX$_{pi}$ (+10%)	−10.60%
lnNetPopDens (+10%)	4.70%
rHighIncome (+1%)	2.80%
SSopp (+1)	4.10%
rAP (+1%)	0.40%
mAPsh100 (+1)	−14.40%

Table 3.5: Marginal effects (**model (7)**)

3.5 Concluding remarks

The discussed results allow several general conclusions on the model specification, testing and accounting for spatial dependence and eventually the geographical determinants of land values in Luxembourg. The hypothesis of spatial lag autocorrelation in the global model is rejected at this global level. These findings suggest that residential land prices in Luxembourg are most likely not autocorrelated with the transaction prices observed within the same section and the neighbouring ones. Nevertheless, the possibility of spatial lag dependence at the more local scale (e.g.: through the different neighbourhoods of a village or within the quarter of a town) cannot be excluded. Since at micro-scale land prices, especially of transactions sold in the framework of larger development projects are likely to be fixed by the same landowner. Similar valuation practices might apply at this scale, that should be further investigated.

The different explanatory variables added step by step allowed to improve the model's explanatory power. The basic trade-off is confirmed, access to Luxembourg and the structural variables had the biggest impact on explaining prices. The additional explanatory variables, the periurban effects, only add little to the overall explanation power of the model. However, they allow more differentiated insights into the determinants of price and show that if omitted, they are mainly captured by the individual transport measures. This underlines the importance of considering these local geographical specificities to gain more differentiated insights into developable land consumers' preferences.

By grand-mean centring distance variables, the estimated marginal effects mostly correspond to a close periurban context of average net density, as identified in appendix B. The estimated intercept is likely to translate the average price paid for a typical transaction of periurban location.

The structural and transaction type variables show the expected impacts. It was shown that if the seller is a professional developer, this increases the price of a typical transaction as expected. Whereas, considering a dummy for transactions of very large size, was not significant. This finding suggests that, all else constant, that no speculative effect could be captured by including this variable.

Beside the importance of travel-time to Luxembourg by car, results further confirm a significant impact of travel-time by public transport. Nevertheless its low part in the modal split, consumers have preferences for sections better connected to the public transport network, travel-time by car held constant.

As discussed in part I, it was impossible to properly account for the local neighbourhood characteristics of transactions, be it urban amenities or green land-uses. Attempts were made to approximate these local effects at section level by including the aggregated counts and diversity measures. It was found that an increased amount of retail and service opportunities is valued positively. Even after accounting for population density, the traditional control for agglomeration effects, this variable captures some benefits obtained from an increased offer in retail and service opportunities.

By including the land-use diversity index in walking distance, the retail and service diversity index turned insignificant. This suggests that the diversity of land-use captures also some urban or centrality effect. But it can also be related to a scale effect, urban amenities might be rather considered at a more aggregated level. Furthermore the general good coverage in shopping malls throughout the country, assumed to regroup in general a diversified offer of products through the country, could be an explanation as well as related to the general high individual mobility.

Increased vacancy rate is valued positively, while its marginal effect on price is rather small. In opposition to previous findings (Irwin and Bockstael, 2001; Geoghegan, 2002; Geoghegan et al., 2003), this positive effect translates that land consumers are expecting positive effects from future developments in the same section (e.g.:increase in local urban amenities), or anticipating an increase of the value of their own properties, or that they simply are not aware of the eventual future conversion of this land and perceive it as permanent open-space. This should be further investigated, especially in light of future planning strategies considering infill development.

Although of low significance, the results of this model confirm the findings of Geoghegan et al. (1997) of a negative effect on consumers from more diverse land-uses in immediate proximity. The insignificant estimates of the proxies for different land-uses in the proximity of all available plots, after accounting for spatial dependence, suggests that they might have captured some unobserved spatial effects in the non-spatial model. As discussed in literature, consumers are expected to value green amenities at a very local scale (Geoghegan et al., 1997; Kestens et al., 2004; Cavailhès et al., 2006; Kestens et al., 2006; Kong et al., 2007; Cavailhès et al., 2009; Goffette-Nagot et al., 2011; Kadish and Netusil, 2012), which we tend to confirm. In the following chapters, these proxies were hence not considered.

Promising results have been found for the land-use diversity measures, that is why these will be further considered and investigated in the following chapters. Suggesting that this measure is not able to capture consumers' preferences for these specific land-uses at this aggregated scale and thus confirms that green amenities are rather valued at a very local scale and that hence the actual neighbourhood of transactions could not be approximated by the section averages.

The aim of this chapter was mostly to confirm the theoretical and empirical findings discussed throughout the introduction and the first part of this thesis. Hence, it aimed at identifying the significant structural and location specific variables, to account for the specific geographical context. Except for the variables discarded previously, most will be kept throughout the following chapters, although some modifications will be made to this global model.

Chapter 4

Varying preferences for local amenities with consumers' income and by market segments: Spatial quantile regression

4.1 Introduction

As discussed in the previous chapter, increasing interest has been turned to the valuation of open-spaces and the spatial land-uses pattern, as neighbourhood green amenities are considered major determinants of quality-of-life. Despite some limitations, we were able to account for some of these determinants and results suggest that beside the standard trade-off described in the monocentric model (Alonso, 1964; Fujita, 1989), local urban amenities and land-use diversity, the periurban effects (Cavailhès et al., 2004), are considered in the valuation of land in the Luxembourgish context.

Expanding periurban areas have been associated to the varying preferences of consumers based on households' composition and their economic background (LeRoy and Sonselie, 1983; Mieszkowski and Mills, 1993; Glaeser, 2008; Brueckner, 2011). The price consumers are willing to pay for different characteristics composing the land depend on the specific constraints these individuals and households face with regard to their economic status (e.g.: employment and income), the households' motorisation, the family structure and life cycle (Thériault et al., 2005). These observations could not be captured in the global model, presented in chapter 3. The method applied there assumed that all consumers share identical preferences, value the characteristics composing land in the same way and that they are in competition on the same market.

However, it was shown that the spatial pattern of exogenous amenities in a city impact on the location of different income groups, and that the valuation of these amenities rises rapidly with income (Brueckner et al., 1999). Further, landscape features are valued differently by different income groups. For example, Des Rosiers et al. (2007) found the higher the respondents' income, the higher the marginal value of high tree cover in front of the property. Liu and Hite (2013) observe that, while there is almost no variation in the negative effect of distance to forest, an increase in the general proportion of the census tract covered in forest is negatively perceived, especially by consumers of middle-priced houses. Further, they find that in general the green space effect was only significant in the middle and high price ranges. Differences in the utility obtained from various transport modes by different income groups have been theoretically discussed in LeRoy and Sonselie (1983); Glaeser (2008), generally associating higher dependence on public transport with lower income groups. Kestens et al. (2006) find for instance that with increasing income the negative effect of distance is less marked than for low-income households, findings that should be considered in the context of Quebec where teleworking is widespread. Differences in the perception and valuation of air pollution and noise have been investigated by Chasco and Sánchez Reyes (2012), stating that only consumers of the most expensive housing units significantly perceive these pollutions as negative externalities. This selection of literature illustrates the importance of variation that might be induced in the valuation of geographical determinants by the consumers' socio-economic background.

A further complication in this case study is related to the lack of information on the consumers (and sellers) of the transactions. Some assumptions were raised in part I (section 1.1.2), on the characteristics of private households purchasing land in the objective of building their home as well as on the implication of professional real estate developers (as sellers and consumers). Although no information on private consumers' income is available, they are assumed a rather homogeneous group[1] of individuals, since

[1]Compared to the consumers of all types real estate transactions without distinction (e.g.: houses, apartments and land). Based among others on the national census data (Fehlen et al., 2003; Berger, 2004; Allegrezza et al., 2014), it was observed that in Luxembourg, single-family homes are the most common form of housing. High home ownership is registered, varying with household socio-economic characteristics (e.g.: nationality, age). Hence we assume that individual land consumers mostly buy land

a substantial financial background is needed to afford land and housing construction in Luxembourg. Mainly three types of consumers have been assumed to act on the developable land market in Luxembourg. First, private households purchasing land for their individual housing needs, aiming at maximising their quality-of-life with regard to a budget constraint. Second, professional or public developers of residential or non-residential projects of different extents with the objective of maximising their profit. Third, individual consumers purchasing land for private use from real estate developers but with already existing development projects. Between these types of consumers, preferences for local amenities are expected to vary. And thus the question arises whether these different actors are in competition on the same market, or whether the data covers different consumer based market segments.

While the first step of the hedonic pricing method (Rosen, 1974) provides insights into the implicit value of the considered attributes composing the land price, it does not allow insights into the underlying demand for and supply of these characteristics (Palmquist, 1984). However, it was shown that the price a consumer is willing to pay, for residential land and the characteristics composing it, depends on several conditions: the consumers' income, individual preferences and the gained utility (Palmquist, 2003).

The second step of the hedonic method allows to estimate the demand for the different price determinants while accounting for socio-economic differences among the consumers (Rosen, 1974; Brown and Rosen, 1982). This demand can be considered as a function of its marginal price (estimated in the first stage), the prices of some related characteristics, the expenditure as well as some socio-economic factors (Palmquist, 1984). This second stage of the hedonic method has been implemented several times (Palmquist, 1984; Epple, 1987; Bartik, 1987; Cheshire and Sheppard, 1998; Day et al., 2006; Carruthers et al., 2010), but a limitation often faced is the unavailability of detailed information on consumers, due to data privacy concerns, experienced as well in this framework.

Quantile regression has been put forward as an alternative to this second step. It allows to statistically examine the extent to which housing characteristics are valued

in the perspective of building a single-family home.

differently across the distribution of housing prices (Zietz et al., 2008), under the assumption that the overall transaction price can be seen as a proxy for the individuals' purchasing power and hence their income. Recently, this method has been applied in the hedonic pricing context; focussing on the agricultural land market (Kostov, 2009; Uematsu and Mishra, 2012; Schreurs et al., 2013) or the housing price determinants (Zietz et al., 2008; Liao and Wang, 2012; Chasco and Sánchez Reyes, 2012; Liu and Hite, 2013). The expansion method was for instance considered by some to account for varying preferences based on consumers' context (Des Rosiers et al., 2002; Anderson and West, 2006; Kestens et al., 2006), nevertheless consumer heterogeneity is often neglected in hedonic pricing models. Spatial autocorrelation among transaction prices is generally related to spill-over processes among prices in a location, with prices influencing those of nearby transactions (Anselin, 1988b). If every effort is made to avoid problems of missing variables and incorrect functional form "*the spatial lag dependence in the hedonic model could indicate deviations from the assumptions of perfectly competitive markets*" (Kostov, 2009, p.57). Thus, the spatial quantile regression should provide insights into market segmentation based on consumers. This might allow further insights into differences between consumer types and their preferences for the determinants of developable land price.

The scope of this chapter is thus twofold. First, with regard to the data limitations, this method should provide a means to identify potential market segmentation depending on consumers. Second, results should allow insights into variations in the residential land consumers' marginal willingness to pay for different geographic determinants in Luxembourg depending on their socio-economic background.

This research is innovative in several ways. In general, the spatial quantile regression approach is still quite uncommon in the hedonic pricing context. To knowledge it has not yet been implemented in this residential land market context, considering local periurban amenities as well as the traditional trade-off. Further, an analysis of the different actors on the developable land market has, to the best of our knowledge, never been applied to the Luxembourgish context.

In section 4.2, the model specification and variables will be recalled in light with the expectations towards them. The spatial quantile approach is introduced in section 4.3 and the results is discussed in section 4.4 and section 4.5 concludes.

4.2 Precisions on data and expectations

4.2.1 Transaction-specific attributes

The AED developable land transaction dataset, covering in total 6,367 developable land transactions remains unchanged[2]. The natural logarithm of the total transaction price was considered as dependent variable, on the right-hand-side of the model parcel size, accessibility measures and net population density were log transformed to account for their non-linear relationship with transaction price.

Property size (*lnSize*) has been identified as the major structural determinant of land price. With increasing prices, the marginal effect of an increase in property size above average is expected to have a positive and decreasing impact on price (Furtado and Van Oort, 2010; Chasco and Sánchez Reyes, 2012). The transaction type variables are mostly kept, except for *BigSize*, which by construction applies only to the highest price ranges. The *dVFA* dummy, for transactions registered as sold with existing development projects, is expected to remain positive, its valuation throughout the price ranges is to be investigated, the same applies to the negative effect of *dTerrain*.

4.2.2 Location-specific variables

Accessibility measures to Luxembourg-city were included via the travel-time by road network (lntLUX_{ci}) and public transport (lntLUX_{pi}) variables presented in part I. Land prices are expected to be negatively related to an increase in travel-time by car, whereas for the time by public transport the negative impact is expected to be more important for the less wealthy consumers (LeRoy and Sonselie, 1983; Glaeser, 2008).

The behaviour of the vacancy rate (rAP) variable throughout the price ranges is to be investigated, the effect is expected to remain positive based on our previous findings.

[2] The model specification has slightly changed from the previous chapter; the neighbourhood land-use variables 100m around AP have been omitted, the focus being now turned to the different diversity indices.

In the quantile regression context, including the green land-use diversity index (*DI-Green*) is expected to provide further insights into its impact on prices identified in the global model and whether the negative sign persists through all price ranges. The mean Shannon diversity indices at two extents (100m :*mAPsh100* and 1,000m :*mAPsh1000*) around all available plots (*AP*) were kept. Following Geoghegan et al. (1997) and the findings of the previous chapter, the 100m radius is expected to have a negative effect on the price of a typical transaction, while the 1,000m radius should seize the positive effect of increased diversity in walking distance. Considering these land-use diversity indices and their behaviour through the different price ranges has to date not been undertaken in the quantile regression context. The results are expected to provide further insights into their valuation with regard to the consumers' purchasing power. It is assumed that more well-off consumers account more for these local green effects.

The marginal value of the diversity and availability of local urban amenities has been considered via the retail and service opportunities diversity index (*DI*) and their supply (*SSopp*). With regard to the insignificant coefficient of *DI* in the global model, more detailed insights are expected through the price ranges. Additional local urban services are expected to be valued positively with increasing price ranges (Brueckner et al., 1999).

Eventually, the socio-economic neighbourhood characteristics provided by the IGSS at section scale, to control for the local socio-economic context, remain the same as in the previous chapter.

4.3 Quantile regression method

As presented in the introduction, quantile regression is proposed to approximate the second step of the hedonic pricing method. It should allow to gain more detailed insights into the demand for the different land price determinants. The quantile regression method allows to model the heterogeneity in the marginal effects (Koenker, 2005)[3].

[3]The quantile regression technique has been introduced by Koenker and Bassett (1978), and detailed reviews of the estimation technique can be found among others in Buchinsky (1997); Koenker and Hallock (2001); Koenker (2005); McMillen (2013).

Variable	Unit	Mean	Min	Max	Impact	Increasing price...
lnPrice	€	12.456	9.579	14.873		DV
lnSize^{+}	are	1.58	-0.46	3.92	+	decreasing
dVFA	dummy	0.19	0.00	1.00	+	?
dTerrain	dummy	0.08	0.00	1.00	-	?
lntLUX^{+}_{ci}	min	3.23	1.51	4.35	-	increasing
lntLUX^{+}_{pi}	min	3.63	2.08	4.80	-	decreasing
DI	index	0.64	0.00	0.91	+	increasing
SSopp	count	19.16	0.00	266.00	+	increasing
rAP	%	11.41	0.72	35.96	+	?
DIGreen	index	0.57	0.08	0.75	+	increasing
mAPsh100	index	1.29	0.99	1.68	-	increasing
mAPsh1000	index	1.63	0.85	2.07	+	increasing
lnNetDens^{+}	hab/km^2	7.05	0.06	8.67	+	increasing
rBelow15	%	20.03	0.00	42.11	-	?
rHighIncome	%	3.75	0.00	11.28	+	increasing
VarPop	%	11.53	-33.33	131.82	+	?
rUnemployment	%	4.33	2.07	9.92	-	increasing
Note:	$^{+}$ grand-mean centred variables					

Table 4.1: Summary statistics: chapter 4

4.3.1 Linear quantile regression

Linear quantile regression first subdivides the sample according to the unconditional distribution of the response variable and then subsequently performs OLS for the subsamples. Hence quantile regression uses the full sample which avoids the truncation problem[4] that arises if OLS estimation is performed separately to the different price segments. Important advantages of quantile regression are generally associated with its superior capability in handling heteroskedasticity, outliers, and unobserved heterogeneity (Koenker, 2005). The linear quantile regression takes the following form:

$$Y = X\beta_\tau + u \tag{4.1}$$

The estimates (β) are τ dependent, τ defining the observed quantile of the price distribution. The coefficients are thus allowed to vary with the quantiles, rather than being assumed fixed (Kostov, 2013a). Although quantile regression can not fully guarantee to eliminate potential impacts of the functional form assumption on spatial dependence, it will alleviate them (Kostov, 2009). Further, quantile estimators are much less sensitive to outliers as no distributional assumptions concerning the residuals (u) are made (Kostov, 2009).

4.3.2 Spatial quantile model

As discussed in the previous chapter and this part's introduction, OLS estimates might be biased in case of spatial autocorrelation (errors or/and dependent variable) and different spatial models have been put forward to account for either type of spatial dependence (mainly error or/and lag). In chapter 3, doubts on the absence of spatial lag dependence emerged by the significant test results for the SAC model. Eventually the SAC and the SAR were rejected and spatial error dependence accounted for via the spatial error model (SEM).

Quantile estimates are most likely not biased by spatial error dependence as Kostov

[4] Due to he sample selection bias discussed by Heckman (1979) and highlighted by Koenker and Hallock (2001) as such segmentation of the response is "*doomed to failure*".

(2009) claim that spatial clustering is not necessarily inconsistent here, since in the quantile regression context no distributional assumptions are made on the error term. This was doubted by Chasco and Sánchez Reyes (2012), who claim that the quantile regression method might not always or only partly allow to correct for spatial error dependence[5]. However, the spatial error model has to date not been implemented in the quantile context (Liao and Wang, 2012). Spatial lag dependence is most likely an issue within the different price ranges (Kostov, 2009).

A spatial lag quantile model has been proposed to account for this spatial spill-over effects within the different quantiles, similar to the SAR presented in chapter 3 and formalised in equation (3.4) (p.105). The SAR completes the linear hedonic equation, by introducing the spatial lag operator (ρ), the spatial weight matrix W and WY, the spatially lagged dependent variable and allows conclusions on the marginal effect of spatial dependence on the dependent variable.

The spatial lag parameter (ρWY) at the right-hand-side of the equation might lead to endogeneity issues that have to be addressed. There are two main techniques to account for this endogeneity either via maximum likelihood estimation (ML) (Anselin, 1988b) or the spatial two-stage least squares procedure (S2SLS) by Kelejian and Prucha (1998). The S2SLS has been widely implemented in the hedonic pricing context and extended to the quantile regression context. Usually the first order spatial lags of the explanatory variables WX are used as instruments (McMillen, 2013).

Different methods have been presented to implement the quantile regression generalisation of the spatial two-stage-least-squares estimator (S2SLS). First, the two step method presented by Kim and Muller (2004), implemented for instance by Zietz et al. (2008) and Liao and Wang (2012). Second, the instrumental quantile regression based on the works of Chernozhukov and Hansen (2006) and extended by Su and Yang (2007) to allow for instrumental variables quantile regression (IVQR)[6]. An implementation of

[5]Recent attempts have been put forward to handle error clustering via Bayesian methods, an overview can be found in Lum and Gelfand (2012). These have however not been considered in this thesis.

[6]The instrumental variable approach has some advantages on the maximum likelihood approach; as it is much less computationally demanding than the ML estimation, mainly because the nxn matrix has not to be inverted McMillen (2013).

the spatial IVQR can be found in Kostov (2009); Furtado and Van Oort (2010); Chasco and Sánchez Reyes (2012); McMillen (2013). Third, the smoothed empirical likelihood estimation presented in Whang (2006); Kostov (2013b), defined as a non-parametric analogue of the likelihood estimation (Kostov, 2013b).

In this chapter, it is relied on the second approach because it is more robust than the first (McMillen, 2013), and since the third is not yet implemented in the used estimation software[7]. As instruments, the spatially lagged explanatory variables (WX) are used as suggested. The spatial quantile generalisation of the linear spatial lag model is formalised as follows:

$$Y = \rho_\tau WY + X\beta_\tau + u \tag{4.2}$$

The spatial quantile estimation allows for a different degree of spatial dependence at different points (τ) of the response distribution. Chasco and Sánchez Reyes (2012) and McMillen (2013) provide a comprehensive overview of the different estimation steps.

As suggested among others by Kim et al. (2003); Anselin and Lozano-Gracia (2009); Abelairas-Etxebarria and Inma (2012), to obtain the marginal effects for every decile the spatial multiplier ρ has to be accounted for. In the quantile regression context, ρ_τ is τ dependent and hence the estimated coefficients per decile were multiplied by the respective spatial multiplier[8].

Kostov (2009) estimated the complete quantile process (as many quantile regressions as observations)[9]. Others considered the quartiles (.25, .5, .75,) (Buchinsky, 1997; Furtado and Van Oort, 2010; Uematsu and Mishra, 2012; Mueller and Loomis, 2013). However, considering a range of deciles (.1 - .9) is the prevailing practice in empirical studies (Zietz et al., 2008; Ebru, 2009; Farmer and Lipscomb, 2010; Furtado, 2011a; Chasco and Sánchez Reyes, 2012; Liu and Hite, 2013), and is therefore considered within this study.

[7] It would have exceeded our programming skills to implement it on our own.

[8] Similar to equation 3.5 (p.106), the marginal willingness to pay per decile can be computed as follows:

$$\text{MWTP}_\tau = \beta_{i_\tau} \times \frac{1}{1-\rho_\tau} \tag{4.3}$$

[9] Possible since a small dataset of 197 observations.

The spatial weight matrix (W)[10] considered in this chapter is the contiguity weight matrix identified in chapter 3.

4.4 Spatial quantile regression results

First, a global model was specified and estimated via OLS, based on which spatial dependence tests have been performed (table 4.2). It reveals similar results as in chapter 3 and identifies significant spatial autocorrelation, except for the robust RLMlag.

Lagrange Multiplier test on Global Model (OLS)				
	LMerr	RLMerr	LMlag	RLMlag
Lagrange Multiplier test	171.16	91.59	79.63	0.06
p-value	0.000	0.000	0.000	0.807

Table 4.2: LM-test results

As announced, by estimating the spatial quantile model, we aimed at identifying whether there is significant spatial autocorrelation within the transactions of the different price ranges[11]. The detailed regression results can be found in appendix F (tables F.1 and F.2). The spatial two-stage least squares model (S2SLS) results reveal a significant spatial lag parameter (ρ) for almost all deciles[12].

After accounting for the spatial multiplier, variations in the marginal effects are observed between the spatial and non-spatial models. The marginal effects for *lnSize*, *dVFA*, *NetPopDens*, *SSopp*, *rAP* and *mAPsh1000*, were underestimated by the OLS and the linear quantile estimation (at least for a large part of the deciles), while the accessibility variables and part of high income residents remain similar. The opposite effect is observed for the marginal effect of *mAPsh100* where the significant coefficients

[10] A discussion on the use of the appropriate spatial weight matrix in a quantile regression context has been provided by Furtado and Van Oort (2010), focussing on a neighbourhood weight matrix. More recently Kostov (2013a) has presented a method to identify the appropriate SWM in the QR context. The questions related to the appropriate SWM being discussed in general in the hedonic pricing context in appendix E, with regard to the scale of the data and the scope of the thesis, this will not be further investigated.

[11] To ease the reading in this section only the main results will be displayed and discussed based on the graphics presented in figures 4.1, 4.2 and 4.3. The legend is found in figure 4.1; This illustration of the coefficients of the spatial models account for the spatial multiplier to be able to compare them with the quantile regression results estimated by QR.

[12] Results of the non-spatial model will thus not be further discussed here, except for means of comparison between the spatial and non-spatial approach, illustrated in the graphs.

were overestimated by QR. Variables not significant in the global models are revealed to have a significant impact on transaction prices of some deciles (*PopVar*, *rUnemployment*, *DI*, *DIGreen* and *mAPsh100*).

The estimated spatial lag parameters for each decile (ρ_τ)[13] are significant for all price ranges, except D7. In general, the estimated lag is close to or below the estimated lag in the S2SLS model, some variation of the magnitude of the spatial lag parameters over the different ranges of the price distribution is observed. While its value oscillates between 0.11 and 0.24 in D2 to D8, highest values are estimated for D1 (0.30) and D9 (0.27). Suggesting that prices of the most and least expensive transactions are more positively influenced by the values of neighbouring transactions, than in mid price ranges; all else constant (also observed by Kostov (2009); Liao and Wang (2012); Chasco and Sánchez Reyes (2012)).

As expected, an above average increase of parcel size of 1% has a significant positive effect through all the price ranges, decreasing with increasing land prices, especially in D1, while it is below the global spatial model estimate in the higher price ranges. All else constant, with increasing prices, consumers able to afford above median prices obtain less utility from a 1% increase in parcel size. The marginal effect of an existing development project (*dVFA*) is confirmed, its impact is particularly important on the price of transactions located in D1 and the middle price ranges. While its impact decreases sharply from D6 onwards and even turns negative in D9. A similar behaviour was observed for the marginal effect of *dTerrain*, especially in D8 and D9 (table F.2). Low impacts are thus observed within the highest deciles, turning even negative in D9, suggesting that more wealthy consumers obtain less or no benefit when purchasing land with existing development plans, whereas especially in the middle price ranges this is considered a premium.

The negative marginal effect of increasing distance to Luxembourg-city by individual transport is increasing with transaction prices (fig.4.1, *Time by...*). Although, for

[13] To find the optimum spatial lag parameter for every decile, the model was estimated for every value between -1 and 1 (interval of 0.1), as described by Chasco and Sánchez Reyes (2012).

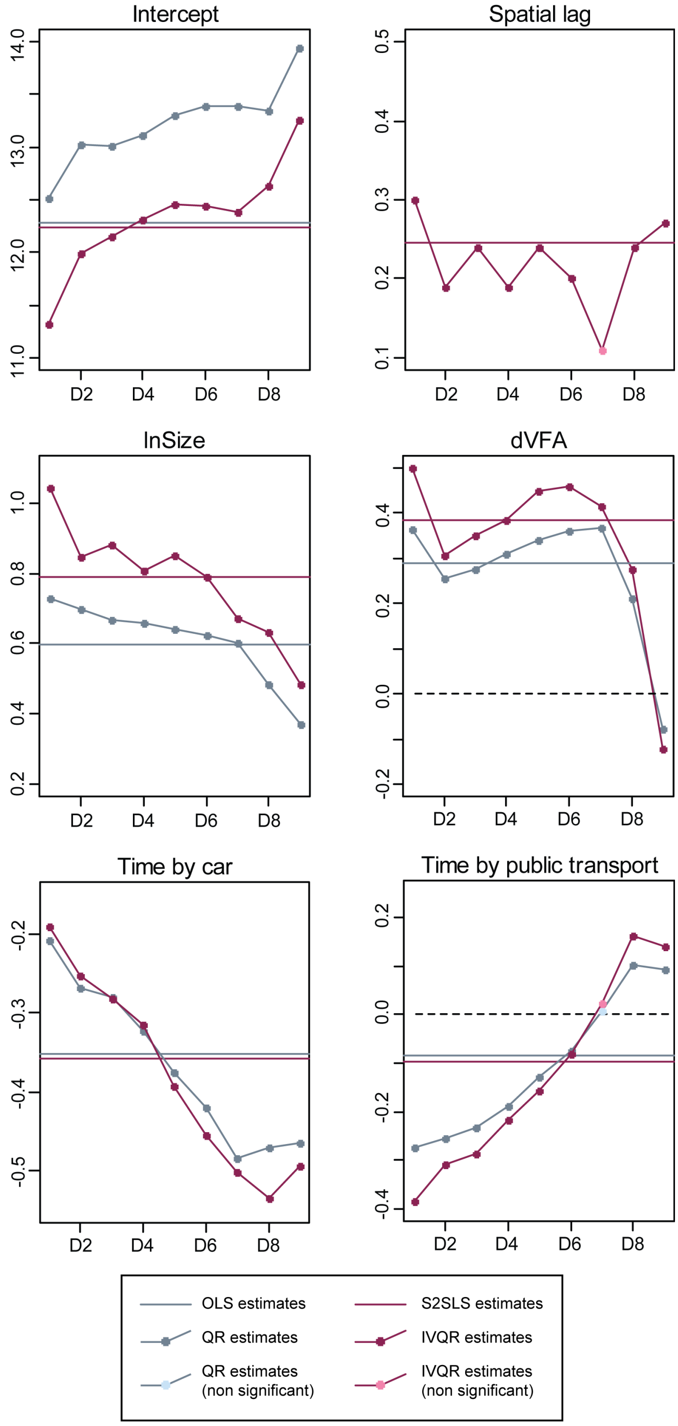


Figure 4.1: Quantile regression estimates (1)

the most expensive land parcels, a slight decrease in the magnitude of the estimate is observed. On the other hand, time-distance to Luxembourg-city by public transport has an expected negative marginal effect for transactions below D6 while the effect turns unexpectedly positive in the highest price ranges (D8 and D9). In other words, the price of a typical transaction in the low price ranges is less negatively influenced by increasing travel-time via road network, while the negative marginal effect of increasing time by public transport is valued strongest by consumers of lower priced transactions, confirming urban economic theory (LeRoy and Sonselie, 1983; Glaeser, 2008). The positive marginal effect of public transport in the highest price ranges suggests that more well-off consumers obtain utility from being located in less well connected sections to the public transport network. We can conclude that shorter commuting time to Luxembourg-city by car is of importance throughout all price ranges, but better-off consumers are willing to pay more to remain close to the employment centre.

An increase in the land vacancy rate has a positive and decreasing marginal impact on land prices, insignificant in the highest price ranges (fig. 4.2). Suggesting that more land available for future construction is perceived positively, especially by less well-off consumers. As discussed in the previous chapter, this might be related to the level of information of the consumers related to future developments in the same section with land consumers considering these open-spaces as permanent green space.

Results from the general model are confirmed for the green land-use diversity index (*DIGreen*), showing significant negative effects in the middle price ranges (D2-D6) (fig.4.2). No significant effect had been estimated in the global model and we expected rather a positive effect, we assume that some additional distance effect is captured by this variable.

The expectations on the marginal effects of Shannon land-use diversity indices are mostly confirmed (lower part of fig. 4.2). In the radius of 100 metre (*mAPsh100*) increased diversity is valued negatively while no significant impact was observed for D1 and from D7 onwards. On the other hand, average section Shannon diversity within walking distance (*mAPsh1000*) has a positive effect for most deciles, although insignificant in D9. Throughout the price distribution, the marginal effect decreases from D2 to D8. Hence all else constant, *mAPsh100* has a particular negative effect on transactions around median price, while the positive effect of an increase in *mAPsh1000* is estimated to have a more important marginal effect for less wealthy consumers. These findings are assumed to be mostly related to differences in mobility of these different consumers and the fact that high diverse land-uses are generally found in more urban areas (as discussed in section 2.3.2.1).

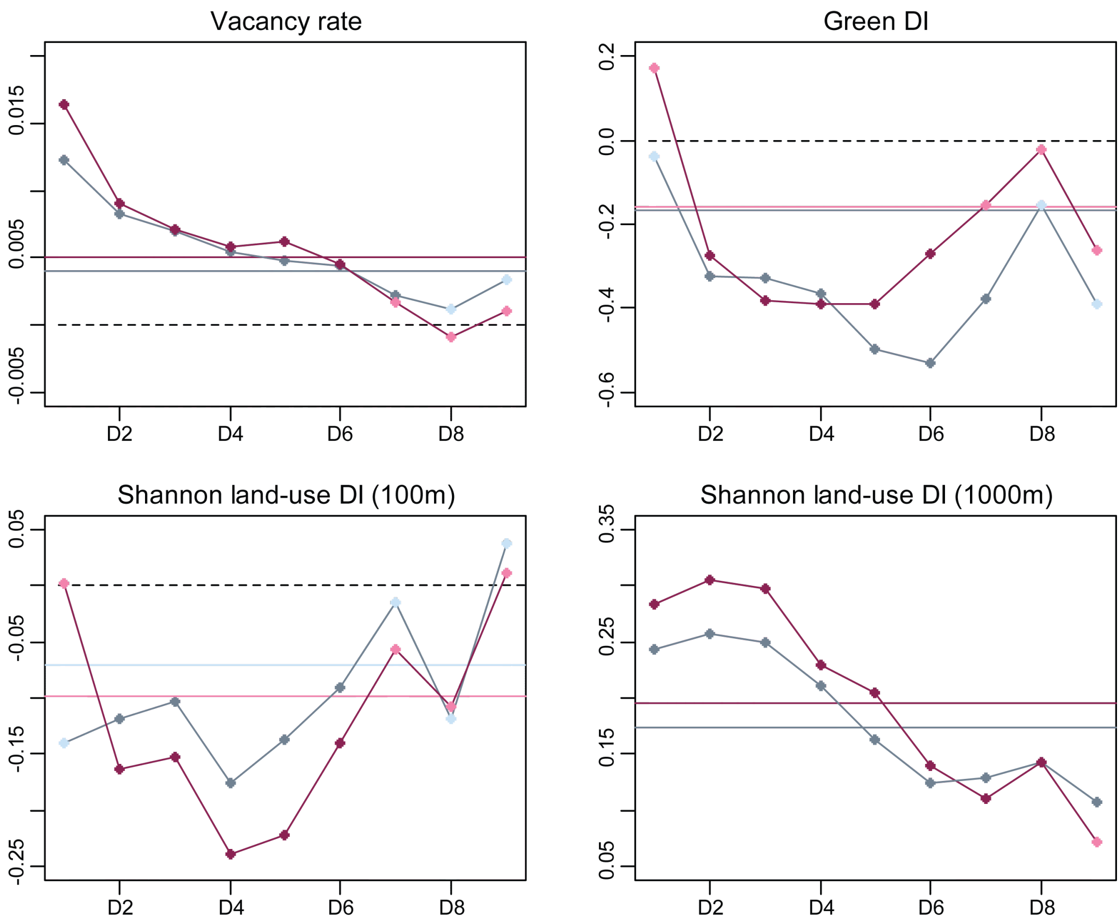


Figure 4.2: Quantile regression estimates (2)

The marginal effect of the availability and diversity of local urban amenities within a section is expected to capture variations in the valuation of retail and service. Results confirm the global model for *DI* with insignificant estimates for almost all deciles (table

F.2). Revealing a negative marginal impact of an increase of local urban amenity diversity only in D6. This is a rather unexpected result, since diversity of services and retail is in general considered a positive amenity. As already suggested in chapter 3, this finding requires a different approach to measure local urban amenities, and might be related to the scale of analysis and the generally high mobility of Luxembourgish residents. The positive but low coefficient of *SSopp* is confirmed for most price ranges ("Retail and service opportunities", fig. 4.3): especially in the middle price ranges. Hence additional retail and service opportunities within a section are valued most positively in the middle price ranges, while for the more wealthy consumers these are of lower importance and mostly insignificant in the lower price ranges.

Although including the socio-economic characteristics was mainly intended to control for neighbourhood effects, some theoretical findings could be confirmed and are worth mentioning here. The positive effect of net population density highlights the premium consumers obtain from being located in sections of above average density, especially in the higher price ranges and in D1 ("Net population density" fig. 4.3). Lowest values are observed for transactions of median price and in the highest price range. All else constant, *NetDens* is assumed to capture an agglomeration effect; more well-off consumers are thus willing to pay more to be located in more urban sections. Further, with increasing prices, and thus consumers' increasing ability to afford high overall prices for land, the positive effect of an increase in the proportion of high income residents (*rHighIncome*) increases. This confirms that more well-off consumers obtain utility from being located close to neighbours of similar socio-economic background, confirming among others the claims of Mieszkowski and Mills (1993).

4.5 Concluding remarks

Spatial quantile estimation results allowed differentiated insights into the valuation of the geographic and structural attributes of developable land prices according to the price ranges. Although quantile regression results are less sensitive to outliers and robust to heteroskedasticity, the ordinary least squares estimation remains inconsistent if

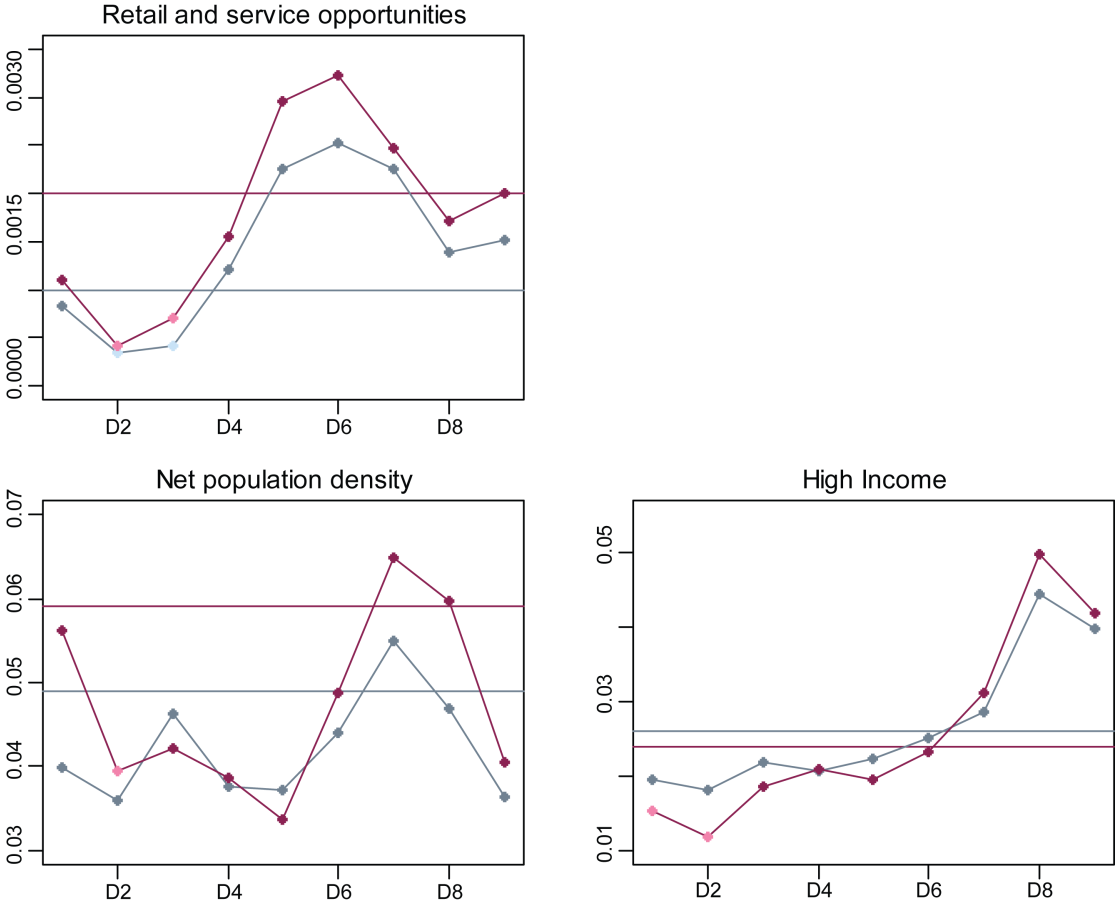


Figure 4.3: Quantile regression estimates (3)

spatial dependence is not accounted for. Therefore a spatial quantile estimation was implemented, which further allowed insights into the distribution of spatial dependence and consumer types. Strongest spatial autocorrelation, in the lowest and highest price ranges, suggests that the prices of these land transactions are more likely to be influenced by the prices of nearby transactions. Based on these findings, we raise the hypothesis of three distinct market segments in the transaction dataset based on consumer types. First, we assume the first decile to capture developable land transactions sold with a non-residential development purpose. Second, transactions located between the second and eighth decile are assumed to most likely represent transactions for residential use, either bought by private households or professional developers aiming at investing in smaller residential projects. Eventually, the last decile is assumed to capture transactions foreseen to accommodate larger development projects of residential or mixed uses. These hypotheses are strengthened if one considers the behaviour of the different attribute coefficients, that will be further exhibited below.

The transactions of very low price (D1) show for most variables extreme estimates, which suggests that these land transactions might not be in competition with the others. This assumption is corroborated by the importance of net population density or the presence of retail and service opportunities, hence the availability of (other) urban amenities within a same section. This points out the attractiveness of an urban context for these transactions, compared to other transactions of lower price. Further, access to Luxembourg-city has the weakest marginal impact on these transactions, suggesting that local accessibility might be of greater importance. The high positive marginal effect of the vacancy rate might be explained by the developers' expectations towards future development of the section. Eventually, the insignificant estimates for green and close by land-use diversity confirm the assumption that local green is not a determinant of the prices of these transactions but that a more urban context is valued by these investors for non-residential land.

A further market segment is assumed for the highest price range (D9), where most green land-use effects are insignificant, and results for accessibility variables and socio-economic neighbourhood characteristics behave differently than in the lower price ranges. We assume that these very expensive transactions, all else constant, rather represent large parcels purchased by professional or public land developers in the perspective of future development, not necessarily exclusively for residential use (e.g.: large scale projects). In this segment, no significant marginal impact is detected for the green land-use amenities, which we assume to be related to the possibility to include local green spaces and urban facilities in these development projects. Distance to Luxembourg-city by car is of major importance, while the estimated positive impact of public transport might be explained by the higher flexibility of the public transport network to be adapted after the project is completed.

Under these hypotheses of market segmentation by consumers, transactions located between D2 and D8 should thus correspond to transactions bought by private households or developers of smaller residential projects. In this perspective, the coefficient estimates provided for these deciles should allow insights into the differentiated valuation of the determinants of residential land price according to consumers' socio-economic background. Additional size being valued highest in the lower ranges, while the marginal effect of existing development projects increases with prices. An explanation might be that land consumers are aiming at building more individual and demanding projects with increasing purchasing power, while less wealthy consumers rather rely on the "carefree" package offered by developers. Estimates for accessibility measures confirm theory, that more well-off consumers depend more on individual transport, travel-time by public transport having no longer a negative impact on prices. An increase in retail and service opportunities is valued positively by consumers in the middle price ranges, while an increased amount of opportunities within the same section is valued less in the higher price ranges, which is probably related to a higher individual mobility of these consumers. The insignificant estimate of DI suggests that diversity of services and retail is either perceived at a different scale, or that the coverage of a diverse supply is satisfactory at national scale and hence not considered in the decision. This urban centrality effect might as well be captured by the net population density. The negative estimates of a sections' green and close by land-use diversity behave in a similar way, being only perceived negatively by the low and middle price ranges, capturing most likely some "Oesling" or rural effect. Increased land-use diversity in walking distance is assumed to capture some local urban effect.

Other approaches to combining the spatial models accounting for spatial heterogeneity and spatial variation of the coefficients through different price ranges have been put forward recently. Geographically weighted regression by quantiles, as implemented by Hallin et al. (2007); Chen et al. (2012) combine QR and GWR to explore spatial non-stationarity through the entire price distribution. A similar approach has been chosen by Des Rosiers et al. (2002); Kestens et al. (2006); Des Rosiers et al. (2007) combining the spatial expansion method to GWR to consider both the non-spatial and the spatial heterogeneity of parameters. With regard to the aggregated location of our transactions the GWR approach would however not be appropriate. The implementation of the mixed effects quantile regression as described by Geraci and Bottai (2013)[14] would be interesting to implement in a next step. However, to date this analysis has not been realised since the three level model, necessary in our case study as will be shown in the next chapter, has not been implemented yet.

[14]Implemented in the '*lqmm*' package Geraci (2014).

Chapter 5

Market segmentation and spatially varying preferences for neighbourhood land-use diversity: Multilevel hedonic analysis

5.1 Introduction

As presented and discussed before, land values mainly depend on the standard trade-off described in the monocentric model, between accessibility to jobs and land consumption (Alonso, 1964; Fujita, 1989). From several extensions to this basic urban model (Brueckner et al., 1999; Glaeser et al., 2001; Turner, 2005; Caruso et al., 2007, 2011) and in particular from the periurban model (Cavailhès et al., 2004), the importance of considering a broad range of location-specific attributes, besides the structural characteristics of a good, is essential to explain land prices. The spatial distribution and diversity of land-uses as well as neighbourhood retail and services around residential places are of particular interest in this thesis, aiming at assessing the attributes' amenity value by relying on the hedonic pricing method.

Today it is commonly accepted that spatial effects have to be accounted for in hedonic modelling. Different autoregressive estimation methods have been developed to account for the different forms of spatial dependence (among others Anselin, 1988b; Ward and Gleditsch, 2008; Elhorst, 2010). These models should allow to identify and correct for the potential bias induced by spatial dependence and have been largely applied in hedonic literature. However the autoregressive functions developed in spatial econometric literature can be seen as "technical fixes" to the problems of modelling spatial data (Orford, 2000), especially as they do not account for problems related to heteroskedasticity and spatial heterogeneity. Spatial heterogeneity is likely to arise if there is variation in the price-attribute relationship by spatial sub-markets (Orford, 2000), so if the prices of land and their characteristics vary substantially with their location in space (Le Gallo, 2004). Consequently OLS estimation, imposing spatial homogeneity, will be misspecified and affect the validity of diagnostic tests (Anselin and Lozano-Gracia, 2009).

The results of the spatial expansion model presented by Geoghegan et al. (1997) suggest that an increase in diversity is valued differently with distance to CBD, and that increased land-use diversity is generally not a desirable feature in the suburban area. The amenity value for quantity and quality of green open spaces has been at the centre of Cho and Roberts' 2008 research, relying on locally weighted regression[1]. They confirm spatial heterogeneity in the valuation of different land-use features with regard to the degree of urbanisation, for instance they find that the positive impact of larger forest patches decreases with distance to the city centre. Cho and Roberts (2008) conclude that this spatial variation in amenity values reveals the need for site-specific land-use policies to fit the local context. With regard to these findings, spatial variation in the valuation of different diversity indices should be further investigated[2]. In this chapter, the focus is hence turned to account and identify spatial heterogeneity in the price-attribute relationship of local and structural attributes and spatial market segmentation in general.

The hedonic pricing method (Rosen, 1974)[3] assumes a unitary market in equilibrium. This assumption prescribes that the implicit prices of the attributes are invariant and it moreover ignores the operational processes and structures that can lead to the disequilibrium in supply and demand (Orford, 2000). Market segmentation might arise when consumers' demand for a particular structural or location-specific characteristic is highly inelastic and that the preference for this characteristic is shared by many other

[1] This method is briefly discussed in appendix D.5, it was not considered in the framework of this research, mainly because of the aggregated location of the transaction data at different levels, although GWR could have been implemented on the aggregated level, the multilevel approach is more appropriate as it allows to consider structural variables at individual transaction level.

[2] Neither rely on the spatial expansion nor the locally weighted regression method has been considered here, a comparison of these two methods is presented by Bitter et al. (2007).

[3] Further discussed in chapter 3 and appendix D.

consumers (Goodman and Thibodeau, 1998). Orford (2000) suggests alternatively to the single-level model, the multilevel approach, which is to him an elegant and conceptually more appealing approach[4]. The multi-level model accounts for the hierarchical structure of the spatial units by modelling the variability at each of the levels and allowing individual observations belonging to a particular spatial unit to be more similar than a random sample (Jones, 1991)[5]. The multilevel approach allows to relax the assumption of a unitary housing market and thus tackle estimation bias that might occur due to spatial heterogeneity. Moreover this approach should be able to capture the effects of spatial dependence, as illustrated in figure 2.21.

The multilevel modelling approach has been widely applied among others in the field of education studies (Hill and Rowe, 1996) and health research (Duncan et al., 1998; Chaix et al., 2005; Lebel et al., 2014) and with increasing interest also in other economic fields (Meijer and Rouwendal, 2006; Fontes et al., 2009), to evaluate for instance business strategies and performance at different levels (Fávero, 2011). Goodman and Thibodeau (1998, p.122) introduced the concept of multilevel modelling in the hedonic pricing context, where the price is determined by the interaction of structural characteristics, neighbourhood characteristics and sub-markets. Orford (2000, 2002) was among the first to investigate means to contextualise the hedonic specification to capture the spatial dynamics of local housing markets to model explicitly the processes leading to spatial autocorrelation in house prices. Other hedonic studies mainly relied on this approach to account for aggregated data at different scales (Djurdjevic et al., 2008) or to illustrate and analyse the distribution of spatial heterogeneity and to discuss the decomposition of unexplained spatial heterogeneity (Brunauer, 2013).

An advantage of the multilevel modelling approach is that it allows random slopes for selected variables at different levels. For instance, Orford (2000) allowed random vari-

[4]A comparison of the multilevel and spatial econometric approaches has been presented by Vanoutrive and Parenti (2009). An alternative to account for the different levels would have been to include spatial fixed effects, as undertaken by Cavailhès et al. (2009). However, this method fails to identify and account for within-unit variation and with regard to the large amount of sections and/or municipalities would have led to problems related to the degrees-of-freedom.

[5]Problems related to different scales of data aggregation are generally summarised under the Modifiable Areal Unit Problem (MAUP), further discussed in appendix D.3.

ation of the neighbourhood quality score within communities level to allow a "*complex geography of social class interactions*"(Orford, 2000, p.1665). Uyar and Brown (2007) accept different structural variables to vary according to neighbourhood affluence or an education quality score, accounting for cross-classified hierarchies, considered as well by Fávero (2011). Chasco and Le Gallo (2013) implement a three-level multilevel hedonic model to measure the marginal impact on house prices of objective and subjective measures of air and noise pollution in down-town Madrid. Beside dwelling size, they consider the noise and pollution variables as random effects at different scales. They find that air pollution only varies randomly at neighbourhood scale, while noise pollution is perceived mainly at the lower level, confirming the local nature of noise compared to air pollution (Chasco and Le Gallo, 2013). Further, it is reasonable to expect the relationship between total land price and the marginal effect of structural attributes to be non-stationary in space due to local demand and supply dynamics (Bitter et al., 2007). Property size is generally considered a source of spatial heteroskedasticity. In the multilevel context, these issues can be addressed by allowing random slopes for the size variable (Jones and Bullen, 1994; Orford, 2000; Djurdjevic et al., 2008; Treg, 2010).

The robustness of these results and their ability to account for spatial dependence, will be tested via a cross-regressive multilevel model (CRMM) as suggested by Chasco and Le Gallo (2012). Since doubts were raised among others by Chaix et al. (2005), claiming that the multilevel approach only accounts for some part of spatial dependence.

The objective of this chapter is, first, to identify the variability of prices within and between the different levels. Second, non-stationary marginal effects of among others land-use and green diversity will be considered. In this perspective a three level multilevel hedonic model will be implemented. The aim is to account for both additional contextual effects and the structure of the available data. This is to our knowledge the first attempt to identify variations in the marginal impacts of land-use diversity throughout different hierarchical levels, while accounting for a variety of transaction and location-specific characteristics at the same time.

The remainder of this chapter is organised as follows, the different datasets are detailed in section 5.2. An review of the multilevel approach, applied to our case study, is given in section 5.3. Results will be discussed in section 5.4, before eventually concluding in section 5.5.

5.2 Precisions on data and levels

The data description presented in this section recalls briefly the different dataset presented in the previous chapters and discussed in depth in part I. The different datasets presented here are allocated to three hierarchically nested levels: transactions (i), sections (j) and municipalities (k), which will be subsequently presented in this section, summarised and illustrated in figure 5.1 and table 5.1. All structural variables are grand-mean centred[6].

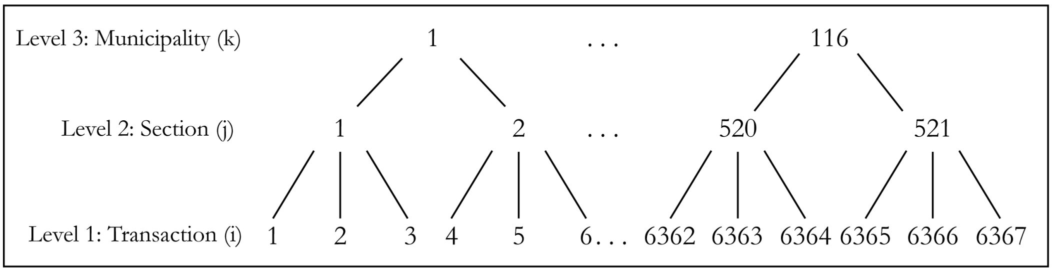


Figure 5.1: Three hierarchically nested levels

5.2.1 Level one: Developable land transactions

The AED dataset for developable land transactions remains unchanged. In this context, the aggregated location of these 6,367 observations (i) located at the sub-municipal level (j), with every section registering at least three transactions. Property size (*lnSize*) is the major structural variable and it should allow insights into the marginal effect of a 1% change to a mean-sized parcel. This effect is expected positive and non-stationary in space. *dVFA* identifies observations sold development plans. In general, these transactions are observed to register higher mean prices per are and to be of smaller size[7].

[6] A discussion of group- and grand-mean centring in the multilevel context can be found in Snijders and Bosker (1999); Shin et al. (2011).

[7] In the models presented of chapters 3 and 4, additional transaction specific variables were considered, however since they did not provide additional information on the agents involved in the transactions,

5.2.2 Level two: Section scale

Accessibility measures to Luxembourg were included in two ways: as travel-time by road network ($tLUX_{ci}$) and by public transport ($tLUX_{pi}$). Travel-time measures has been identified in the previous chapters as a main determinant of land values, negatively impacting on prices. Further travel-time by car has revealed of higher importance than public transport in the consumers preferences. Although in the previous chapters the log transformed travel-time to Luxembourg was considered, in this chapter we followed the hedonic multilevel literature relying on the linear specification of accessibility measures (Treg, 2010; Chasco and Le Gallo, 2013).

The datasets on local urban amenities and public services are based on different official and/or on-line sources, and represent the main retail opportunities and (public) services, based on a wide variety of data sources described in chapter 2.2. The diversity index (DI) was generated following Youssoufi (2011), with the different local urban amenities, weighted according to the frequency of their use, should capture an urban centrality effect. It is expected to have a positive marginal effect on land price. Although this variable was not significant in the previous models, a random slope will be allowed to test if there might be significant spatial variation in this price-attribute relationship.

The Shannon land-use diversity index was generated at two extents: 100m (immediate proximity) and 1,000m (walking distance) around plots registered as available (AP) according to Geoghegan et al. (1997). As already presented, a positive marginal effect of increased land-use diversity price is expected within walking distance ($mAPsh1000$), while in close proximity ($mAPsh100$) the effect is expected to be negative. The marginal effects of land-use diversity are expected to be non-stationary in space. *DIGreen* should further allow to identify the marginal effect of natural land-use diversity on land prices, as for the retail and service diversity, spatial heterogeneity in the valuation of green diversity will be tested for.

they have not been further considered here.

5.2.3 Level three: Municipal scale

Eventually, the socio-economic context variables based on census data from STATEC are added to the model[8]. These explanatory variables are introduced at municipal scale the second most important administrative and planning level at which most spatial planning policies are implemented (Chilla and Schulz, 2011). The part of population above 65 years as well as unemployment rate are expected to have a negative marginal impact on prices and to translate residents' socio-economic weakness. While population variation between 2001 and 2007 and density are expected to have a positive effect on prices.

5.2.4 Spatial weight matrix

For the cross-regressive multilevel model, spatially weighted explanatory variables were generated based on the contiguity matrix, as already presented in chapter 3 and detailed in appendix E.

5.3 Multilevel modelling approach

With regard to potential shortcomings related to the use of spatial weight matrices, discussed in appendix E, in hedonic pricing models and the need to deal with estimation problems related to spatial dependence and heterogeneity the multilevel modelling approach has in the last decade been put forward in the hedonic pricing context[9]. The advantages of the multilevel modelling approach have been summarised by Jones and Bullen (1993) as lying in the presumption of auto-correlation, the correct estimation of the effects for variables aggregated at higher levels, the precision-weighted estimates and the borrowing of strength.

In the single level hedonic model (eq.D.1 p.234) illustrated in graph (A) in figure 5.2. Y_i refers to the natural log of the transaction price; β_0 the overall intercept and β_1 the

[8] The model specification has changed from the previous chapters; in this chapter we rely on the socio-economic context variables provided by STATEC rather than the IGSS, since the multilevel approach should allow to better handling the aggregated scale of the former and since we experienced difficulties in geo-referencing from the postal code to the section scale (appendix C.4).

[9] A more exhaustive presentation of the multilevel approach and technique is provided in Snijders and Bosker (1999); Luke (2004) and with focus on the hedonic model specification by Jones (1991); Jones and Bullen (1993, 1994); Orford (2000).

Variable	Description (Unit)	Mean	Min	Max	Impact
Level-one:	*Transaction*				
lnPrice	ln of price deflated to 2007€	12.46	9.58	14.87	DV
lnSize	ln of parcel size (are)	1.58	−0.46	3.92	+
dVFA	Development project (dummy)	0.19	0.00	1.00	+
Level-two:	*Section*				
tLUX_{ci}	Time to Luxembourg-city by car	28.64	4.52	77.73	−
tLUX_{pi}	by public transport (min)	41.30	8.00	122.00	−
DI	Shopping and service diversity (index)	0.64	0.00	0.91	+
rAP	Vacancy rate (%)	0.11	0.01	0.36	+
DIGreen	Green diversity (index)	0.57	0.08	0.75	+
mAPsh100	Shannon Index in radius of	1.29	0.99	1.68	−
mAPsh1000	100/1,000m around AP (index)	1.63	0.85	2.07	+
Level-three:	*Municipality*				
PopVar	Population variation (2001-2007) (%)	9.39	-2.96	36.30	+
rAbove65	Part people above 65 years (%)	12.67	6.84	20.04	−
rUnemploy	Unemployment rate (%)	4.33	2.07	9.92	−
PopDens	Population density (hab/km^2)	330.58	22.39	2,080.35	+

Table 5.1: Summary statistics : chapter 5

coefficient of the explanatory variable X. In a single-level model the random element is supposed to be captured by the error term (ϵ); the variance is assumed to be zero and the errors independent and normally distributed (Jones, 1991).

The multilevel modelling approach allows to model the structure of the variation not accounted for by the explanatory variables, as it does not assume a constant variance captured by a single error term (Orford, 2000). With the independence assumption relaxed and the intra-group correlation explicitly modelled, more efficient estimates are obtained and thus inference becomes more reliable (Chasco and Le Gallo, 2013). The different steps of the multilevel modelling approach, as considered in this chapter, will be presented in the following sections.

5.3.1 Unconditional model and intraclass correlation coefficients

The standard procedure in the multilevel modelling literature is to start with an unconditional model, without explanatory variables, to determine whether there is variability at the three levels identified in our data, namely the transactions (i), the sections (j)

and the municipalities (k). In case the variances equal zero, no sub-market effects are present and a single-level model would be sufficient to fit the data.

Based on the obtained variances at the three levels, different intraclass correlation coefficients (ICC) can be computed. First, the **level-two ICC** (5.1), as described in Snijders and Bosker (1999), expresses the similarity of mean section prices for sections within a same municipality.

$$\text{Level-two ICC} = \frac{\varphi_0^2}{\tau_0^2 + \varphi_0^2} \tag{5.1}$$

Where, τ_0^2 is the variance of the mean section price between sections within municipalities compared to mean municipal price and φ_0^2 is the variance of the mean price between municipalities compared to the estimated overall mean price. The interpretation would be that if one selects randomly two sections within one municipality and calculates the mean transaction price in one of the two sections, the average transaction price in the other section could be predicted reasonably accurately (Snijders and Bosker, 1999).

Second, the **ICC for level-two (and three) relative to level-one** (eq. 5.2) expresses the likeness of transaction prices in the same section within the same municipality, measuring cluster homogeneity (Jones and Bullen, 1994). Where σ^2 denotes the variance between transactions within sections within municipalities.

$$\text{ICC for level-two relative to level-one} = \frac{\tau_0^2 + \varphi_0^2}{\tau_0^2 + \varphi_0^2 + \sigma^2} \tag{5.2}$$

5.3.2 Random intercept model

In a second step, the random intercept model (RIM) is then estimated, considering the selected explanatory variables. To identify to what extent the intercept varies according to the three nested spatial levels, two macro-models are defined at section (eq.5.3) and at municipal level (eq.5.4):

$$\beta_{0jk} = \delta_{00k} + U_{0jk} \tag{5.3}$$

Where, β_{0jk} is the intercept in section j within municipality k. The level-two intercept is thus as a function of the average municipal price plus a differential U_{0jk} (the section random effect) for each sub-market. In the level-two macro-model, δ_{00k} is the intercept of municipality k, defined by the level three micro-model, where V_{00k} is the random effect and γ_{000} the overall intercept.

$$\delta_{00k} = \gamma_{000} + V_{00k} \tag{5.4}$$

Combining the micro-model (D.1) and the macro-models (5.3 and 5.4) for the higher spatial levels leads to the random intercept model (5.5); with three residual terms that translate the price variability at the three levels. In figure 5.2 (B), the regression lines for the different sections within a two-level version of the RIM is illustrated. In the RIM the only random group effect is the random intercept, considering all explanatory variables as fixed through the three levels. The three level random effects model is denoted as follows:

$$Y_{ijk} = \gamma_{000} + \beta_{1jk}x_{ijk} + (V_{00k} + U_{0jk} + R_{ijk}) \tag{5.5}$$

The dependent variable, Y_{ijk}, refers to the natural log of the price of a transaction i in section j within municipality k, with a single explanatory variable $\beta_{1jk}x_{ijk}$ and γ_{000}, the overall intercept. The error part of the model contains three random terms (V_{00k}, U_{0jk}, R_{ijk}), one for each level describing the differential to the higher level intercept[10]. The level-one error term (R_{ijk}) is assumed to follow a normal distribution, with mean zero and constant variance; the higher level residuals (V_{00k} and U_{0jk}) are assumed to be independent from the lower level residuals. The variance between transactions within sections within municipalities, $var(R_{ijk})$, is denoted by σ^2. And τ_0^2 is the variance of the mean section price between sections within municipalities, $var(U_{0jk})$, compared to mean municipal price and φ_0^2 the variance of the mean price between municipalities, $var(V_{00k})$,

[10]With R_{ijk}, the error term at level one, describing the differential to the average section price; U_{0jk} and V_{00k} being respectively the random term estimated at section or municipal scale, describing the differential to the mean municipal respectively to the overall intercept.

compared to the overall mean transaction price.

The RIM implies that the price-attribute relationship does not vary through the different levels, however this relationship can differ between the spatial levels in more ways (Snijders and Bosker, 1999). The RIM being a restrictive specification if the marginal effect of the explanatory variables vary in space, since the estimated explanatory variables' slopes would not be fixed through the spatial units considered.

5.3.3 Fully random model

In this perspective, the fully random model (FRM) allows the explanatory variables to vary according to a higher-level distribution, by specifying, in the three level context, one or two additional macro-models. Thus, the FRM should control for residual heteroskedasticity at the individual level. The spatial variation of the marginal price of land-use diversity and other local amenities will be controlled as wall as for size. The macro-models for the two higher levels (eq.5.6 for level-two and eq.5.7 for level-three) can be denoted as follows:

$$\beta_{1jk} = \delta_{10k} + U_{1jk}, \tag{5.6}$$

$$\delta_{10k} = \gamma_{100} + V_{10k} \tag{5.7}$$

By including these additional macro-models to the RIM (eq.5.5), the FRM takes the following form:

$$Y_{ijk} = \gamma_{000} + \gamma_{100}x_{ijk} + (V_{10k}x_{ijk} + V_{00k} + U_{1jk}x_{ijk} + U_{0jk} + R_{ijk}) \tag{5.8}$$

The random part of the model is completed by two additional error terms ($V_{10k}x_{ijk}$ and $U_{1jk}x_{ijk}$), allowing conclusions on the marginal effect of the explanatory variable with regard to the intercepts of the section and municipal scale; $\gamma_{100}x_{ijk}$ describing the overall coefficient of the explanatory variables. Their variances are denoted $var(U_{1jk}) = \tau_1^2$ for the random section slope and $var(V_{10k}) = \varphi_1^2$ for the municipal slope. The covariance terms ($cov(U_{0jk}, U_{1jk}) = \tau_{01}$ and $cov(V_{00k}, V_{10k}) = \varphi_{01}$) allow the random intercepts and slopes "*to co-vary according to a higher-level, joint distribution*" (Jones and Bullen, 1994, p.257).

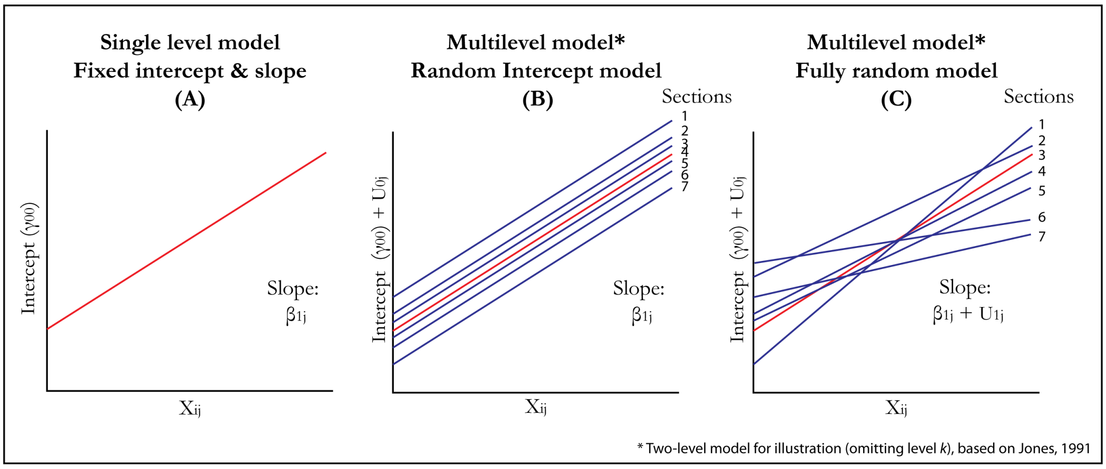


Figure 5.2: Single-level and two level random models

5.3.4 *Cross-Regressive Multilevel Model*

To test the multilevel model's ability to account for spatial autocorrelation, Chasco and Le Gallo (2012) estimated a Cross-Regressive Multilevel Model (CRMM). By including a set of spatially lagged explanatory variables, they tested whether their multilevel model was able to capture all the spatial processes present in the housing prices. This approach relies on the spatial multiplier model as presented in Anselin (2003), adding to the right-hand-side of the equation a set of spatially lagged explanatory variables. This model assumes that spatial effects operate only through the observed explanatory variables (Morenoff, 2003). As the spatial lags $\rho W x_{ijk}$ are at the right-hand-side of the equation, no endogeneity problem arises and thus the model can be estimated via multilevel approach.

$$Y_{ijk} = \gamma_{000} + \gamma_{100} x_{ijk} + \rho W x_{ijk} + (V_{10k} x_{ijk} + V_{00k} + U_{1jk} x_{ijk} + U_{0jk} + R_{ijk}) \quad (5.9)$$

If the multilevel model accounts properly for spatial autocorrelation, the spatially lagged variables should not be significant (Chasco and Le Gallo, 2012).

5.4 Multilevel results

The single level model results (**OLS (1)**) are presented in table 5.3. In general, as shown in previous chapters, the structural and accessibility measures confirm the expectations based on the findings from urban economic theory. As in the previous chapters, an insignificant global effect for retail and service diversity (*DI*) is found. While the general green diversity index (*DIGreen*) has a negative impact of low significance, mean Shannon land-use diversity indices show the expected signs[11]. The average predicted price for a typical transaction of mean size and of type other than *dVFA*, with all variables grand-mean centred (except *dVFA*), is of 283,064€[12].

Lagrange Multiplier test				
	LMerr	RLMerr	LMlag	RLMlag
LM-test	174.64	96.04	78.62	0.01
p-value	0.000	0.000	0.000	0.915

Table 5.2: LM-test results

Spatial dependence tests confirm significant spatial error autocorrelation while no significant spatial lag dependence is detected by the robust LM-test (table 5.2). This spatial error dependence is expected to be accounted for by the three-level model presented in the next section[13].

[11]The functional form of hedonic pricing models has been largely discussed in literature (Ahlfeldt, 2011; Dubé et al., 2011). In the following models the semi-log functional form will be applied following the approach suggested by Verbeek (2008), often applied in multilevel modelling context (Giuliano et al., 2010; Treg, 2010; Shin et al., 2011; Chasco and Le Gallo, 2013).

[12]Applying the method suggested by Verbeek (2008); Multicollinearity was measured by variance inflation factors test (VIF) and test scores are all below 5. The null hypothesis of homoskedastic residuals was rejected by the Breusch-Pagan test, suggesting non-constant variance of the error terms.

[13]The multilevel models, accounting for the three-level nested hierarchical structure of the data, have been estimated by restricted maximum likelihood (REML), according to Snijders and Bosker (1999) the difference between the maximum likelihood (ML) and REML method is that the REML estimates the variance components while taking into account the loss of degrees of freedom resulting from the estimation on the regression parameters, while ML does not. We used the "*lme4*" package (Bates et al., 2013) in R (R Core Team, 2013). The significance levels of the fixed terms have been computed using the "*lmerTest*" package by Kuznetsova et al. (2013), obtaining p-values by implementing "Satterthwaite approximation" for denominators' degrees of freedom.

5.4.1 Random intercept models

The unconditional model (**model (2)**)[14], identifies 7.5% of the total variation in the transaction price is located between sections within municipalities, while 16% of the total price variance is between municipalities. Single- or two-level models are hence rejected. The **level-two ICC** being 68% suggests that mean section prices within a municipality are quite alike, while the **ICC for level-two relative to level-one** indicates that transaction prices are not varying homogeneously between sections of a same municipality. Map **UM** (fig. 5.3) illustrates the variation of the mean municipal price to the overall intercept, as expected estimated municipal mean prices above the overall intercept are observed for most of the municipalities in the southern part of the country, while in the north below average means are estimated in general.

The ability of the fixed explanatory variables, added in **model (3)** (RIM), to account for a significant part of the price variability is confirmed by the LR-test. Only 9.13% of the price variance is now located at the two higher levels, according to **ICC for level-two relative to level-one**. A considerable part of between-transaction variance remains unexplained; which is not surprising with regard to the poor information on the structural transaction characteristics available and the aggregated scale of the contextual variables. Only 1.7% of the total unexplained price variance is located between municipalities. The decrease of the **level-two ICC**, indicates that most of the variation at the higher levels is located between sections. Suggesting that the explanatory variables were able to capture a large amount of the between municipal price variability.

Map **RIM** (fig.5.3) illustrates the variation of the municipal intercept relative to the overall intercept. Mainly transactions located in municipalities within the agglomerations of Luxembourg and regional urban centres (e.g.: Clerveaux and Bettembourg) are observed to register higher variance from the overall mean. The high coefficient for

[14] The transaction level (1) should capture the variation in residential land price due to differences related to the parcel size and type. The section level (2) should capture local variations at village or city district level via location-specific characteristics at section scale. The municipal level (3) should account for the variation in prices at municipal scale, considering the socio-economic composition of the population.

Erpeldange, located between Ettelbruck and Diekirch[15] is also most notable. All else constant, a typical transaction within this municipality is on average 7% more expensive than the overall national mean. Explanatory variables insignificant in the OLS model remain insignificant and *DIGreen* and most socio-economic variables turn insignificant (table 5.3) after considering the different levels. In general, no substantial changes are observed for the other fixed effects coefficients.

- Tests confirm residual spatial dependence and heteroskedasticity
- The utility of a three-level hierarchical model is confirmed
- Large proportion of price variability located between transactions due to omitted structural variables
- Explanatory variables capture an important part of the variance

5.4.2 Fully random model

In a next step, we aimed at accounting for the spatial non-stationary marginal effect of size discussed in literature. This should explain further between transaction variance (**model (4) - FRM Size**). The LR-test indicates that the FRM with random size coefficient between sections performs best and improves the model fit, whereas no significant spatial variation was identified between municipalities in the marginal price consumers are willing to pay for a 1% size increase to an average-sized parcel. Accounting for this random effect explains additional 0.8% of the between transaction variation, while it increases the between municipal variance (φ_0^2) (**FRM Size** fig.5.3). Not considering the spatial variation of marginal size price underestimated the average transaction price for some municipalities, while already low mean prices have in general been lowered in **model (4)**. The gap between the estimated overall mean price and the municipal averages estimated for the municipalities of Luxembourg, Mersch and Erpeldange has increased, similarly to what is observed in the east and north-west of the country. This suggests that the municipal intercept has increased for these municipalities after accounting for within municipal variation of the marginal price of an above average increase in

[15] Together known as the “Nordstad”, the major urban agglomeration in the northern part of the country.

	Dependent variable: *lnprice*				
	OLS	RIM	FRMsize	FRMdiv	CRMM
	(1)	(3)	(4)	(5)	(6)
Level 1:	*Transaction*				
Intercept	12.420*** (0.007)	12.420*** (0.014)	12.420*** (0.015)	12.420*** (0.016)	12.430*** (0.027)
lnSize	0.612*** (0.010)	0.613*** (0.011)	0.609*** (0.019)	0.608*** (0.019)	0.608*** (0.019)
dVFA	0.188*** (0.017)	0.214*** (0.018)	0.227*** (0.017)	0.230*** (0.017)	0.227*** (0.017)
Level 2:	*Section*				
$tLUX_{ci}$	−0.012*** (0.001)	−0.013*** (0.002)	−0.014*** (0.002)	−0.014*** (0.002)	−0.013*** (0.002)
$tLUX_{pi}$	−0.003*** (0.001)	−0.002* (0.001)	−0.003* (0.001)	−0.003* (0.001)	−0.003* (0.001)
DI	-0.030 (0.026)	-0.014 (0.042)	-0.031 (0.039)	-0.029 (0.041)	-0.038 (0.041)
rAP	0.007*** (0.001)	0.005* (0.002)	0.004* (0.002)	0.005* (0.002)	0.005 . (0.002)
mAPsh100	−0.297*** (0.070)	−0.324** (0.113)	−0.373*** (0.108)	−0.375* (0.138)	−0.387** (0.116)
mAPsh1000	0.214*** (0.040)	0.178* (0.070)	0.216** (0.067)	0.187* (0.076)	0.239*** (0.066)
DIGreen	-0.142 . (0.083)	0.005 (0.146)	-0.086 (0.136)	-0.096 (0.140)	0.093 (0.168)
Level 3:	*Municipality*				
PopVar	-0.001 (0.001)	-0.001 (0.002)	-0.002 (0.002)	-0.002 (0.002)	-0.002 (0.003)
rAbove65	-0.006 . (0.003)	-0.007 (0.006)	-0.007 (0.006)	-0.007 (0.006)	-0.003 (0.008)
rUnemploy	−0.018** (0.006)	-0.009 (0.012)	-0.019 . (0.011)	-0.020 (0.012)	-0.008 (0.015)
PopDens	0.000*** (0.000)	0.000** (0.000)	0.000** (0.000)	0.000** (0.000)	0.000* (0.000)
Spatial multipliers					
dVFAW					-0.084 (0.119)
DIW					-0.031 (0.081)
rAPW					-0.002 (0.004)
DIGreenW					-0.419 (0.276)
PopVarW					-0.001 (0.004)
rAbove65W					-0.007 (0.012)
rUnemployW					-0.015 (0.018)
AIC	9,674	9,580	9,295	9,323	9,346
Log Likelihood	-4,822	-4,773	-4,629	-4,638	-4,642
Note:	Signif. codes: . p≺0.1; *p≺0.05;**p≺0.01; ***p≺0.000; (Standard Errors)				

Table 5.3: Fixed effects

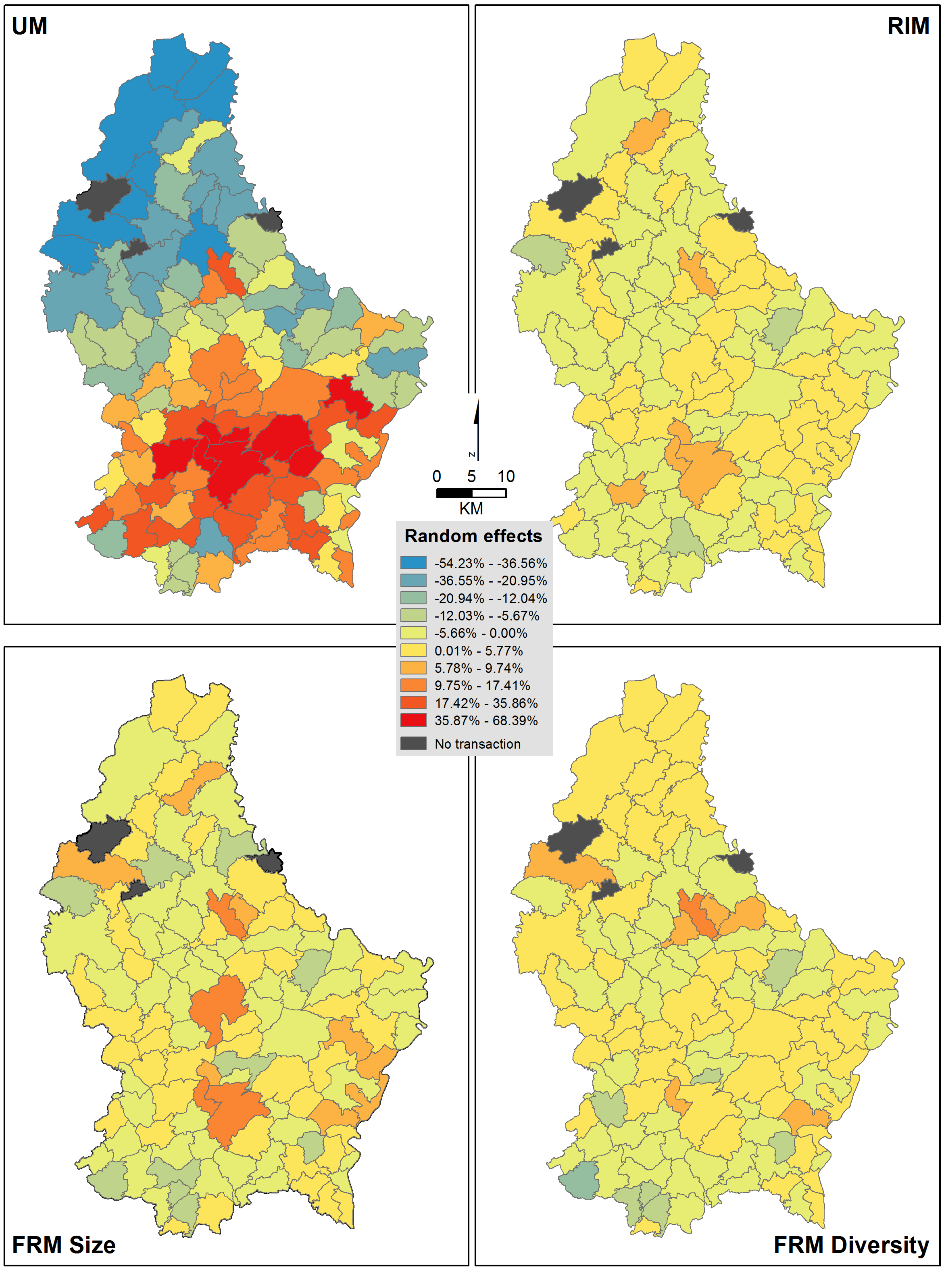


Figure 5.3: Random intercepts at municipal level

	UM (2)	RIM (3)	FRMsize (4)	FRMdiv (5)
	Variance	Variance	Variance	Variance
Level 3: Municipalities φ_0^2	0.079	0.005	0.007	0.006
mAPsh100 φ_1^2				0.501
$cov(\varphi_0^2,\ \varphi_1^2)\ \varphi_{01}$				*0.044*
mAPsh1000 φ_2^2				0.077
$cov(\varphi_0^2,\ \varphi_2^2)\ \varphi_{02}$				*0.020*
$cov(\varphi_1^2,\varphi_2^2)\ \varphi_{12}$				*0.132*
Level 2: Sections within muni. τ_0^2	0.037	0.020	0.024	0.024
lnSizeM τ_1^2			0.064	0.062
$cov(\tau_0^2,\tau_1^2)\ \tau_{01}$			*-0.030*	*-0.030*
Level 1: Transactions σ^2	0.377	0.246	0.226	0.225
Level-two ICC	0.682	0.190	0.216	0.197
ICC for level two relative to level one	0.235	0.091	0.120	0.118

Table 5.4: Random effects

parcel size.

The **Level-two ICC** confirms that the section level contributes still more to the price variability than the municipal level, hence some local specificities impacting on the estimated average section prices have not been captured yet. The covariance term (τ_{01}) translates a negative linear relationship between the estimated section intercept and the marginal effect of size (fig.5.4), which, except for one section, is always positive and negatively correlated to the intercept. With all variables centred, a negative correlation between slope and intercept means that sections with higher marginal mean price for additional size have a lower within-class size effect. Thus, the higher average marginal price tends to be achieved more by prices of the smaller parcels than by higher prices of the bigger plots. In other words, in sections with higher average price for a typical transaction, consumers are willing to pay less for a 1% size increase (a relationship also observed in chapter 4 with regard to the price distribution).

Confirming the findings of Geoghegan et al. (1997), the Shannon land-use diversity measures (*mAPsh100* and *mAPsh1000*) were allowed to vary at municipal scale in **model (5)**. Random variation of *DIGreen* and *DI* were considered as well, but did not lead to a significant improvement of the model and were thus not further considered.

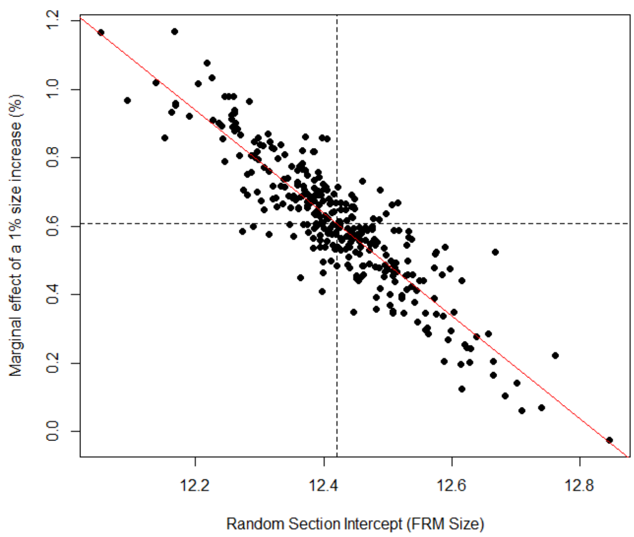


Figure 5.4: Marginal effect of size vs. section intercept

Model (5) accounts for variability on the coefficients of the mean Shannon diversity indices between municipalities. The between municipal variance is further explained by these spatially varying marginal values for average Shannon land-use diversity close and distant to *AP*. In map **FRM Diversity** (fig.5.3), especially the high variances to the overall intercept, observed for Luxembourg or Mersch have been mitigated compared to **model (4)**. Nevertheless, the high variation from the overall intercept observed in the municipalities of the "Nordstad", suggests that the average prices remain higher there. Furthermore, in the former industrial south, the opposite is observed. This suggests that around these urban centres, consumers value a typical transaction differently and these potential sub-markets should be further investigated.

The fixed effects of the land-use diversity variables indicate the expected marginal effects. The global model's results are confirmed for most municipalities and for both variables. However, exceptions are identified, an inversion of the signs of the estimated marginal effects is observed for both extents within some municipalities. For instance, in the "Nordstad" municipalities, land-use diversity in proximity to all available plots is

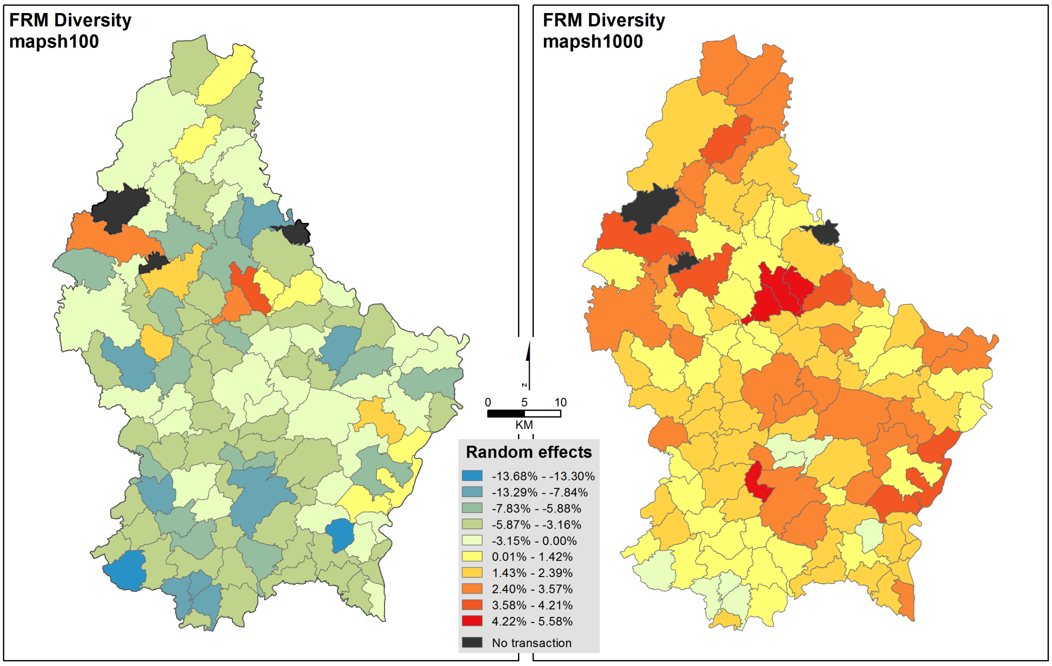


Figure 5.5: Random coefficients: Shannon diversity indices

valued positively, similar estimates are observed in the "Moselle Valley" region and the municipalities around the "Lac de la Haute-Sure". Further, increased land-use diversity in walking distance is found to be valued negatively in some of the municipalities in the former industrial south (i.e. Esch/Alzette, Differdange, Schifflange, Kayl) but as well in some of the periurban municipalities north of the capital city.

The covariance term, φ_{12}, describes a positive relationship between the random coefficients of the Shannon diversity indices; the more a 0.1 increase in diversity in *mAPsh1000* is valued positively, the weaker is the negative effect of *mAPsh100*. Not allowing random coefficients would have led to wrong conclusions, as there are local variations in how diversity in different extents is valued.

- Marginal effect of size varies between sections within municipality
- Sub-markets identified for "Nordstad" and "Industrial South"

- Spatial variation in land-use diversity coefficients was confirmed
- No significant impact of neither service nor green land-use diversity

5.4.3 Cross-regressive multilevel model

In **model (6)** (table 5.3)[16] the results of the CRMM estimation suggested by Chasco and Le Gallo (2012) are displayed.

The LR-test confirms that the model is not significantly improved and as none of the spatial lags are significant and thus neighbourhood sections' values do not significantly impact on transactions.

- Spatial autocorrelation is confirmed to be captured by the final FRM

5.5 Concluding remarks

In this chapter, the multilevel approach was applied it allows to account for the hierarchically nested data structure, to gain further insights into price variability at different levels and to control for spatial dependence and spatial heterogeneity. While the **unconditional model** confirms the usefulness of the three levels, the **RIM** confirmed the important part of price variability explained by the fixed explanatory variables; although a large part of the transaction price remains unexplained, in particular at transaction scale. Allowing random slopes for the structural and the land-use diversity indices provides further insights into the variability of their marginal prices in space. In general, the fixed estimates remain rather stable through the different models and have not been further discussed here.

Parcel size is not valued homogeneously within municipalities, these variations are most possibly related to local urban policies, imposing a hierarchy between the different sections of a municipality. These are considered to be related to different goals of urban development and most likely to the specificities of the supply of developable land. This should be a track for further investigation.

[16] Based on the findings of Morenoff (2003), Chasco and Le Gallo (2012) did not include the spatial lag of all explanatory variables (not considering accessibility and pollution variables) because of issues related to multicollinearity. Similar observations were made in this case study, we could neither consider the spatial lags of the random slope variables, nor those of population density and the accessibility measures.

The spatially varying estimates of the Shannon land-use indices revealed insights into the spatial distribution of the amenity value of land-use diversity. Differences throughout the country could be pointed out in opposition to the fixed effects results. The picturesque and rather diverse landscapes of the "Moselle Valley" and "Lac-Haute-Sure" have a positive marginal effect on the price of a typical transaction, even close-by proximity. On the other hand, estimates of the municipalities characterised by an important industrial past suggest that land-use diversity is rather negatively valued here, at either extent. These results are assumed to be related to a negative perception of the remaining brownfields and the persisting presence of industrial land-uses.

In general, the municipalities register average prices below the overall intercept, all else being constant. Even after accounting for these negative effects of land-use diversity, the below average municipal intercepts suggest a distinct market segment for this region. Moreover, in the "Nordstad" municipalities, in the north of the country, the random intercepts as well as the marginal effects for land-use diversity are positive and largely above the estimated fixed effects. We assume that this might capture some effect related to the political will, in the last decades, to promote urban development within Ettelbruck and Diekirch to form an entity and strengthen its position as the urban agglomeration of the north. The particular high municipal mean for Erpeldange might be explained by its central position between these two municipalities and hence related to speculation behaviour as well as the commercial land-uses allocated along the main road connecting Ettelbruck and Diekirch. A further understanding of these observations would require an in-depth analysis of this area and project and is a track for further investigation.

The green diversity index, expected to capture more directly the effect of "natural" land-use diversity at section scale, did not yield significant results in the FRM. Suggesting that consumers do not account for green diversity at this aggregated scale, which confirms previous findings of this thesis. A similar scale effect is assumed at the origin of the inability to capture the amenity value of local urban amenity diversity, which we assume to be valued at a more aggregated scale. These results require further investigation, ideally relying on accessibility measures at a fine scale or more sophisticated and detailed indices.

The spatial econometric techniques presented in chapter 3 and the multilevel approach should be seen as complementary. In our case study, the advantage of the multilevel approach over traditional spatial econometric techniques lies in the additional information gained on the variability of the transaction prices and the marginal effects of land-use diversity measures between and within the different administrative levels with regard to the dataset restrictions. By simply looking at the *adjusted* R^2 we can conclude the fully random model is able to account for 57% of the variance and the most of this unexplained variance is located at transaction level. Whereas the *adjusted* R^2 in the OLS only explained about 46% of the price variance. Allowing random intercepts and slopes allows thus, on the one hand, to explain a bigger part of transaction prices and, on the other hand, highlights the importance of the need for further structural or local context measures at transaction scale.

Despite the conclusions of Chaix et al. (2005) and Chasco and Le Gallo (2012), the **CRMM** confirms that in our case study this is given. However combining the spatial autoregressive and multilevel techniques in a spatial multilevel model as discussed in Elhorst and Zeilstra (2007) or Corrado and Fingleton (2011) should provide additional insights.

Conclusion to part II

With regard to the research questions raised and the importance of considering spatial effects to understand the location decisions of consumers, the three chapters of this part provided insights into the composition of developable land prices in Luxembourg.

The general question of the valuation and measurement of local amenities has been addressed in chapter 3, where the need to account for spatial dependence, related to omitted spatial variables, was highlighted. Results confirm that land prices are mostly determined by parcel size and access to Luxembourg-city, but despite the aggregated location, the importance of the geographical determinants was underlined. Although it revealed difficult to account for the impact of neighbourhood green and local urban amenities at this aggregated scale, results confirm that consumers account for land-use diversity at different extents differently.

Based on the global model and with regard to the question whether the observed transactions are sold on a single market where all land consumers are in competition, the quantile regression results suggest that we are confronted to three market segments. Further, within the segment representing mostly individual consumers, preferences for structural and geographical determinants were found to vary according to their purchasing power. This finding indicates that local policies should account for the residents socio-economic background and the individual needs of the different consumers to design social and sustainable planning policies.

Eventually, the issues related to spatial market segmentation have been addressed via the multilevel modelling approach, allowing to more fully account for issues related to spatial dependence and spatial heterogeneity. Results highlight that most of the unexplained variance is located at transaction scale, but that additional contextual effects are captured by accounting explicitly for the higher spatial levels. Further, the estimated average municipal prices for a typical transaction, all else constant, where found to vary to the global intercept for some municipalities, suggesting a that particular attention should be drawn to the "Nordstad" and former industrial south. Furthermore, the valuation of additional parcel size is found to vary within municipalities and land-use diversity is not valued homogeneously throughout the country, which we relate to local specificities in local policies and the geographical context.

The results of the three empirical chapters confirm that periurban features, and especially land-use diversity matter in the individual consumers' decision to purchase land, especially by individual consumers.

General Conclusion

In the context of sustained demographic and economic growth, the Grand Duchy of Luxembourg has experienced some substantial changes in urban structure. A large periurban belt has emerged around the agglomeration of Luxembourg-city, characterised by low-density settlements and tight functional links with the capital. Negative consequences associated with periurbanisation are mainly progressing land consumption, intense traffic congestion and pollution as well as at the same time exploding real estate prices. Luxembourgish spatial planning policies mainly aim at promoting a denser and more compact urban development, encouraging the growth of regional urban centres, while restraining land conversion in the more rural areas. The lack of success of these measures, aiming at changing consumers' preferences has given rise to the questions underlying this thesis. To contribute to a further understanding of the emergence of this process, this thesis aimed at analysing the consumers' preferences for the geographical determinants of land prices at local scale.

Based on urban economic theory, the rise of the periurban area can be seen as the result of individual consumers' decisions to locate close to local urban and rural amenities while at the same time being in the proximity of the urban core area (Cavailhès et al., 2004). While in general the negative consequences of urban spatial expansion are highlighted, the periurban concept emphasises the benefits obtained at the individual's level. Besides the standard trade-off recognised in the monocentric city model (Alonso, 1964; Fujita, 1989), Cavailhès et al. (2004) explain the emergence of the periurban area by the utility consumers obtain from local urban and green amenities generated from the co-existence of urban and rural land-uses.

In part I, the objective was the quantification of the structural and location-specific attributes as identified in the literature. The focus was turned to the monocentric-city trade-off (i.e. parcel size and job accessibility) and the periurban effects (i.e. local urban and green amenities). In this perspective, a large database was generated and the expectations, on how these explanatory variables impact on land prices, exposed. The generation and management of the different datasets was related to some challenges and limitations, mainly due to the absence of fine-scale data, recalled below.

These explanatory variables were then considered in the hedonic models, presented throughout part II, providing insights into the extent to which land prices are determined by the structural and in particular the local periurban amenities. The main aim of chapter 3 was the specification of a global hedonic model to shed light on the explanatory variables valued in the consumers' decision, applying advanced spatial econometric techniques. Moreover, theory suggests that consumers' preferences depend on their socio-economic background and that the attributes composing the land parcel are valued differently according to the households' income and composition. Due to some data limitations, market segmentation by consumers was expected since no distinction could be made between the type of consumers (i.e. private or professional) or with regard to the future use of the parcel (i.e. residential or non-residential). The implementation and findings of the spatial quantile hedonic model, implemented in this perspective, have been presented in chapter 4. The aim was to approximate the second step of the hedonic pricing method to identify market segmentation by consumers and varying preferences according to the price distribution within these different segments. Eventually in chapter 5, the multilevel approach was used to identify spatial market segmentation, spatial variation in the implicit attribute prices and to account for the different levels of data aggregation. This approach allowed to account for additional contextual effects and for spatial dependence and heterogeneity at the same time.

Before presenting the general findings of part II and giving some perspectives for further investigation, the limitations encountered in the framework of this thesis will be briefly recalled.

Weaknesses and limitations

Initially, it was foreseen to consider fine-scale time-distance measures to the local periurban amenities, to specifically identify consumers' preferences and demand for proximity to these geographical amenities. Due to data privacy concerns, expressed by the Administration of Deeds (AED) and the Data Privacy Commission (CNPD), data constraints were faced that requested a reformulation of the initial research objectives to a more aggregated scale. The unavailability of fine-scale and detailed transaction as well as consumer information is a common problem for real estate research, in general, and in Luxembourg in particular. This limitation required to fall back on more aggregated measures of these geographical determinants (mainly via diversity indices at section scale), that have been discussed mainly in part I and considered throughout the hedonic models in part II.

A further drawback related to the data privacy constraints is the unavailability of information regarding the different actors involved in the real estate transaction. The second step of the hedonic pricing method (Rosen, 1974; Brown and Rosen, 1982) would have required more detailed information on the socio-economic background of the consumers and sellers. This weakness was partly addressed by falling back on the spatial quantile regression technique, providing insights into the variation of the determinants estimates by price ranges (chapter 4). Eventually, we were confronted to three levels of analysis (transaction, section, municipality) which has been in particular discussed and addressed in chapter 5.

The major limitations related to the land property dataset and the resulting issues relative to the explanatory variable generation process, required an intense time investment. Besides the aggregated scale of the transactions added complexity to the selection of the spatial weight matrix. Accounting for more local spatial relationships among transactions at micro-scale would have allowed further insights into potential spatial lag dependence. Further, the identification of appropriate instruments to account for different endogeneity issues (discussed in appendix D.2.2) could not be accurately addressed.

The lack of detailed and micro-scale transaction data has as well recently been deplored by the Conseil Economique et Social (2013), requesting the CNPD to reconsider their decisions with regard to the aggregated real estate transaction data, with limited success. Further they complain the missing guidelines for the contents that have to be specified in notary deeds, at the origin of the very heterogeneous administrative database, demanding substantial transformations to be suitable for statistical analysis, a shortcoming as well highlighted in Observatoire de l'Habitat (2013).

Main findings and perspectives

"*The purpose of models is not to fit the data but to sharpen the questions.*"

Samuel Karlin, 11th R.A. Fisher Memorial Lecture, Royal Society of London on 20 April 1983

Based on the data described in part I, step by step the significant determinants of land prices were identified in chapter 3. Although in the global model only around 48% of transaction prices were explained, results generally confirmed the expectations raised in part I. However, the fully random model was able to account for 57% according to the adjusted R^2. The multilevel model indicates that additional variance is to be explained at the higher levels but that most of the remaining price variability is located between transactions: additional structural variables are thus required. To better grasp consumers' preferences, more detailed and precise registration of transactions by the notaries should largely improve the quality of estimation results, by adding further structural information and allowing insights into the planned development projects.

The spatial quantile hedonic model suggests a segmentation of the market by three different kinds of land consumers. Assumptions were made that these can be divided into those purchasing land for non-residential use (lowest price range), those purchasing land for residential construction of small (for private use) or medium building projects (professionals) and eventually the highest price ranges are assumed to regroup those transactions bought by professional developers with the scope to invest in larger scale residential projects. To gain further insights into the different valuation practices within these different segments relying on qualitative methods would be appropriate. Interviews with the different actors on the residential land market could be a mean to investigate what the different agents expectations and preferences are while buying/selling land. Furthermore, the notaries and the different administrations involved in the land transactions might as well provide some additional insights on the functioning and the role of the different actors of the real estate market, in general, and the developable land market in particular. This would be particularly interesting in the Luxembourgish case, with regard to the uncommon relationship between policy makers at national and local scale, who are often involved at the two scales (e.g.: mayors being members of the parliament) as already addressed by Hesse (2014b). The different scales of governance and how policies are influenced formally and informally, have been further analysed by Affolderbach and Carr (2014). Their observations reveal a blending of scales on the land-use processes in Luxembourg with weakly differentiated limits between the different authorities and responsibilities (Affolderbach and Carr, 2014). Further, qualitative approaches might help to disentangle the role of the social context and network in the purchasing decision. These are all effects that could not be considered in the framework of this research but certainly would add to the understanding of individuals' decisions. Especially since such an analysis might shed light on the willingness to sell of landowners, be they private or professional, which also largely impacts on urban form and spatial expansion of the urban area.

Limiting the land transaction dataset from developable land to residential building land and additional information on the sellers and buyers would allow much more concise estimation results with regard to individuals' preferences. Further this information would allow to identify appropriate instruments to tackle endogeneity issues. Eventually, such detailed information would allow to confirm the quantile regression results via the spatial expansion method, among others considered by Kestens et al. (2006); Des Rosiers et al. (2007) or ideally the estimation of the consumers' implicit demand for the specific local amenities with regard to consumers' income and social context, hence Rosen's second step.

Parcel size and time-distance to Luxembourg have been identified as the main determinants of land prices. Results confirm urban economic theory claiming that prices decrease with distance to the employment centre and that an increase in size is valued positively. This is particularly true for transactions in the market segments by consumers assumed to consider other than private land buyers. Allowing random variation of the price-size relationship shows variations within the sections of a municipality, suggesting some local policy or hierarchy effect. This local effect between sections within a municipality should be further investigated. Although the quantification of municipal policies is not straight forward, it should allow additional insights into the supply related impacts on residential land prices in Luxembourg. Further, the hedonic modelling context could provide a means to identify how these policies are perceived by consumers in their purchase decision and hence how they impact on land prices.

The negative impact of increased time-distance to Luxembourg-city is strongest for individual transport. However, results show that land consumers account for access to Luxembourg by public transport and are willing to pay more when located in sections with shorter commuting time. Especially, less well-off consumers benefit from a good public transport connection, and consumers of more expensive parcels receive higher utility from shorter commutes by individual transport, confirming Glaeser (2008). This underlines the importance of further promoting public transport, in a context of wide availability of individual transport as observed in Luxembourg. Meanwhile, real estate promoters of large scale projects are found to value mostly access to Luxembourg by individual transport.

Results suggest that retail and service diversity is not significantly impacting on land prices, whatever price range considered, and no significant spatial variation could be identified. Nevertheless, an increase in the number of opportunities has the expected positive effect, especially in the mid price ranges. These findings suggest that, although consumers value more opportunities within a section, diversity of the latter is not a significant determinant of their choice. Assumptions were made of the relation of these results to the general high individual mobility of residents or the good coverage and accessibility of shopping centres. To better access the consumers' preferences for local urban amenities, more differentiated indices (e.g.: by categories like public, education, retail) or a typology of the supply of retail and services (e.g.: relying on factor analysis), could provide insights into which local amenities are valued in Luxembourg and handle the multicollinearity issues faced here.

The quantification of local land-use and green amenities was approximated via diversity indices at aggregated scale. The attempts to approximate fine scale neighbourhood green measures, by considering all available plots' neighbourhood land-use, did not provide significant results. However, throughout the models, the expectations towards the land-use diversity measures, close-by and in walking distance, were met. The negative impact of increased diversity in 100m was particular strong for middle priced properties, while the positive effect of land-use diversity in walking distance is especially true for consumers of less expensive parcels. The observation of the lowered impact of land-use diversity in the 1,000m radius is assumed to be related to higher mobility of the consumers that can afford more expensive parcels. With regard to the price ranges associated with either non-residential developable land or land bought by professional developers in an investment perspective, local periurban effects have no significant marginal effect on land prices. In chapter 5, spatial differences in the valuation of the Shannon land-use diversity indices were identified and the importance of considering the specific local context was highlighted. For instance, in the picturesque landscapes at the Moselle shore, land-use diversity is generally valued positively, while in general negative impacts are estimated in the former industrial south. These differences are likely related to the geographical specificities of these regions and need to be accounted for in planning policies.

Section green diversity did not provide very significant estimation results through the empirical models. However, in the market segment expected to include individual land consumers, a significant negative effect was identified. It is assumed that this variable captures some additional distance effect, or "Oesling" effect and therefore yields the unexpected negative result. These results underline that green amenities are mostly perceived at a local scale and should be further addressed at this level.

The marginal effect estimated for increased vacancy rate, as proxy for general supply, suggests that prices are higher within sections with more land available for construction. On the one hand, this positive perception of vacant land might be related to the consumers' unawareness of the future development of these parcels, considering them as green open-spaces (Geoghegan, 2002; Geoghegan et al., 2003; Irwin and Bockstael, 2001). On the other hand, consumers might be anticipating a future increase in their properties' value, related to future construction projects, as observed by Ooi and Le (2013). Either way, this finding highlights the caution that should be taken by planning policies aiming at regulating land prices by increasing the supply of developable land.

In general, the periurban local effects were insignificant in the market segments associated with non-residential land-uses or larger development projects. In the first segment, this is not surprising; for the second segment we assume that the provision of local urban and green amenities, as well as the connection to the public transport network, will probably be part of the future development project and are thus not major determinants of the price of these transactions.

Throughout the different models, the socio-economic composition of a neighbourhood was controlled for. Land prices are found to be positively influenced by an increased part of high income residents within a section. This positive effect is even stronger for more wealthy land consumers, confirming Mieszkowski and Mills (1993) who highlighted the attractiveness of richer neighbourhoods fostered by the wish to locate close to similar neighbours in terms of socio-economic background. Increased population density has a positive effect on prices and is assumed to translate some agglomeration effect, beside the retail and service opportunities considered. It should be considered as a proxy for urban amenities, as identified as well by Geoghegan et al. (2003); Goffette-Nagot et al. (2011) and it was found to be of particular interest for consumers of higher priced parcels. . The IGSS dataset should be further exploited to better approximate socio-economic changes in sections over time, via a longitudinal analysis on households' mobility for instance. Additional neighbourhood indices, based on the range of socio-economic and demographic neighbourhood characteristics available, could be considered to further approximate neighbourhood quality, similar to other hedonic models (Orford, 2000, 2002; Kestens et al., 2004; Uyar and Brown, 2007; Leishman, 2009).

Further, the multilevel model allowed to identify market segmentation at municipal scale, pointing out especially that transactions sold within the "Nordstad" municipalities are on average more expensive than the typical transaction, all else constant. Moreover, since the “Nordstad” was not identified as major urban agglomeration in the typology presented in appendix B, further investigation is necessary with regard to a more polycentric organisation of the country. We assume that the past policies promoting the “Nordstad” as the regional capital of the north, have supported additional speculative behaviour in these municipalities compared to the rest of the country.

Spatial error dependence was identified to be a major issue in the global model and was accounted for via the spatial error model. Although the quantile regression context did not allow to specifically grasp this form of spatial dependence, the multilevel approach is found to account for spatial heterogeneity and spatial autocorrelation. Although the global model did not provide evidence for spatial lag dependence, in the quantile regression context it was shown that within the different price ranges, prices are positively correlated to prices of nearby transactions. Test results on spillover effects between transaction prices, within the same section and the neighbouring ones, did thus not provide very conclusive outcomes, which we tend to explain by the aggregated neighbourhood scale considered in the contiguity matrix.

The results of this thesis underline the importance of periurban amenities in consumers' location decisions. Despite some limitations and the aggregated scale of transactions, we find that land consumers in Luxembourg have preferences for these amenities and value them differently, between and within the different (spatial) market segments.

If future spatial planning measures find ways to account for these differences and allow the policies a certain flexibility with regard to local geographical specificities and individuals' needs, this might allow for tackling the problems related to urban expansion at the global level and allow at the same time acceptable and liveable neighbourhoods in line with individuals' preferences for periurban amenities.

Appendices

Appendix A

Additional conceptual and theoretical aspects

Urban growth is commonly associated with demographic and/or economic growth as well as the generalisation of the auto-mobile since the middle of the last century. These fundamental drivers of urban spatial expansion have been formalised in urban economic literature since the 1960s, mainly in the US context. In the European context, although the same fundamental drivers apply, different extents and forms of urban spatial expansion have been observed. Differences between European and U.S. cities have been highlighted by Brueckner et al. (1999), while Caruso (2002); Siedentop and Fina (2012) focus on inner European differences.

The conceptual framework, is presented in section A.1, where distinction is mainly made between the different terms and concepts, pointing out to the definition of *periurbanisation*. The major assumption in urban economics is that location results from a choice made by consumers and that this choice is not entirely irrational, nor random in space (Glaeser, 2008). This theoretical background and how individual choices lead to urban spatial expansion and eventually the periurban belt is presented in section A.2. The causes and consequences of urban spatial expansion, that entail challenges for urban and spatial planning at a global and the individual level, will be discussed in section A.3. The debate on the consequences of compaction policies and their effects on urban form and at the individual level are discussed in section A.4.

A.1 Conceptual framework

Urban growth and the resulting spatial expansion of cities has been largely observed in recent decades, in particular by geographers and urban economists, proposing different, yet similar, concepts. Generally, the term *urban sprawl* is used to refer to the processes of urban spatial expansion as well as the emerging patterns and associated challenges for urban planning. Malpezzi (1999) points out to the lack of consensus in discussions on urban sprawl[1], that either fail to define it or present alternative discussions. Malpezzi and Guo (2001) illustrate how the term *urban sprawl* is nowadays commonly used in a wide range of academic and policy oriented discussions. Several different terms and similar concepts have been put forward to describe the causes, process and consequences of urban spatial expansion. Malpezzi and Guo (2001) highlight that, although urban economists used to prefer less "value-laden terms" such as urban decentralisation, sub-urbanisation[2], counter-urbanisation[3], urban sprawl prevails in recent urban economic literature.

The missing consensus might be explained by the fact that the urban sprawl literature confuses its causes, consequences and conditions (Galster et al., 2001). Urban sprawl is a major topic in current urban policy discussions, in particular *New Urbanism*, however it is often not defined or simply positioned in contrast to "*abstractions such as 'orderly', 'well planned' or 'compact development'*" (Gordon and Richardson, 2007, p.13).

Galster et al. (2001) identify six categories of urban sprawl definitions, where urban sprawl is either defined as an example, an aesthetic judgement, a cause of an externality, a consequence or effect, one or more existing patterns of development or as a process of development. In a more general manner, urban sprawl can be defined as "*the less compact outgrowth of a core urban area exceeding the population growth rate and having a refusal character or impact on sustainability of environment and human*" (Bhatta, 2010, p.9), resulting in different land-use patterns and urban forms (e.g.: leapfrog development, scattered and low-density residential land-use,...). Brueckner (2011, p.69) defines sprawl as "*spatial growth of cities that is excessive relative to what is socially desirable*". A

[1] The definitions of urban sprawl mainly applies to U.S. American cities. An overview on the evolution of the urban form in the U.S. since the 19th century can be found among others in Anas et al. (1998).

[2] *Suburbanisation* is generally used as a synonym to urban sprawl (e.g.: Mieszkowski and Mills, 1993; Vaughan et al., 2009), we refer toVaughan et al. (2009) for an overview on the complexity of the suburban and the often value laden perception of these areas.

[3] *Counter-urbanisation* describes the emergence of new urban areas from the perspective of population migration from large to small settlements leading to a more balanced urban pattern (Hosszú, 2009).

more positive perception of urban sprawl is presented by Glaeser and Kahn (2004), summarizing sprawl by two basic concepts: decentralisation and density of individuals as well as economic activities.

There have been several attempts to measure urban sprawl, even though some experts are said to believe that they "*know sprawl when they see it*" (Ewing, 1994, p.1). Ewing (1994) provides a detailed literature review of urban sprawl, defining it via the land-use patterns often observed in relation to population growth. Going beyond the density/distance gradient often referred to in urban economic literature (Mieszkowski and Mills, 1993; Anas et al., 1998). Different theoretical and empirical approaches for measuring sprawl that rely on geographical information systems (GIS) and remote sensing data have been proposed in the US context (e.g.: Galster et al., 2001; Malpezzi and Guo, 2001; Sudhira et al., 2004; Song and Knaap, 2004; Burchfield et al., 2006) and with focus of the patterns of urban growth in Europe (Siedentop and Fina, 2012)[4].

In the 1990s the concept of *Zwischenstadt*, developed by Sieverts (2001) mostly with regard to the polycentric organisation in north-western Germany, has known considerable popularity. Literally translated to *between city*, Sieverts (2006) refers to it as areas that are *neither city nor countryside but which are unlike what we think of as suburbs* (Sieverts, 2006). At first glance it is a diffuse, disorganised structure of different urban spaces with fragmented land-use patterns, a structure without a definite core, but with many more or less functionally specialised areas, networks and nodes (Sieverts, 2001, p.15). Although its patterns appear quite haphazard, the emergence of this *Zwischenstadt* is the result of innumerable rational individual location decisions since the 1950s according to Sieverts (2001), who defines this area as developing independently from

[4] A more detailed review of different studies on measuring sprawl and exploring its causes and outcomes can be found in appendix I of Ewing et al. (2002). Galster et al. (2001) present and illustrate eight conceptually different dimensions of land-use: Residential density, development contiguity, concentration of development, clustering, centrality with regard to the CBD, nuclearity (maximised in the monocentric context), mix of land-uses and income and eventually proximity of different land-uses. They point out the complexity of the processes and urban forms resulting from urban sprawl. Irwin and Bockstael (2007) claim that a shortcoming of these measures of urban sprawl is the lack of considering land-use patterns at fine scale. By investigating the dynamics and spatial distribution of land-use at two points in time, they find a relationship between land-use fragmentation, low-density residential development and urban growth generally located at some distance from the urban areas. Further they highlight the limitations of sprawl measures not accounting for these fine scale changes in land-use pattern.

the urban core and has its origins in individuals' longing for the combination of urban comfort and pastoral romance. This concept emphasises the relationship between urban amenities and individuals' quality-of-life associated to rural land-use patterns. According to Sieverts (2001), the characteristics of the neighbourhood in which a consumer purchases land is getting more important than the distance to employment. Hesse (2012) highlights that the concept of the *Zwischenstadt* requires an independent view on the urban periphery, focussing explicitly on this area, not with regard to its relationship with the core urban area or its consequences.

Similar to the *Zwischenstadt*, the periurban concept emerged in particular in the French framework (Cavailhès et al., 2004; Caruso, 2002, 2005) and the European context (Caruso, 2002; Nilsson et al., 2013). Cavailhès et al. (2004) define the periurban area by two major characteristics. First, periurban areas are under urban influence through a strong functional link with the urban centre that is characterized by commuting flows. And second, periurban areas show a rural morphology due the presence of agro-forestry activities covering and dominating a large part of the area and hence implying low population densities and local employment. More recently, Ravetz et al. (2013, p.13) claim that "*periurban can be seen as not just a fringe in-between city and countryside, a zone of transition, rather it is a new kind of multifunctional territory*". Compared to the urban area where urban land-uses dominate, periurban areas are, morphologically spoken, characterised by a mix of residential and agricultural land-uses. At the same time a strong functional link with the main employment centre persists, translated by commuting flows. Compared to the rural context, the periurban area offers a selection of local urban amenities (e.g.: shopping opportunities, public services), that in the traditional rural economy are less present. To sum up, the periurban area is thus mostly a residential area, characterised by low construction densities where rural activities and land-uses continue to exist and where important commuting flows with the urban centre are maintained.

Given the lack of consensus and the strong negative connotation associated to the term *urban sprawl*, we step back from using this term in this thesis. Contrary to the *urban* *sprawl* literature, distinguishing mainly between urban and rural area, the periurban area arises at the urban fringe and is characterised by the coexistence of urban and rural land-use. It can be seen as a morphologically distinct entity, with regard to the neighbouring urban and rural area. However, emphasis is put on the functional link maintained with the main urban centre, in contrast to the concept of *Zwischenstadt* where this area *between city and countryside* is considered independently.

Before presenting and discussing the causes and consequences of periurbanisation, the theoretical approach developed by urban economic theorists to explain urban growth and periurbanisation is exposed.

A.2 Growing cities theoretically

The common ground of the different concepts and definitions presented in section A.1 is that urban spatial extension is generally triggered by urban growth, governed by some fundamental drivers: population increase, economic growth and/or decreased transportation costs. These fundamental drivers can be derived from the monocentric city model,presented below.

The monocentric city

The monocentric city model developed by Alonso (1964); Muth (1969); Mills (1967), is mainly based on land rent models developed by Ricardo (1821)[5] and Von Thünen (1826)[6]. This basic model is considered as the simplest framework to understand urban growth as it helps to understand the basic determinants of urban decentralisation (Glaeser and Kahn, 2004). In a market economy, land-use allocation occurs through the price of land. Therefore, in order to understand land-use allocation, it is primordial to understand how land prices are determined in a competitive economy (Fujita and Thisse, 2013). The land-use equilibrium is determined by the bid-rent function, defined as the rent different consumers would bid for land at alternative distances to the city-centre to

[5] Ricardo (1821) introduced the notion of scarcity, land rents exist due to a fixed supply of land with particular attributes.

[6] The Von Thünen model states that the existence of a market place is sufficient to structure land-use by different activities in a competitive land market and sought to explain the pattern of agricultural activities around cities (Fujita and Thisse, 2013).

maximise their utility. In the competitive land market, each unity of land is affected to the highest bidding use, the price of land at any location is the highest bid consumers are willing to pay at that location. The outer edge of the urban area, or urban fringe, is located at the distance to the Central Business District (CBD) where the residential bid-rent is below the agricultural bid-rent. The external boundary of land-use is at some distance from the CBD where the opportunity cost of land is assumed to be zero (Fujita and Thisse, 2013).

The spatial extent of the urban area is an increasing function of population and income and a decreasing function of agricultural land price. The urban area only expands if agricultural land is made available for transformation, that would be the case if the profit from agricultural land-use are below those of any other activities (Brueckner and Fansler, 1983).

To identify the determinants of residential equilibrium in the monocentric city, some additional assumptions of the model should be highlighted[7]. The consumer location decision occurs within an urban area (emerging from the land-use equilibrium) organised around a single, dimensionless, CBD in a featureless plain where all jobs are located in the CBD and space is assumed to be homogeneous except for the distance to the CBD (Fujita and Thisse, 2013). Households are assumed to have identical preferences and income. They are confronted with commuting costs depending on the distance they have to commute to work[8] and the same per-kilometre cost of commuting applies to all. Households are assumed to consume only two goods: housing and a composite of all consumed goods and all households are renters. Landowners are assumed to be exogenous and land developers or landowners location choices and profits are not considered[9]. Supply is assumed to be fixed.

[7] We mainly fall back on the annotation and abbreviations chosen by Brueckner (2011).

[8] Commuting costs are divided in two components according to Brueckner (2011): the money cost and the time cost, the latter being ignored in the basic model for simplicity.

[9] For further insights on the implications of profit maximisation behaviour of land developers on the urban structure can be found among other in Glaeser (2008); Brueckner (2011).

The housing good[10] represents the size of a land unit available for construction[11] and the rent is given per unit. The total rent is thus the unit rent times its size, representing the total rent paid with regard to dwelling size. Households' disposable income corresponds to the gross income after reducing transport costs. Consumers maximize their utility under a budget constraint, where the total expenditures (on housing and all other goods) equal the total disposable income. Hence, land consumers obtain utility from consuming land and all other goods.

The most important concept of urban economics is the spatial equilibrium assumption (Glaeser, 2008). The housing market can be seen as the result of the residential location choices made by households, under market equilibrium assumption, the prices paid to purchase a good translate the demand of households for a location with regard to distance to the CBD. Residential equilibrium within the urban area is reached when consumers obtain the same level of utility at all locations as they offer an equivalent package of amenities. This spatially uniform utility is only given if the per unit price decreases with increasing distance to the CBD. To prevent utility from decreasing, and thus maintain equilibrium, "*some offsetting benefit must be present*" (Brueckner, 2011, p.31). The lower unit price compensates for the increased transport costs, induced by moving further from the CBD, hence the per unit price decreases with distance to the city. On the other hand, unit size increases with distance to the CBD, as the consumption of all other goods is decreasing with distance. The increase in unit size in response to decreasing prices with distance represents the *substitution effect* related to the change in unit price (Brueckner, 2011).

In the basic model the urban area emerging around the CBD fits its population, as exogenous landowners allocate land to the highest bidding use. The fundamental drivers of urban growth that can be derived from this model are a growing population, increasing incomes or decreasing transport costs (Brueckner, 2000).

We focus on the residential viewpoint of urban economic theory and how residential

[10] This model allows to analyse the consumers choice of land size, as we consider in this thesis residential land, the housing production is not presented.

[11] Plot size is considered to be a normal good (Fujita and Thisse, 2002).

location choice is determined by the trade-off property consumers make between access to employment and land consumption. The monocentric model is often criticised, but although it is an extreme simplification of reality, it manages quite well to describe urban spatial expansion and to identify the main drivers of urban growth (Brueckner, 2011).

Within this theoretical framework, urban spatial expansion is considered as the outcome of some fundamental changes (at global level) and the result of individual location decisions (local, individual level), where consumers aim to maximise their utility under a budget constraint.

Extensions to the monocentric model

However it is too simplistic to reduce households' decisions to the sole desire for larger properties. First, locations in or around the urban area are more diversified than by their distance to the employment centre and consumers are most likely to consider local specificities of different neighbourhoods in their location decision. Second, the demand for land is confronted to individuals and households with heterogeneous socio-economic backgrounds and preferences. Housing is a basic need of people but quality-of-life is determined by more than property size and short commutes to work. Third, urban spatial expansion is a process, the dynamic nature is not accounted for in this one dimensional context. As pointed out by Mills (1981, p.210) "*the monocentric city model equilibrium is the result of a sequential decision process*". The importance of the temporal aspect of urban expansion at the same time as the resulting patterns, is frequently highlighted (Mills, 1981; Galster et al., 2001; Bhatta, 2010; Caruso and Cavailhes, 2010)[12].

In the 1990s, emerged the field of new economic geography (Fujita, 1989; Krugman,

[12]Some analytical models considering the dynamic framework of urban growth have been presented among others by Wu and Plantinga (2003); Cavailhès et al. (2004); Turner (2005); Caruso et al. (2007), a more detailed review can be found in Caruso and Cavailhes (2010). Caruso and Cavailhes (2010) distinguish between static models considering exogenous economic agents promoting the creation of open space, dynamic models considering the emergence of non-developed spaces within a competitive land market and eventually models considering the interactions among open-space producers and consumers of the resulting green amenities leading to a land-use equilibrium where both types of agents co-exist. They illustrate these different models of land-use allocation by a selection of figures. As an example they illustrate the spatial urban patterns with regard to the optimal location of green-belts. Urban structure in a dynamic context where landowners anticipate future land conversion potentials and speculation is resulting in leapfrog development in a first stage and densified via infill development in a second stage as presented by Mills (1981).

1991), aiming at explaining the formation of urban areas and the spatial structure of these areas as the result of interactions among individual economic agents and their distribution in space. These models are more robust than the traditional bid-rent models as they explain the emergence of agglomerations and their urban spatial structure as an endogenous process resulting from the interactions among individuals and their distribution in space (Irwin and Geoghegan, 2001). Agglomeration economies exist "*whenever an individual productivity rises when he or she is near to other individuals*" (Glaeser, 2008, p.116). Their role in location choice of firms, have been extended to residential choice of households by Turner (2005). Individuals are on the one hand assumed to benefit from being together (e.g.: social interactions, knowledge spillovers) while on the other hand these interactions might as well be avoided as they might have negative impacts on individuals well-being (e.g.: crowding and related vanishing open-spaces) (Caruso and Cavailhes, 2010). Extensions to the basic model have been made, considering different income groups and how they locate within an urban area (Brueckner et al., 1999; Glaeser, 2008; Brueckner, 2011).

It is beyond the scope of this thesis to detail the different extensions of the 'Alonso-Muth-Mills' model put forward in the last decades, considering in example heterogeneous consumers, polycentric organisation or land-use policy implications. For a detailed review of the common extensions to the urban model we refer among others to Fujita (1989); Anas et al. (1998); Glaeser (2008); Brueckner (2011). Here, the focus will be turned to the extensions considering the effects on urban form by the perception of green and local urban amenities as well as local urban amenities, as considered eventually in the periurban model (Cavailhès et al., 2004).

Importance of local urban and green amenities While the first objective of urban economics is to explain the allocation of land-uses and size of urban area as a result of market processes (Anas et al., 1998), a major concern is the importance of space in economics. The limitation of the monocentric model is partly due to its treatment of space, which is reduced to a simple measure of distance from the urban centre (Irwin and Geoghegan, 2001). As highlighted by Nilsson (2014, p.46) "*the spatial distribution of amenities is an important determinant of urban development patterns and plays a major role in shaping the urban spatial structure*" and is thus considered in the demand of residential land consumers. There has been an increasing number of theoretical and empirical analyses of the effect of green amenities on residential location choice. A review of theoretical models from urban economic and geographic literature considering the heterogeneity of urban patterns with regard to green amenities is provided by Bockstael and Irwin (2000) and more recently by Caruso and Cavailhes (2010).

Although generally considered as positive amenities, green land-uses can have negative effects on quality-of-life, "*if it is unsightly, odorous, or insect ridden*" (Geoghegan et al., 2003, p.34). For instance, the role of agriculture is quite mitigated as it generates at the same time positive and negative externalities for residential land consumers (Bockstael and Irwin, 2000).

Distinction is made between private and public open-space or between undeveloped and protected and open-space available for development (Brueckner, 1983; Yang and Fujita, 1983; Wu and Plantinga, 2003; Irwin and Bockstael, 2004; Turner, 2005), finding positive impacts on property prices if open-space is protected while the proximity to open-spaces available for construction generally has a negative or insignificant impact (Irwin and Bockstael, 2001; Geoghegan, 2002; Geoghegan et al., 2003), further discussed in part I.

Local urban amenities. The importance of considering local urban amenities in households purchase and location decision was among others put forward by Tiebout (1956); Brueckner et al. (1999); Glaeser et al. (2001); Cavailhès et al. (2004) or Brueckner and Helsley (2011). Beside access to job, individuals obtain utility by the availability and proximity to urban amenities (e.g.: shopping opportunities, public services). A more detailed literature review of the role of local urban amenities on urban forms is provided in part I.

The monocentric city with periurban belt

In perspective of the importance of green amenities in residential location choice and the role of social interactions, Cavailhès et al. (2004) extended the basic urban model and formalised the rise of a mixed belt of agricultural and residential land-uses, calibrated on the French context. A residential location model where residential land consumers and farmers that are in competition on the same market was developed. They rely on urban economic models with amenities as presented above and allow interactions among agents.

The assumptions rely mainly on those of the basic urban model, a featureless plain where all non-agricultural jobs are concentrated in a dimensionless CBD. They consider interaction among two types of agents, residential land consumers, working in the CBD, and identical farmers are competing on the land market. Both types of agents are free to migrate to maximise their utility or profit.

Three consumption goods enter the consumers utility function: a composite good, land consumption and different local public amenities consumed at the place of residence. Distinction is made between different urban (e.g.: retail opportunities, public services) and rural amenities (e.g.: green open-space), where the latter are considered as a by-product of agricultural land-use, a mix only found in the periurban area. These amenities are considered as the spatial attributes of properties, directly entering the residential utility function but not the budget constraint (Cavailhès et al., 2004).

For the periurban belt to emerge, agricultural activities have to produce a certain number of rural amenities. Periurban land-use equilibrium is reached when the bid-rents are at equality (Cavailhès et al., 2004). A central assumption of the periurban model is that rural features of the periurban area enter the utility function of residential consumer, while at the same time, farmers benefit from proximity to the CBD through lower shipping costs.

Residential location choice in the periurban area is thus determined by the trade-off made by residential land consumers between accessibility and land consumption, as known from the basic model. In addition, residents of the periurban area benefit from spatial externalities related to the land-use mix and the neighbourhood interactions that occur in the periurban area.

The urban area is larger with the periurban belt, and thus natural land consumption and transport costs are higher. However in the periurban model, the substitution between land and local amenities is accounted for in the budget constraint. Cavailhès et al. (2004) conclude that periurban development is less space consuming as it makes lots smaller in the urban area, but that it pushes the outer edge farther from the CBD, thus encouraging the spatial expansion of the residential area. In general, they claim that the periurban configuration is however more efficient than the one suggested by the basic model, as the taste for large plots is partly substituted by the implicit value of the rural amenities.

The periurban model highlights the importance of local urban and green amenities in the emergence of a periurban area, as a result of the interaction between city dwellers and agricultural activities.

A.3 Causes and challenges of periurbanisation

Urban spatial extension and the emergence of the periurban area are generally associated to a wide variety of negative impacts on society and individuals well-being. The negative consequences can be summarised to the fragmentation of undeveloped land, to the degradation of natural resources and to the elimination of the functional open-spaces, to increased public service costs and traffic congestion and the resulting pollution (Irwin and Bockstael, 2007). These negative consequences are mainly affecting the whole society and entailing global challenges for the planning policies. Some studies relativise these negative consequences and point out the positives effects of the periurban areas, especially at the individual level.

Brueckner (2000) claims that policy makers need to be aware that urban expansion is first determined by these fundamentals of urban growth. Distinction should be made between causes that might lead to sprawl but must not (fundamentals and the lack of or inappropriate policy measures) and those that inevitably yield sprawled growth (individual preferences for large properties, geographical context, or budget constraints) (Bhatta, 2010). Brueckner (2000, 2011) identifies three major market failures as causes of excessive urban growth[13]: The failure to account for the social costs related to traffic congestion, for additional needs in public infrastructure and for the social value of vanishing open-spaces. If these costs are not taken into account when purchasing residential land, this disrupts the land-use equilibrium in the competitive market, leading to more expanded urban areas than socially acceptable (Brueckner, 2000). Not considering the real costs of land has impacts on individuals space consumption, as the real costs induced to society (e.g.: by pollution, time spend in traffic jams, vanishing productive farmland...) do not flow in the individuals budget constraint.

The fundamental causes of urban growth identified in the monocentric urban model are related to societal and technological changes occurring at a global level. Ravetz et al. (2013) claim that periurbanisation should not only be seen as a negative change of the rural areas, in particular with regard to the growing part of population moving to these areas. The main challenge for urban planning is to develop realistic plans to deal with it. The positive benefits of population growth, as a fundamental cause of urban growth, bears positive benefits for the city by generating for instance new jobs.

The desire for larger parcels and single-family housing can be seen as the main driver of periurbanisation, due to individuals choosing to move to the urban-fringe where larger parcels are available and affordable. The land-use patterns generally associated with the periurban area are low-density settlements, discontinuous to the urban core area and emerging leapfrogging urban development said to waste and miss-allocate productive farmland (Mills, 1981). Glaeser and Kahn (2004) however highlight that this allows access to cheaper and larger properties, and hence has a positive effect on individuals.

The importance and omnipresence of individual transport, related to the increased amount and reduced duration of commutes to the employment centre, as well as the ownership of multiple cars per household, are generally put forward as a main cause of

[13] A market failure arises when economic agents face incentives that are distorted because of institutional failings or some other reason, leading to economic outcomes that are bad from society's point of view Brueckner (2001).

the spatial extension of the urban area (Glaeser and Kahn, 2004). On the one hand, the improvement and extension of the road infrastructure has allowed faster and more comfortable commuting, but on the other hand these network extensions and the indirect promotion of individual transport via low gas taxes have largely contributed to periurbanisation processes. The increasing and generalised motorisation of the households is associated with an increase in traffic congestion, and negative consequences on the environment (e.g.: air and noise pollution). Glaeser and Kahn (2004) argue that the technological advances considerably reduced emissions of pollutants in the last decades. In general they conclude that the negative externalities of the increased traffic on the environment exist but are not overwhelming.

Another negative effect of the emergence of low density settlements is related to increased social isolation in these areas. In the American context, Putnam (2000) claims that living in the low-density areas at the urban fringe weakens social interactions and leads to social isolation. Based on these findings, Brueckner and Largey (2008) developed a theoretical model on the relation of residential density and social interaction and tested it empirically. Results revealed that density has a rather negative effect on social interactions at local scale. In low density areas interactions are even shown to be stronger than in the dense urban centre[14]. In other words, social isolation is more likely to occur in the dense urban areas.

The attractiveness of the periurban area is thus also determined by the socio-economic specificities of periurban neighbourhoods and different lifestyles associated to different income classes and household compositions. As summarized by Caruso (2005), households' preferences vary with different life-cycle stages and their socio-economic specificities. Families have been shown to prefer to locate outside the core area, as the periurban area offers a wider range of recreational activities and spaces as well as associated to

[14]Possible reasons for less social interactions in the urban area might in example be higher density spurring the need for privacy, more sources of entertainment, or higher criminality increasing suspicion towards neighbours ad thus more reluctance to interact. Higher social interactions in low-density neighbourhoods might be related to the spatial layout of the properties, in favour of outdoor activities possibly leading possibly to unplanned encounters, related as well to different demographic characteristics (e.g.: families with children) of theses areas households (Brueckner and Largey, 2008).

school quality (Uyar and Brown, 2007; Clapp et al., 2008). Further, individuals have been shown to prefer locations with more similar neighbours (Mieszkowski and Mills, 1993; Glaeser, 2008; Brueckner, 2011). On the one hand, this influences the social composition of neighbourhoods and, on the other hand, it impacts on residential land prices. In example high land prices in some areas might, beside the location-specific characteristics, be determined by the benefit consumers obtain from locating in a rich neighbourhood, leading to homogeneous communities of higher-income households. Hence high residential prices may be perceived as guarantee of quality and have a positive impact on neighbouring prices.

These urban developments entail major challenges to spatial planning policy makers, whose general objective should be to develop means to tackle the negative consequences and however respecting the choices and preferences indirectly expressed by individuals. Caruso (2005) argues that the costs related to periurbanisation should be compared to the positive effects of the periurban area on well-being and quality-of-live that are generated at the individual level. Both aspects have to be considered by planners and policy makers to allow sustainable spatial development.

A major urban function lies in the concentration and availability of a wide variety of goods and services both public and private, low construction and population densities have in this perspective been associated with increased public service costs (Mills, 1981; Brueckner, 2001). Brueckner and Helsley (2011) show the relation that exists between sprawl and the decline of the central urban areas, urban blight, that has been observed mainly in the U.S.. They claim that the market failures leading to sprawl could also cause a degradation of the urban structure in the central urban areas, meanwhile, the gentrification of the core urban area is predicted by LeRoy and Sonselie (1983), if one accounts for different income classes and different transport modes.

Brueckner (2000) further points out the political causes of sprawl due to land-use zoning, tax regulation measures or mortgage interest deductions, dividing responsibility between local and national governments. A further discussion on the role of politics and the failure to prevent the negative effects of sprawl is provided by Glaeser and Kahn

(2004).

In planning and economic literature, urban spatial extension is thus not perceived as a sustainable form of development. However, means to measure the degree of sustainability is to be further explored, according to Bhatta (2010). Neuman (2005) considers *sustainability* as a broad and vague term with many meanings. A general accepted meaning would be the balance among equity, economic and environmental concerns[15].

A.4 Spatial planning measures to the "compact city"

With regard to the consequences and challenges of urban spatial expansion, different urban planning policies, to tackle the consequences and to slowdown urban expansion, have been developed.

The *Smart Growth* movement[16] appeared in the 1990s in the U.S. to face urban sprawl and its negative consequences. The Smart Growth concept advocates "*compact, transit oriented, walkable, bicycle-friendly land-use, including neighbourhood schools, compete streets, and mixed-use development with a range of housing choices*" (Bhatta, 2010, p.39-40)[17]. The aim is, via a selection of public and private subsidies, the creation of an encouraging framework to refocus a share of regional growth in the urban core area linked to the ideals of *new-urbanism* and *neo-traditional* town planning[18] (Burchell et al., 2000), to avoid a further spatial expansion of this area (Bhatta, 2010).

The aim is to control the spatial expansion of the city, while in the meantime revitalising the central urban areas to make them more attractive (e.g.: support community and neighbourhood revitalisation efforts, infill of vacant lands and conversion of brownfields). The Smart Growth principals aim at protecting the land and natural resources (e.g.: by more compact urban growth forms) and reorientating transport (e.g.: by promoting public as well as non motorised transport, fostering regional access to goods, services

[15]It is beyond the scope of this thesis to further investigate on the definition of sustainability.

[16]We restrict the discussion to the Smart Growth principals, of which some can be identified in recent spatial planning policies in Luxembourg.

[17]For a detailed review of the *Smart Growth Principals* see among others to Burchell et al. (2000); Bhatta (2010); Smarth Growth Network (2014).

[18]*New Urbanism* promotes similar claims than *Smart Growth* from a more local perspective of urban planning and sustainable development discourses (Hesse, 2014a, p.1-2).

and recreational amenities). Further, they aim at decreasing socio-economic segregation and foster a sense of community (Bhatta, 2010) in the urban area by promoting higher population densities. By generating more compact neighbourhoods, characterised by a mixed-use development, they aim at promoting affordable housing in the city centre and urban parks and recreational areas (Bhatta, 2010). Eventually, they should promote attractive and functional communities in the inner urban area (Burchell et al., 2000). These principals hence aim at promoting more attractive urban areas for individuals and by these means slow down urban spatial expansion.

Critiques to these measures. In the last decade, several critics have raised and pointed out the negative consequences of such measures on quality-of-life and the possibility that they might not guide residential location choice in the expected way. Burchell et al. (2000) raise the question on how to convince individuals of the attractiveness of non single-family housing and reduced spaces for cars (e.g.: inner-suburban and urban areas).

Neuman (2005) questions that striving for more compact cities[19] is the solution to more sustainable urban forms. The paradox of the compact city refers to the negative relation of the sustainability of cities and their liveability, a paradox that can not be solved with the actual "*modes of thinking about, acting on, and living in the city*" (Neuman, 2005, p.16). They claim that liveability is a matter of personal preference and not only a matter of urban form.

Hesse (2014a) emphasises some contradictions that arise when policies intervene to foster more compact and dense urban development, claiming that relying on simple density measures to deal with the rather complex processes of urban growth is subject to limitations, that are difficult to overcome. Others claim that the preservation of urban open-space may yield more sprawled development patterns (Irwin and Bockstael, 2004; Glaeser, 2008). Turner (2005) shows that infill development policies are not improving the well-being of residents, and that the spill-over effects of these developments on the

[19]According to Neuman (2005) there is no commonly accepted definition of the *Compact city*.

quality-of-life of already installed residents should not be neglected. Caruso et al. (2007) confirm that if undeveloped gaps are filled in, neighbouring locations' rent decrease due to the vanishing rural amenities provided by these undeveloped parcels. Ooi and Le (2013) find positive effect of infill development on local housing prices; but claim that in their case study the developers act as price leaders and contribute to price discovery in housing markets.

Brueckner (2000, 2007) points out that the definition of zoning tools, such as urban growth boundaries (UGB)[20] is quite difficult, given that inappropriately restricting spatial growth might result in example in an excessive increase of housing prices, degradation of central urban area or density and thus (Brueckner, 2000) claims that policy makers should resist to impose too strict UGBs. As put by Burchell et al. (2000): Smart Growth is missing the preservation of useful sprawl.

Further Brueckner (2000) highlights some drawbacks of development taxes that might arise due to the difficulty of evaluating the social costs of open-space. The problem of such a tax lays in the subjective nature of open-space benefits as considered by residents. However, overestimating of the implicit value of open-space would have substantial consequences on land-use and price and thus these kinds of policy measures should be handled with care.

Echenique et al. (2012) criticise the planning and policy practices promoting compacter urban forms for not providing credible supportive evidence that they are more efficient and less polluting than lower density development. They highlight that the *Smart Growth* principals on their side bear negative consequences, in particular socio-economic (less choice, congestion and crowding), that might outweigh the environmental impact of urban spatial expansion. Echenique et al. (2012, p.135) claim that planning policies should allow for flexibility with regard to local specificities. They find that the land-use and transport policies, favouring denser development, have no long term impact

[20] "*A UGB is a zoning tool that slows urban growth by banning development in designated areas on the urban fringe. In effect, imposition of such a boundary involves drawing a polygon around a city and prohibiting development outside it.*" (Brueckner, 2000, p.167). See Evans (2004) for more detailed presentation of UGB policies and the residential land market.

on the consumption of natural resources and energy. On the contrary they rather tend to increase housing costs and weaken economic competitiveness. With regard to protecting the environment from increased traffic pollution, Echenique et al. (2012) suggest to rather rely on technological improvements (e.g.: non-fossil energy) than on residential compaction measures.

The trade-off between air pollution emission and exposure resulting from urban form have been investigated in the analytical model of Schindler and Caruso (2014). In line with Echenique et al. (2012), they find that compaction policies can reduce consumers utility, they should always be "*addressed in comparison to expected benefits*" (Schindler and Caruso, 2014, p.21). The range of variation of exposure to pollutants can outweigh the benefits obtained from emission reduction by shorter and fewer commutes. Within the periurban area, exposure to air pollution is found to be weaker than in the more dense urban areas. Densification strategies seem less effective than an increase in public transport provision and local design strategies seem more appropriate than land-use controls or transport policies, according to Schindler and Caruso (2014).

This discussion on the causes, consequences and solutions to face periurban land-use patterns and their negative impacts on society and individuals highlights further the importance of accounting for the local scale and that urban processes and forms are the result of individual choices taken in a specific geographical, historical and political context.

Appendix B

Notes to part I: Periurban Luxembourg

As highlighted in the thesis' introduction, periurbanisation is observed around most urban areas and the causes and consequences are largely discussed in different fields. To gain insights into the spatial organisation of the Grand Duchy of Luxembourg a typology is generated by cluster analysis. Given the definition of Cavailhès et al. (2004), the periurban area can be summarised as depending on cities for the jobs while having strong rural characteristics in terms of land-use, landscapes and housing. An overview on the definition of periurbanisation in Europe can be found in Caruso (2005)[1], summarising the periurban area by two main characteristics. First, they are under urban influence and a functional link with the centre, characterized by commuting flows, exists. Second, periurban areas show a rural character and low population densities, and arise due to the benefits obtained by individuals from the coexistence of residential and agricultural land-uses.

The objectives of generating this classification are threefold. First, we aimed at identifying different morphological and functional areas in Luxembourg, and mainly the periurban area as defined above. Second, it aims at identifying if the assumptions of the monocentric model applied in the Luxembourgish context, with Luxembourg-city as the main urban area. And third, with regard to the existing typologies, accounting for the sub-municipal scale to allow a finer analysis of the local context of the land transactions.

This typology is considered complementary to the existing ones, presented in section B.1. The main originalities of this typology are that it accounts at the same time for morphological and functional characteristics and is at section scale. This finer scale accounts more accurately for the Luxembourgish context and will support the description

[1]The concept has been further detailed in appendix A.1.

of the different datasets in the following chapters[2]. The datasets underlying this analysis are discussed in section B.2, before presenting the results in section B.3.

B.1 Review of existing typologies

Different classifications of Luxembourg at municipal scale have been undertaken in the last decade by different institutions and in different contexts. The "Programme Directeur de l'Aménagement du Territoire" (PDAT) (MIAT, 2003) suggests a subdivision in six planning regions. Adding a typology of the main urban areas into "Centres de Développement et d'Attractivité" (CDA), with the objective to structure (around Luxembourg), to regenerate (South-east and former industrial region) and to develop ("Nordstad" in the northern part). No indications are provided on the criteria underlying the selection of the planning regions' extent and the identification of the different CDA.

Fehlen et al. (2003) presented a municipal classification in the framework of the national census data analysis in 2001. They applied the methodology and definitions of the French Institute of Statistics and Economic Studies (INSEE) to Luxembourg, mainly dividing the country in urban and rural areas based on economic and functional criteria. Six main urban centres emerge ($\succ$ 3,000 jobs) according to this classification, illustrated in map A fig. B.1[3]; underlining the importance of the city of Luxembourg, Esch/Alzette and Differdange. The periurban belt is mainly defined by at least 40% of residents working in an urban centre and the multi-polarised municipalities, with residents working in one of the urban areas in proximity (highlighting the former industrial south). These three areas describe the area under mostly urban influence, while the remaining municipalities are considered as under rural influence. Hence, a clear distinction between the northern and the southern part of the country can be observed.

More recently, Carpentier (2006) presented a typology based on urban structure, accounting for daily mobility, with the scope to evaluate the role of space in the differ-

[2]At some point considered to be used as fixed effects, however rejected with regard to the critics raised towards this method, see appendix D.5.

[3]A simplified version of their original map.

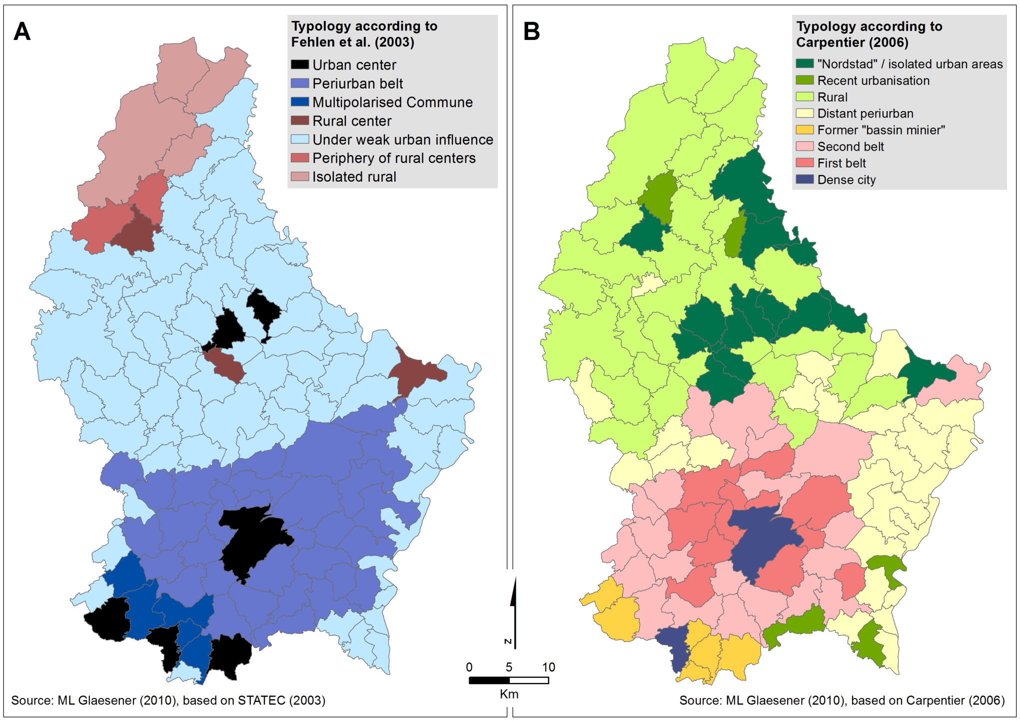


Figure B.1: Typologies for Luxembourg

entiation of practices and mobility in Luxembourg. Via a principal component analysis (PCA) and considering five factors, based on 20 variables[4], a hierarchical classification at municipal scale was generated (Map B fig. B.1). The municipalities are classified according to their level of urbanisation, the construction period and the employment areas focussing on urban structure and daily commuting flows. Eight types of municipalities are identified, belonging either to the urban or the more rural area. Within the rural region Carpentier (2006) distinguishes between the "Nordstad" and the isolated urban centres, the municipalities recently urbanised as well as the distant periurban area and the rural area. Within the area mainly under urban influence, distinction is made between the former industrial south, a first and a second belt around the densely urbanised cities.

These different classifications of the municipalities do not, except for Carpentier

[4]Morphological: land-use, construction types and period; Functional: commutes job/residence.

(2006), account simultaneously for all the criteria retained in our definition of urban and periurban areas. Further, they are at aggregated municipal level and do not consider the sub-municipal level.

B.2 Morphological and functional criteria

In this section a brief description of the different variables considered in the cluster analysis[5] to identify groups of sections with similar morphological and functional characteristics. Data at two hierarchically nested levels (sections and municipalities) had to be combined due to data limitations. At section scale we account for the morphological structure while the functional relationship with the two major urban agglomerations will be accounted for at municipal scale.

B.2.1 Morphological variables

The different morphological variables were generated in ArcGIS10 (2010) based on the raster images from the digital land registry plan (PCN) provided by the Administration du Cadastre et de la Topographie (2008b)[6]. A first variable considered was the overall built-up density (URB_s) per section (Table B.1 and Map A fig. B.2). The built-up area (B) regroups all constructions considered as "buildings"[7]. URB_s represents the part of area covered by residential land-use compared to the total section area. A first area of high urban densities is observed in and around the municipality of Luxembourg. A second concentration of high urban densities is observed in the south-east of the country between Esch/Alzette and Differdange. Around Luxembourg, a cluster of very high density that appears to cover the whole area and link the different urban centres was identified. Further the regional centres of Clerveaux, Echternach, Ettelbruck, Diekirch,

[5]Cluster analysis is a hierarchical ascendant classification method, a multivariate statistical procedure. By starting with a database with information about a sample of entities, this method attempts to fit the entities into relatively homogeneous groups (Aldenderfer and Blashfield, 1984). We rely on the method developed by Ward in the 1960s. By joining at every stage the two clusters which yield to the lowest increase in the intra-group variation (Beguin, 1979) and allow to optimise the minimum variance within the clusters (Aldenderfer and Blashfield, 1984).

[6]Further described in section 2.3.

[7]The "*bâtiment*" shapfile regrouping all public, residential, industrial, commercial, agricultural, etc. buildings, all in all 30 categories with different functions and uses from the PCN, however not considering outdoor sport facilities (e.g., football field,...) or airport facilities (e.g., runway), that would be important features while determining continuous built-up area (Sohn, 2006).

Troisvierges et Wiltz register high densities. In the sections in the north and partly in the east register in general lower densities of urbanised area.

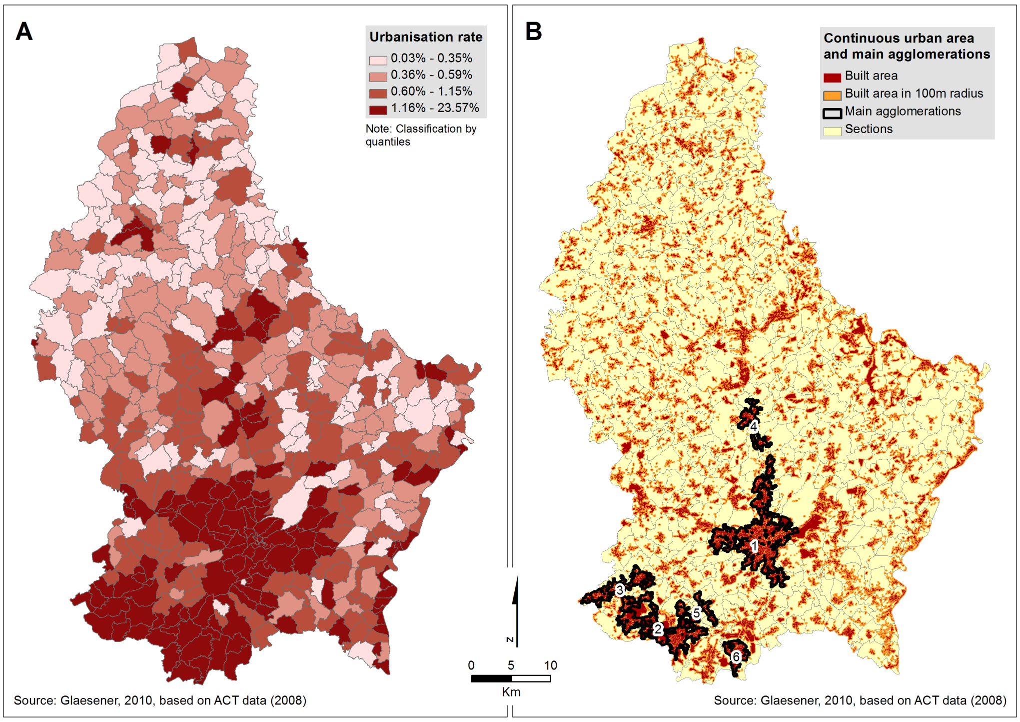


Figure B.2: Built-up density and continuous urban areas

According to INSEE, an urban area is defined by a continuous built-up area without gap exceeding 200m between buildings (INSEE, 2010). The 200m threshold is frequently applied and commonly accepted in literature to describe a continuous urban area in Europe (Le Gléau et al., 1996). To identify the extent of the major urban agglomerations in Luxembourg the continuous built-up area within a 200m radius was computed (in red Map B fig. B.2), using neighbourhood statistics in ArcGIS10 (2010)[8]. In table B.2 the six largest continuous urban areas in Luxembourg (1-6 on Map B fig. B.2) have been listed.

The regional centres composing the "Nordstad"[9] (MIAT, 2003) were expected to

[8]By considering a 100m circular buffer around the built-up area (orange) we identified the main continuous built areas in Luxembourg (black outline Map A fig. B.2).

[9]Mainly Ettelbruck and Diekirch.

Name	Description	Formula
URB_s	Density of built area	$= B/S_s$
$rURC_{LUX}$	Part of built area continuous to Luxembourg divided urban density	$= URC_{LUX}$ / URB_s
$rURC_{ESCH}$	Part of built-up area continuous to Esch/Alzette divided urban density	$= URC_{ESCH}$ / URB_s

Note: s = section ; B = built cells (5x5m) ; S_s = Sum of cells per section (5x5m); URC = continuous urban area at 200m

Table B.1: Morphological variables

out-stand as the agglomeration of the North, they however only reach ranks seven respectively ten. Morphologically spoken, the sections of these municipalities do thus not reply to the criteria of a continuous urban area and will not be accounted for as an urban agglomeration. Hence the morphological extent of the agglomerations of Luxembourg[10] and Esch/Alzette[11] are the only ones to be considered as urban agglomerations. Together they form about 16% of the continuous (agglomerated) urban area in Luxembourg.

Rank	Agglomeration*	Contiguous built area (km^2)	Part in the total agglomerated area (%)
1	**Luxembourg**	51.19	10.79
2	**Esch/Alzette**	26.23	5.53
3	Petange	12.90	2.71
4	Mersch	7.75	1.63
5	Mondercange	7.01	1.48
6	Dudelange	6.83	1.44

*Note:** Agglomerations are named after the major urban centre to which they belong, with regard to the spatial extent of the urban area around Esch/Alzette, we will rather refer to it as Agglomerated South in what follows.

Table B.2: Main areas of continuous urbanisation

In the cluster analysis, the density of urban area within the sections belonging to the two major agglomerations, $rURC_{LUX}$ and $rURC_{ESCH}$, illustrated in Maps A and B in figure B.1 will eventually be considered. This variable accounts for the strong morphological link between the urban centres and their respective neighbouring sections. The ratio indicates the part of built area (B) belonging to the continuous urban area of the agglomeration. Maps in figure B.3 highlight the appropriateness of the section

[10] Covering among others sections of the municipalities of Walferdange, Steinsel, Strassen, Hesperange.
[11] Covering sections of the municipalities of Differdange, Sanem and Schifflange.

scale, as in general only some sections of the neighbouring municipalities belong to the continuous urban area of Luxembourg and Esch/Alzette. Some unexpected interruptions of the continuously built area can be explained, especially for Luxembourg, by artificial and natural barriers (e.g., the highway or the "Grengewald" a large forest north west of Luxembourg-city).

The morphological criteria eventually kept in the cluster analysis describe on the one hand the general construction density (URB_s) and on the other hand the morphological links of sections with one of the two major urban agglomerations: Luxembourg and Agglomerated South.

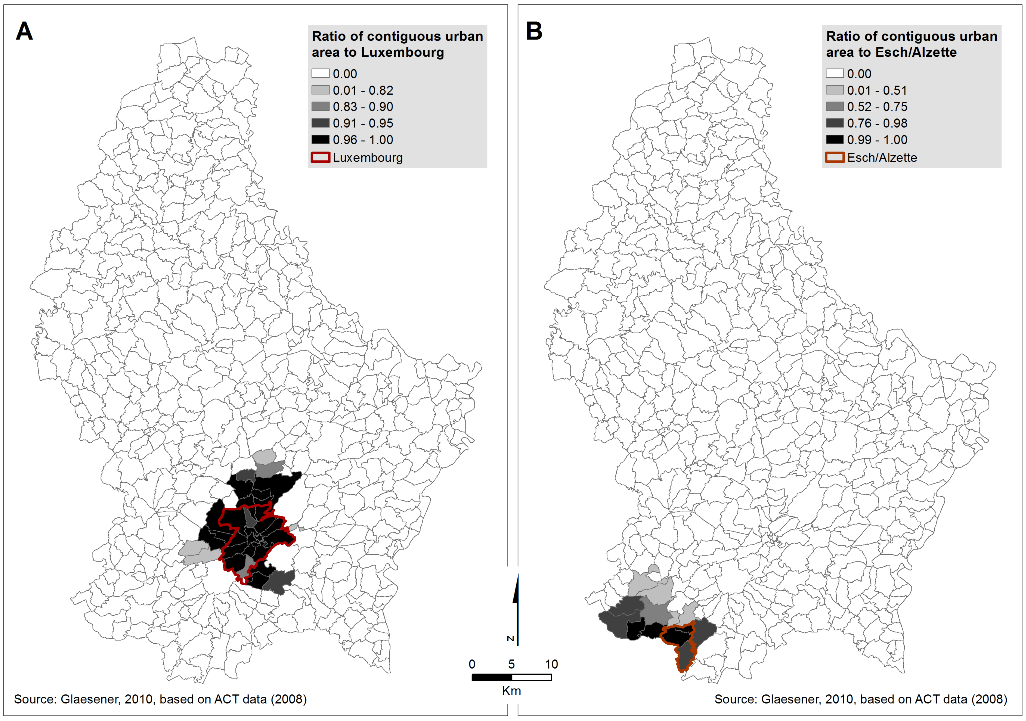


Figure B.3: Part of built-up area contiguous to the major urban areas

B.2.2 Functional variables

According to the presented definition of the periurban area, it is characterised by the functional links it maintains with an urban core area, considered as the main employment centre (Cavailhès et al., 2004). To translate this link, the commuting flows (home-job) from the census 2001[12] have been used to approximate this link between the different municipalities and the main urban agglomerations. The aggregated scale of this data at municipal level is a limitation in this analysis, however no information on the commutes between sections is available. Further, the missing international commuter flows from the border region to the different municipalities can be considered a further limitation.

A commuter is defined as a person leaving the municipality of residence for his job, working outside the municipality of residence by STATEC. Distinction is made between the entering or leaving commuters, from the respective point of view of the municipalities where the job is located or the municipality of residence. The three variables considered to describe the functional link of the municipalities with the major agglomerations identified in section B.2.1 are presented in table B.3. The maps in fig. B.4 represent the part of resident commuters to either Luxembourg (Map A: NAV_{LUX}), Esch/Alzette (Map B: NAV_{ESCH}) and eventually those commuting within their municipality of residence (Map C: NAV_{INTRA}). Maps A and B illustrate the dominant position of the capital as major destination for commuting flows and its position as a major employment centre. All municipalities register commuters to Luxembourg, especially in the southern part of the country, with municipal averages above 28% of commuters working in the municipality of Luxembourg. Although most municipalities register commuters to Esch/Alzette, municipalities with more than 2% of commuters to Esch/Alzette are concentrated in the south-east of the country[13].

Further, the the intra-municipal commuters (NAV_{INTRA}) have been considered, those not leaving their municipality of residence. High percentages can be considered as a proxy for the local availability of job opportunities as well as of the more regional organisation of the area and availability of local jobs (i.e. agricultural activities). Except for the main agglomerations and some regional urban centres, municipalities registering above 17% of intra-municipal commutes are generally located in the north and in some

[12]Provided by Isabelle Pigeron-Piroth based on STATEC (2001) data, at the moment of this analysis (2010), the 2011 census data was not available.

[13]Except for the municipality "Lac-Haute-Sûre" in the north, with 2% of its leaving commuters working in Esch/Alzette.

municipalities along the 'Moselle Valley'.

Name	Description	Formula
NAV_{LUX}	Part of leaving commuters, working in Luxembourg	$= T_{iLUX}$ / T_i
NAV_{ESCH}	Part of leaving commuters, working in Esch/Alzette	$=T_{iESCH}$ / T_i
NAV_{INTRA}	Part of intra-municipal commuters	$= T_{ij}$ / T_i, if $i = j$

Note: i = origin (municipality of residence) [$_{iLUX}$ = origin to Luxembourg]; j = municipality job is located in, T_{ij} = commute between origin and destination ; T_i = sum of entering commuters in municipality i.

Table B.3: Functional variables

A high correlation between $rURC_{ESCH}$ and NAV_{ESCH} is observed in table B.4. This correlation is most probably due to the important part of commuters to Esch/Alzette living in the sections belonging to the continuous urban area of this agglomeration. The part of commuters to Esch/Alzette was thus omitted, since this municipality does not seem to be a particularly attractive employment centre at national scale nowadays.

Eventually, we thus considered five variables in the cluster analysis[14]: NAV_{LUX}, NAV_{INTRA}, URB_S, $rURC_{ESCH}$ and $rURC_{LUX}$, results are presented in the following section.

	NAV_{ESCH}	NAV_{LUX}	NAV_{INTRA}	URB_s	$rURC_{ESCH}$	$rURC_{LUX}$
NAV_{ESCH}	1.0000					
NAV_{LUX}	0.1454	1.0000				
NAV_{INTRA}	-0.1014	-0.1343	1.0000			
URB_s	0.2909	0.3705	0.4532	1.0000		
$rURC_{ESCH}$	**0.6780**	-0.0069	0.0169	0.2451	1.0000	
$rURC_{LUX}$	-0.0484	0.4888	0.4001	0.5952	-0.0381	1.0000

Table B.4: Correlation among variables

B.3 Classification results

To illustrate the spatial organisation of Luxembourg, a sub-municipal typology based on morphological characteristics and the functional relationship of the municipalities with the capital was generated. The variables included in the cluster analysis are: the density of urbanised area per section, continuous urban area to Luxembourg and Esch as well as commuters to Luxembourg and those working in their home municipality. Five classes

[14]The cluster analysis was conducted in STATA 11 (StataCorp, 2009).

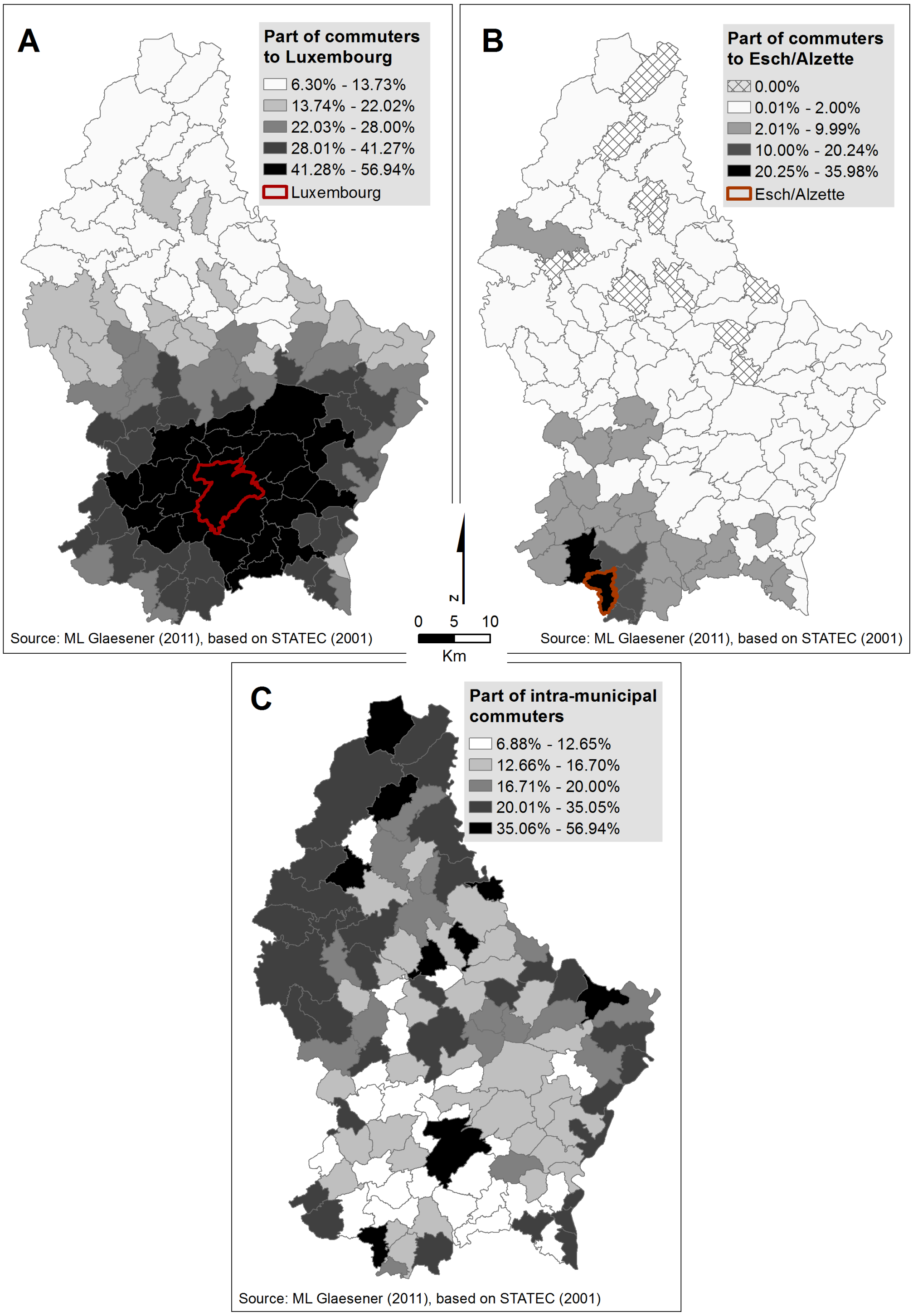


Figure B.4: Part of commuters

have been kept after a graphical analysis of the dendrogramm, identifying the major morphological and functional clusters presented in table B.5. Figure B.5 illustrates the distribution of the different clusters through the country. The agglomeration of Luxembourg is in the centre-south of the country and mainly surrounded by the periurban a belt around the agglomeration of Luxembourg and the northern part of Esch/Alzette. The close periurban area covers 156 sections located in 40 municipalities in a radius of approximately 20 kilometres around Luxembourg.

	Obs.	NAV		URB_s	rURC	
		LUX	*INTRA*		*ESCH*	*LUX*
Agglo. of Luxembourg	36	0.549	0.371	0.073	0.000	0.968
Agglo. South	10	0.265	0.222	0.060	0.909	0.000
Rural area	186	0.111	0.238	0.006	0.000	0.000
Distant periurban	133	0.244	0.233	0.011	0.000	0.000
Close periurban	156	0.424	0.138	0.011	0.001	0.001

Table B.5: Cluster analysis results

Distinction is made between the two major agglomerations, as expected based on the contiguity criteria fixed by $rURC_{LUX}$ and $rURC_{ESCH}$. The agglomeration of Luxembourg is especially defined by its high building density (7.3%) and intense functional links with the capital (NAV_{LUX}). Despite the aggregated commuter information, the extent of the agglomerated areas exceeds the administrative boundaries of the respective municipalities (fig. B.5), and thus confirms the usefulness of section scale.

The sections of the Agglomerated South presents high urban densities (6%) and a strong functional link with the capital (26.4%). The agglomeration largely exceeds the municipal boundaries of Esch/Alzette, especially to the west, including the municipalities of Differdange and Schifflange.

The rural area is characterised by low urban densities (0.6%), mainly located in the north of the Grand Duchy. The functional links with Luxembourg are limited, the part of commuters to Luxembourg being on average around 11%. On the one hand the part of intra-municipal commuters is high (23.8%) and as expected there are no morphological links with the agglomerations.

The distant periurban area is distinguished from the rural area by a denser urban structure and increased part of commuters to Luxembourg, although the intra-municipal commutes are still important. It can be seen as the interface between the rural area and the zone defined as 'close periurban'. We assume that this might be considered as the rural area in transition, since distinction between distant and close periurban area is mainly based on the functional link with the agglomeration of Luxembourg while there are no differences in the density of urban structures. In the close periurban area a morphological link with either agglomeration exists although it is not very pronounced. The close periurban area is characterised by a very strong functional link with the capital (43% of commuters to Luxembourg) and a low part of intra-municipal commuters.

B.4 Concluding remarks

Five distinct regions based on these morphological and functional criteria have been identified, illustrated in figure B.5. The northern part of the country corresponds to a more rural structure while in the south is marked by the capital and the former industrial area in the south-west. Despite a second agglomeration identified in the southern part of the country, we do not think that it puts into question the monocentric organisation of the country around the agglomeration of Luxembourg as observed today.

Compared to the typologies available for Luxembourg, presented in section B.1, the spatial organisation of Luxembourg is mainly confirmed. A main originality of our typology lies in its ability to capture the extent of the agglomeration beyond the municipal boundaries. The distinction between a close and distant periurban area as well as the Agglomerated South, as presented especially by Carpentier (2006), is confirmed. A major difference of our classification is that the "Nordstad" is not considered a major urban centre.

The unavailability of fine scale commuting data is a limitation to this analysis. Although a finer consideration of the daily commutes would certainly have added precision, it is assumed that it would not have led to substantially different results. The choice of five classes might as well be subject to discussion, since a lower level of aggregation would

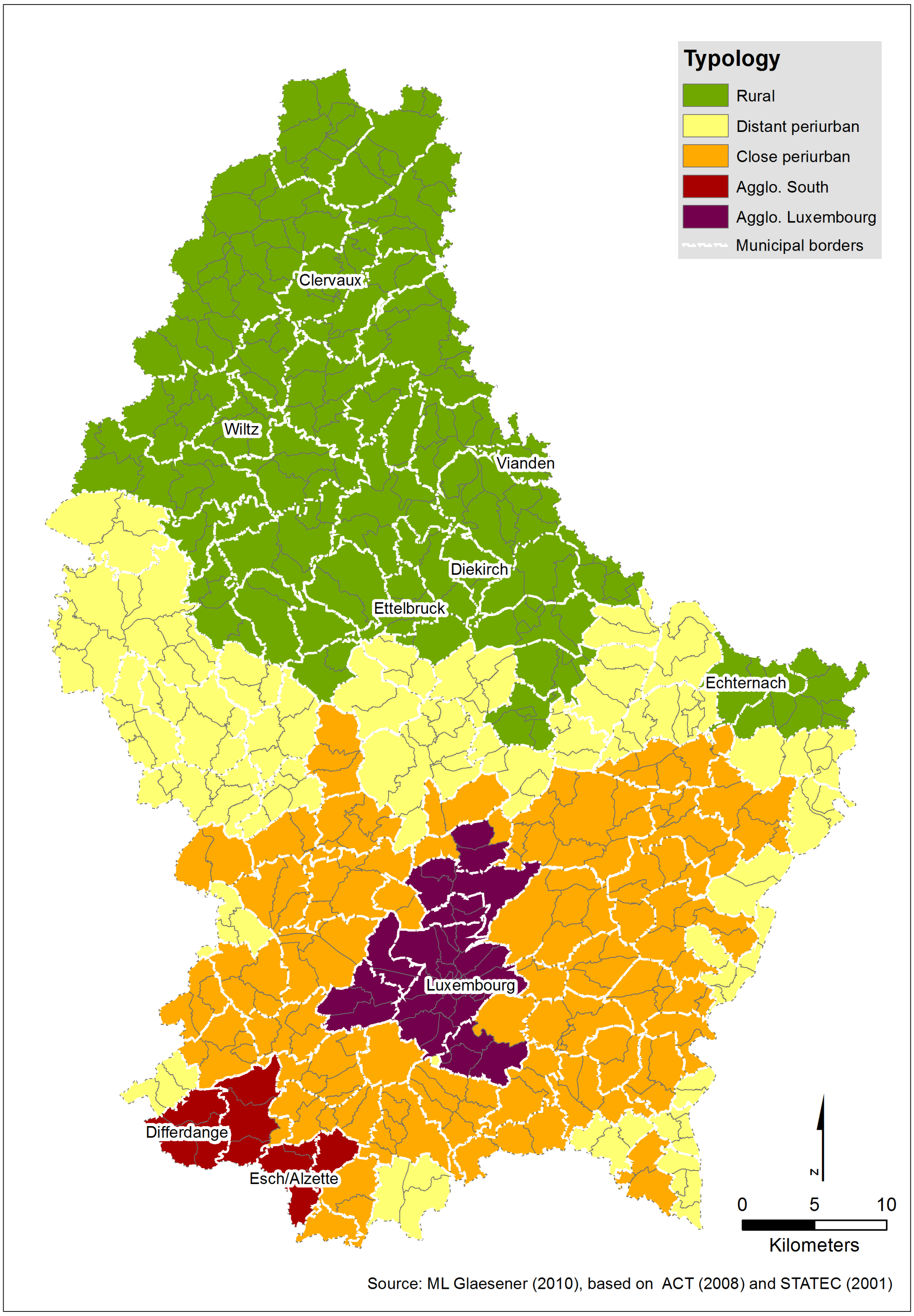


Figure B.5: Functional and morphological typology have highlighted some smaller regional urban centres[15]. With regard to the dominant position of Luxembourg-city, this choice is justified in the framework of our research, considering the residential land market at national scale.

With regard to the spatial extent of the periurban areas, the possibility of this area crossing the national border is most probable. It would have been interesting to extend the study area in order to identify the total extension of Luxembourg's area of influence and border crossing periurban area, however this was beyond the scope of this thesis.

Although the proposed typology accounts for a restricted amount of the sections' morphological characteristics, the selected variables allowed an appropriate classification of the sub-municipal units of the study area. A classification is a simplification of the observed reality, it should allow insights into the overall organisation of the country according to the morphological and functional criteria considered.

[15] Highlighted by the labelled municipalities in figure B.5, these regional urban centres would have corresponded to the rural CDA identified in the PDAT (MIAT, 2003)

Appendix C

Notes to part I: Data generation and clean-up

In this chapter, some additional information will be provided regarding the different datasets presented in part I. Although the different contextual variables have been collected at different dates[1], it is assumed that the considered averages remained rather stable over time and space and will thus not have an impact on inference (Chasco and Le Gallo, 2013). In this appendix further information is provided on the different datasets presented in part I.

C.1 AED dataset clean-up

The real estate transaction dataset was provided by the Administration of Deeds (AED)[2]. The dataset includes all transactions registered by the notaries at the AED in the period between January 2007 and December 2011, the latest update was provided in June 2012[3]. The overall quality of the database is limited,its purpose is purely administrative and was hence not created for the purpose of statistical or econometric analysis. Some substantial clean-up was necessary to make it suitable for our research. The general description of this dataset can be found in section 1.1, here some additional details, especially on the different steps of data management, will be given.

Observations are localized at sub-municipal scale, but will only be represented at aggregated municipal level due to data privacy constraints. Address-based localisation of transactions is not available due to data privacy concerns, the data privacy commission (CNPD) had doubts on the need of providing the fine location of transactions for this

[1] AED from 2007 to 2011, IGSS 2007, Accessibility 2010 and 2012, STAETC 2001 and 2007.

[2] The database is managed by the "Centre des Technologies de l'Information de l'Etat" (CTI).

[3] However, there might still be some transactions not figuring in the dataset, since there is a certain decay in the registration.

research. Consequently, we faced some problems related to the aggregated level of data.

From the initial dataset some observations had to be eliminated for ease of interpretation. Only developable land transactions[4] were considered in this study. A residential land transaction can be defined as a parcel of land available for construction[5], located within the municipal building perimeter and not supposed to be occupied by any construction at the moment of transaction. Land with other uses or other purposes (e.g.: forest, road, public green space) was not considered.

To summarise the transaction information under a unique ID (*NoTrans*), some transactions had to be deleted:

- if with *Quotité* was $\prec$1.00
- if covering more than one section
- if more than one legal framework type was registered
- if transaction date was prior to 2007

For some transactions a distinct price was provided for "*Terrain*", indicating that only a part of the total transaction price was spent for the land, often these were "complex" transactions, where the buildings were sold separately. In this case (almost 7% of transactions) the price indicated for "Terrain" was considered instead of the global price. This initial dataset contained 8,185 residential land transactions. Main summary statistics before the clean-up are represented in table C.1. Only by minima and maxima, the many extreme values reveal that there are some problematic transactions, which do not seem to describe standard land transactions or might be encoding errors. The three main variables (price, surface, price per square metre) were examined in order to identify what transactions should be kept for the analysis and fixed some conditions to exclude outliers and to normalize the distribution of the dependent variable (*lnPrice*), without dropping too many observations.

Only land transactions were considered where the consumer purchases the totality of the parcel. We do not consider transactions in which the consumers buy only one pro-

[4]Throughout the thesis we refer occasionally to them as residential land transactions, but in fact we don't know whether the "end-use" will be for residential use or any other use (e.g.: offices, retail).

[5]Code 500 "place" in the Administration du Cadastre et de la Topographie (2008b).

C.1. AED dataset clean-up

Variable	Obs	Mean	StdDev	Min	Max
Price (€)	8,185	394,864	1,102,182	1	50,000,000
Price per m^2	8,185	3,281	31,222	0	1,350,000
Surface (are)	8,185	6.78	18.93	0.1	946.70

Table C.1: AED data before clean-up

portion of the parcel as registered in the land register, as might be the case of apartment dwellings for instance; hence the final dataset should be cleaned of land transactions bought in the framework of an apartment sale. Eventually, all transactions with price per square metre between percentile 10 and 95 and price and surface between percentile 10 and 99 where kept. Further, those transactions located within sections with less than four transactions have been deleted. This leaves us with 6,367 observations for the reference period (2007-2011), summary statistics of the main variables can be found in table C.2, with the deflated prices. The price distribution after clean up is represented in figure C.1. Only 77.79% of the initial dataset were kept for the following study[6].

Variable	Obs	Mean	StdDev	Min	Max
dfPrice	6,367	327,724	274,920	14,454	2,878,744
dfPrice per m^2	6,367	684.44	597.19	91.31	5,103.84
Surface (are)	6,367	6.07	5.10	0.63	50.52

Table C.2: AED data after clean-up

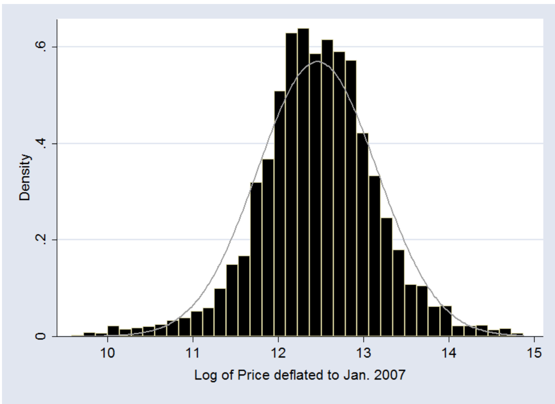


Figure C.1: Price distribution of the final dataset

[6]The final dataset covers an area of 38,658 are of land available for construction.

Other transaction specific variables

Other variables contained in the dataset inform about the transaction date, type and legal framework as well as the section the transaction is located in. Several variables were created to consider time and seasonal effects on prices:

- *Days010107*: number of days elapsed since the 1^{st} of January 2007 until transaction date
- *Year dummies*: binary variable for the year (2007-2011) of transaction date
- *Season dummies*: binary variable for the season (winter, spring, summer, autumn)
- *Registration trimester dummies*: binary variable per trimester in which the transaction was registered at the AED

Some other variables contained in the AED dataset, most of which only exist for a small part of the observations, e.g.: architect fees were only registered for 3% of the transactions, while 17% are complex transactions; the already mentioned price for *Terrain* concerns 8% of transactions (section 1.1.2). 94.11% of the observations are registered under legal framework droit1, *"droit de propriété"*, and 5% under droit7, "Droit de construction"[7]; other legal framework only rarely applies on residential land transactions. Eventually these two types of legal framework were considered, relying on the *dTerrain* variables if a non-residential building existed.

C.2 Local urban amenities: sources

This appendix provides further detail on the sources of the different datasets considered in part I section 2.2. For most of the amenities presented in that section the procedure was identical. The addresses of the amenities (e.g.: bakeries) were collected from the online phone registry[8], according to different relevant keywords (e.g.: boulanger, boulangerie, pâtisserie). The addresses were copied and geo-referenced via an online tool AUS-EMAPS.com (2010)[9] and then imported to ArcGIS10 (2010) where we joined

[7] A transaction sold as developable land, but already including a building or a construction site.

[8] Editus.lu, `www.editus.lu`, 2010-2012; Except if different sources are mentioned e.g.: for the supermarkets provided by CEPS/INSTEAD.

[9] The correct location of the input addresses had to be controlled for manually, which represented a considerable workload.

them to the section dataset. The spatial distribution of the different urban amenities has been represented in figures C.2 and C.3. Although all neighbourhood amenities were generated at micro-scale, they could only be considered at section scale, mainly as dummies or counts.

Basic service and shopping opportunities. It is assumed that residential land consumers in Luxembourg benefit from being located in or close to an area offering different products and services. In this perspective we generated a selection of basic shopping and service opportunities, described in table C.3 and illustrated in figure C.2 (maps A, C, D) and figure C.3 (map C).

Variable	Total Obs.	Mean	Max
Butcher	141	0.26	7
Bakery	199	0.37	16
Supermarket	255	0.49	12
Press & Tobacco	66	0.13	6
Bookstore	78	0.15	8
Bar	642	1.22	42
Restaurants	1,262	2.42	102
Hairdresser	460	0.87	40
Service station	330	0.62	12

Table C.3: Shopping and service opportunities

Health infrastructure. Health infrastructure is also assumed to be essential for at least part of the land consumers (related to life-cycle) and has been represented by three main categories: general practitioners, pharmacies and hospitals (see Map B in figure C.2). According to the address based online list of the "*Collège Médical*", 506 general practitioners were registered in Luxembourg in 2010. The location of the 91 pharmacies was geo-referenced based on the addresses available online from the union of pharmacists[10]. Summary statistics are presented in table C.4, providing an overview on the distribution of these variables across our dataset.

[10] "Syndicat des Pharmaciens luxembourgeois asbl", http://www.pharmacie.lu/les_pharmacies.

Variable	Total obs	Source	Section mean	Max
Pharmacy	91	Union of pharmacists	0.16	7
General practitioner	506	Collège Médical, 2010	0.96	40
Hospital	12	Editus.lu	0.02	3

Table C.4: Health infrastructure

Education. Education infrastructure was geo-referenced based on the official address lists from the Ministry of Education[11] we generated the information of primary schools as well as public and private high-schools (table C.5 and figure C.3 B). The different children day-care facilities ("crèches" and "Maison relais") were geo-referenced based on the addresses provided by the "guichet public" website (`http://www.guichet.public.lu/.../liste-maisons-relais/index.html`) in 2012.

The Grand Duchy of Luxembourg counts 50 high schools, located in 29 sections, and 153 primary schools, in 123 sections. In our dataset 9% of the residential land transactions are located in a section with a high school, while 18% of the sold properties have a primary school within the same section. The spatial organisation of primary schools is coordinated by the Ministry of Education, its distribution is quite homogeneous throughout the country with regard to the needs regarding municipal population. LISA tests revealed high spatial correlations especially around the capital as well as in the south. Additionally, high schools are mainly located in regional and main urban centres[12].

Variable	Total obs.	Source	Section mean	Min	Max
Daycare	293	Min. Educ.	0.55	0	13
School	153	Min. Educ.	0.28	0	12
High school	50	www.guichet.lu	0.10	0	5

Table C.5: Education infrastructure

Further local public services. The localisation of the town houses (*Mairies*) at section scale is based on the official address list of SYVICOL[13], see table C.6 and Map A figure

[11] "Ministére de l'Éducation Nationale" `http://www.men.public.lu/.../ecoles_services_externes` in 2009.

[12] The one high school located outside the national boundaries is the "Schengen Lyzeum" in Perl. A trans-border cooperation project not considered in the dataset.

[13] "Syndicat des Villes et Communes luxembourgeoises".

C.3. Monocentric and radial distribution along the road network for most of the amenities, while the distribution of public, educational and partly the health services show a more random distribution better covering the whole country, which can be explained by the top-down location of these services.

Label	Unit	Level	Mean	Min	Max
dMairie	Dummy	address	0.43	0	1
Postal office	113	Editus.lu	0.22	0	4

Table C.6: Public services

Correlation among urban amenities Table C.7 shows high positive correlations between the different local amenity dummies. For instance, we notice very high correlation of shopping opportunity variables among themselves, with pharmacies as well with schools. Some of these relations are obvious, e.g.: among the transport infrastructure dummies. We notice high positive correlations for *dhighschool* and *dHospital* with the municipalities of Luxembourg (LUX) and (ESCH), this can be explained by the concentration of these amenities in main urban centres (see corresponding maps).

The previously presented local amenities present high correlations among themselves, as shown in table C.7, and in space as observed by comparing the different maps in figures C.2 and C.3. For instance, high correlation of retail opportunities[14] among themselves as well as high positive correlations for health and education infrastructure with the main urban centres are observed. As it is not possible to compute accessibility measures to different amenities from residential land transactions and in order to get around the problem of multicollinearity, a diversity index was generated, as presented in part I.

C.3 Land-use data management and transformation

In this section, additional information on the different land-use datasets provided by the Administration du Cadastre et de la Topographie (2008a,b) will be detailed. Two land-use databases, both managed and provided by the "Administration du Cadastre et

[14]An additional variable generated is the count of shopping and service opportunities per section (*ssopp*) to account for the general offer of opportunities in a neighbourhood, illustrated in map D fig. C.3.

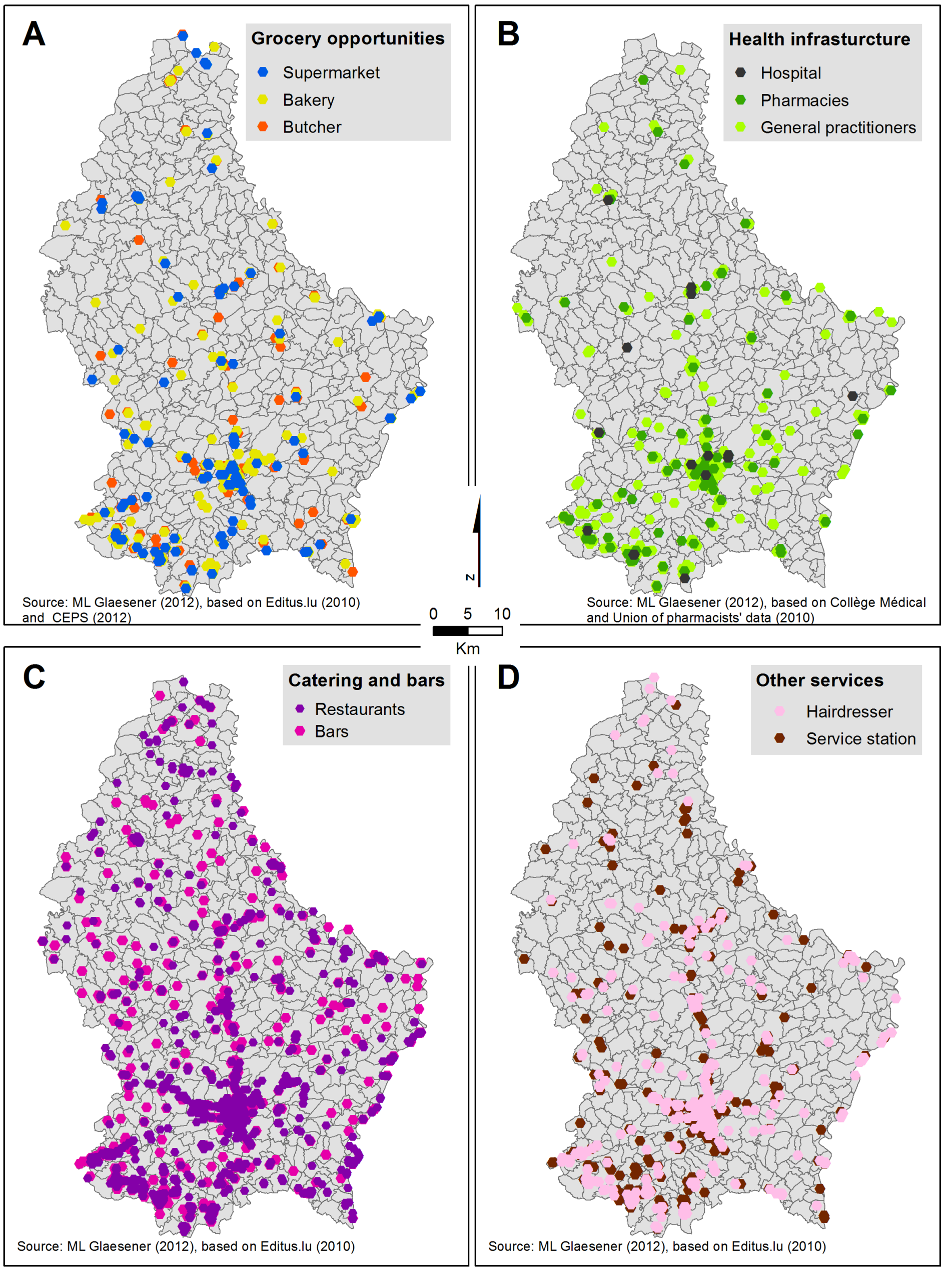


Figure C.2: Local urban amenities I

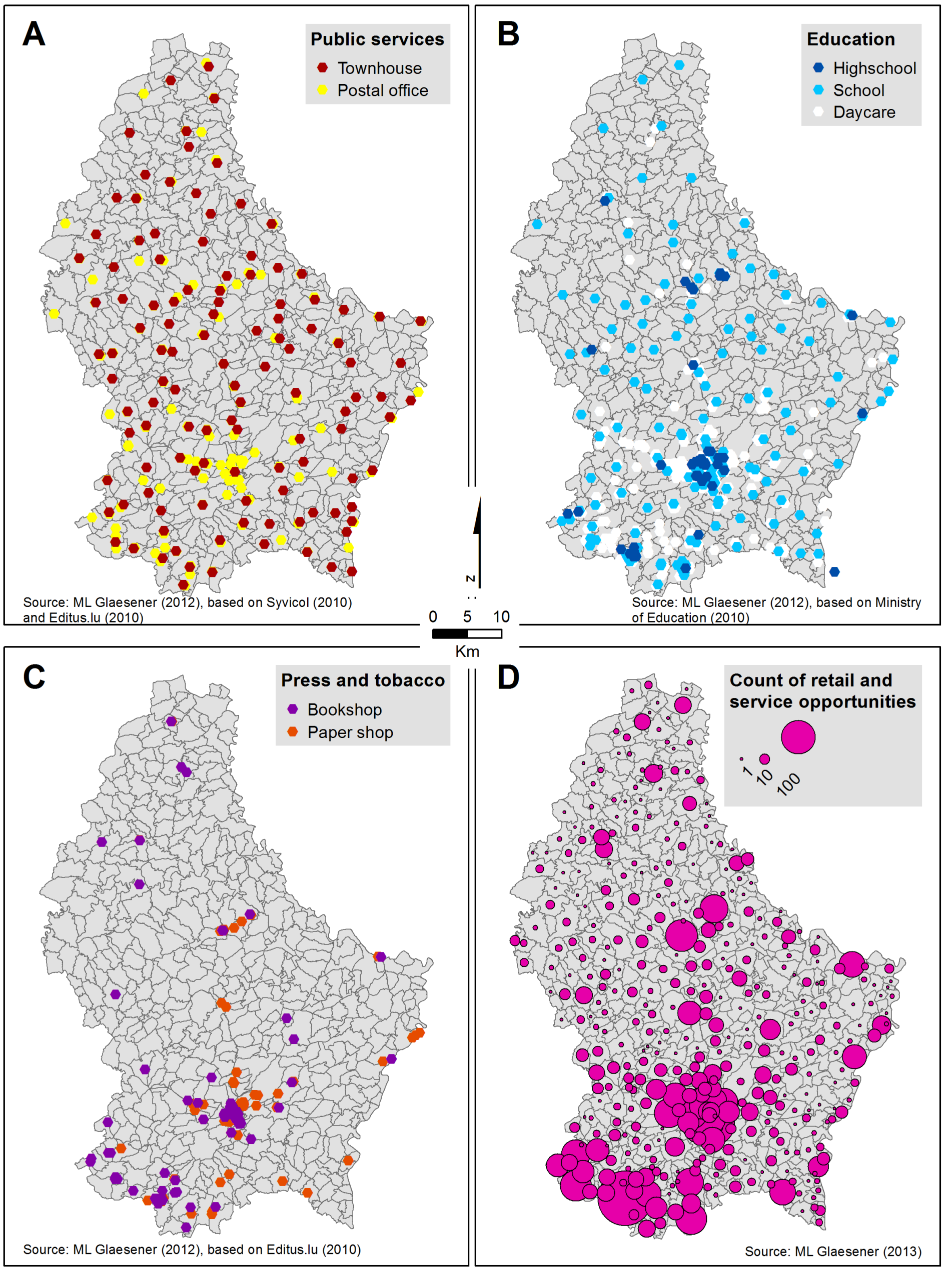


Figure C.3: Local urban amenities II

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)
Hairdresser (1)	1																			
Bakery (2)	0.92	1																		
Butcher(3)	0.78	0.79	1																	
Supermarket (4)	0.80	0.74	0.78	1																
Book store (5)	0.80	0.76	0.60	0.66	1															
Paper shop (6)	0.77	0.77	0.65	0.62	0.55	1														
Car service station (7)	0.57	0.47	0.59	0.65	0.45	0.44	1													
Bar (8)	0.90	0.87	0.74	0.73	0.70	0.68	0.49	1												
Restaurant (9)	0.87	0.84	0.66	0.70	0.81	0.79	0.44	0.78	1											
Pharmacy (10)	0.88	0.85	0.77	0.70	0.66	0.70	0.44	0.79	0.77	1										
General practitioner (11)	0.79	0.80	0.68	0.74	0.65	0.68	0.45	0.80	0.78	0.77	1									
Hospital (12)	0.45	0.41	0.28	0.34	0.36	0.36	0.19	0.39	0.35	0.35	0.46	1								
Primary school (13)	0.65	0.61	0.55	0.64	0.38	0.52	0.45	0.71	0.42	0.60	0.60	0.32	1							
High school (14)	0.56	0.59	0.46	0.49	0.48	0.59	0.34	0.52	0.46	0.52	0.61	0.53	0.49	1						
CMR (15)	0.47	0.45	0.38	0.54	0.32	0.46	0.54	0.46	0.46	0.36	0.60	0.43	0.50	0.41	1					
Town-house (16)	0.22	0.21	0.28	0.24	0.11	0.10	0.11	0.15	0.16	0.24	0.12	-0.03	0.18	-0.01	-0.02	1				
Postal office (17)	0.47	0.40	0.46	0.53	0.44	0.53	0.42	0.41	0.46	0.38	0.47	0.18	0.36	0.26	0.40	0.23	1			
Highway junction (18)	0.28	0.19	0.28	0.29	0.08	0.18	0.48	0.16	0.20	0.20	0.20	0.08	0.20	0.09	0.35	0.19	0.16	1		
Train station (19)	0.27	0.28	0.39	0.31	0.11	0.16	0.30	0.39	0.16	0.21	0.25	0.05	0.32	0.07	0.22	0.21	0.22	0.21	1	
Bus stop (20)	0.74	0.73	0.68	0.70	0.58	0.67	0.51	0.75	0.63	0.67	0.71	0.51	0.68	0.63	0.62	0.09	0.49	0.30	0.41	1

Table C.7: Correlation among urban amenities at section scale de la Topographie" have been considered. First, the "Plan Cadastrale Numérisé" (PCN) (Administration du Cadastre et de la Topographie, 2008b), a digital cadastral plan for 2008. Second, the "Base de Données Topo-Cartographique" (BD-L-TC) (Administration du Cadastre et de la Topographie, 2008a) for the same year. Both datasets classify the Grand Duchy in different types of land-uses, but with different objectives.

The PCN is subdivided into two main categories, either of type "Occupation" or "Nature". In general, the first describes the type of construction occupying a parcel (e.g.: public, residential, industrial, ...), graphically this shapefile represents the contours of buildings. The second category represents the cadastral parcels with a very detailed indication on the type of land-use it represents (e.g.: hardwood, softwood, brushwood,...). We were in general interested by the nature of land-use in this database, the zonal shapefile of the PCN being the starting point for the land-use raster datasets we created. However, the PCN does not always include detailed information on linear land-uses, such as for instance rivers and road network, these were completed by information from the BD-L-TC. The BD-L-TC dataset is based on aerial photographies resulting in different kinds of shapefiles (lines, polygons and points), subdivided into ten major topics (Administration du Cadastre et de la Topographie, 2008a)[15] further divided into 60 sub-classes of different shapes. We cross-checked the two vector databases by comparing the information on buildings, which provided the dataset on the parcels[16] considered available for construction. In a next step they were converted to raster (cell size 10m); divided in 15 main classes[17] (table C.8 in C.3 and figure C.6 (figure C.4). The abbreviations are explained in table C.8 below.

Particular emphasis was put on the identification of the land available for construction, the available plots (AP), corresponding to all parcels coded as "place a bâtir" in the PCN dataset from 2008, cross-checked by the BD-L-TC, comparing the shapefiles to the *PCNbuilt* and *BDLTCbatiment* datasets. Despite the one year difference between

[15]Road network, railways, energy infrastructure, buildings and various equipments, hydrographic network, vegetation, orography, altimetry, toponymy and administrative boundaries.

[16]Parcel, plot and land are used as synonyms for the part of land potentially considered as transaction.

[17]14 + 1the +1 describing the area we could not associate to any of the other land-uses.

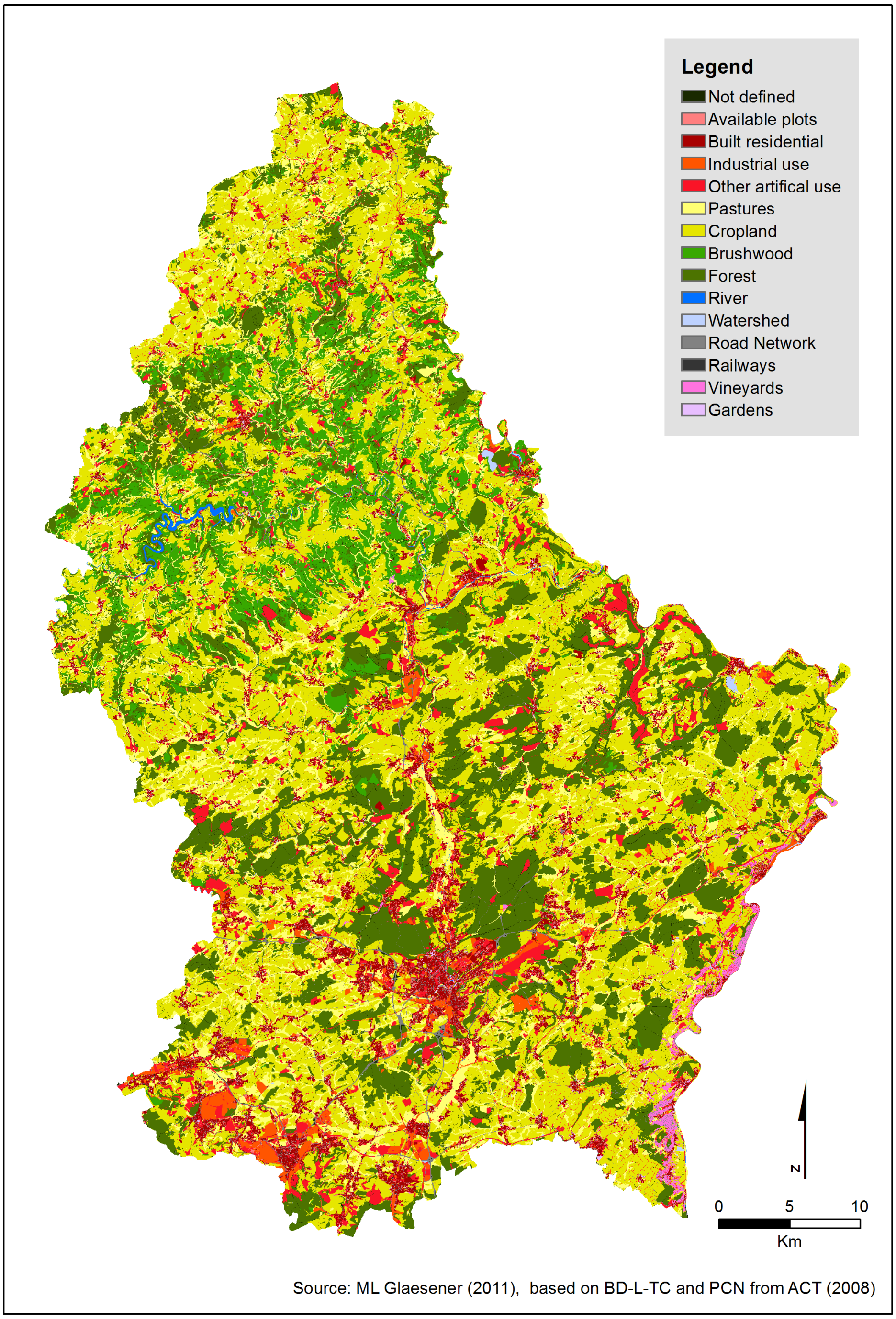


Figure C.4: Land-uses

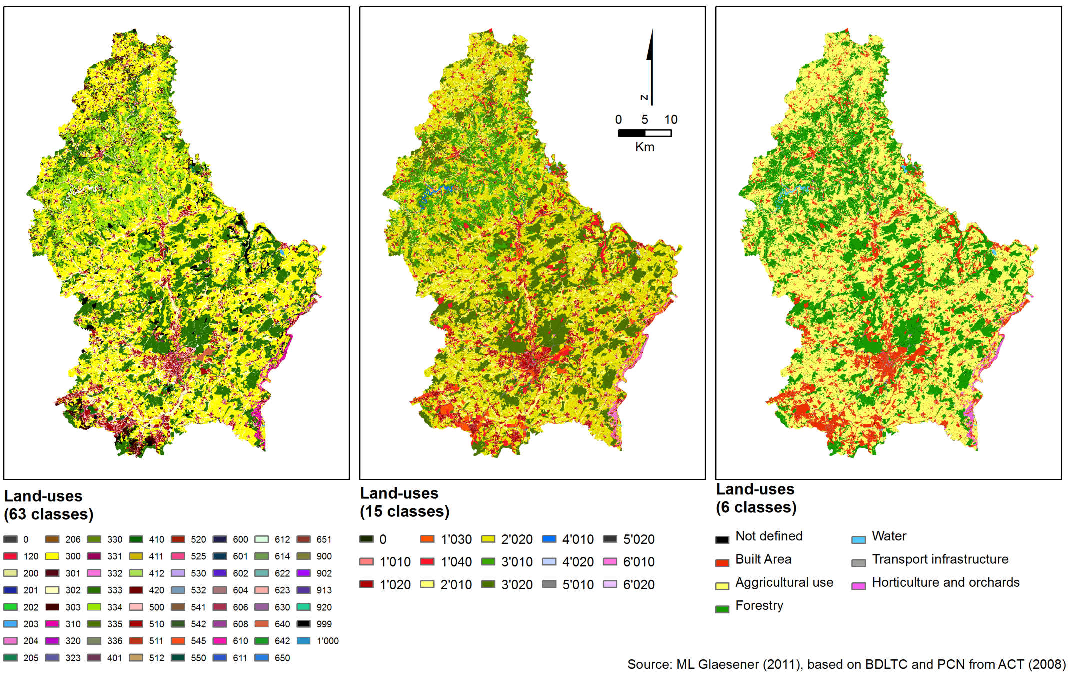


Figure C.5: Land-use : different levels of aggregation

Abbr.	Code	Name & Description
AP	1010	Available plots : land available for construction
BR	1020	Built residential: land occupied by buildings
IND	1030	Industrial land-use
BA	1040	Other artificial land
RESO	5010	Road network
RAIL	5020	Railways
AGG	2010	Agricultural land: Meadow and pasture
AGB	2020	Agricultural land: fallow and crop-land
BW	3010	Brushwood: low vegetation
FOR	3020	Forest: high vegetation
RIV	4010	River: water flows
WS	4020	Lakes: watersheds
VY	6010	Vineyards and orchards
GD	6020	Gardens

Table C.8: Abbreviations for land-uses

the Administration du Cadastre et de la Topographie (2008a) and the AED dataset, we assume that most of the listed parcels were mostly also available in 2007. By crossing the PCN and BDLTC dataset we eliminated all parcels coded as available (by the PCN) which are marked as occupied according to the BDLTC, under the assumption that land registered as occupied in the BDL-T-C was most probably sold before 2007[18] and thus does not figure in our residential land transaction database as transaction. To better approximate our dataset and to eliminate outliers and encoding mistakes, we only considered parcels bigger than 6.3are and smaller than 51.57are[19]. By this selection criteria the parcels identified as available should thus correspond to those sold in the reference period plus those still available (potential future transactions).

Further, we could not account for modifications of the building perimeter, additional parcels could have been added between 2008 and 2011 by modifications to the "Plan d'Aménagement Général" (PAG). Almost 350ha of AP are found to be located outside the PAG boundaries fixed for 2007, however these differences might be explained by the different dates of the data, between 2007 and 2008 (PCN and BD-L-TC) these might have been extended in some municipalities, increasing the total area of available parcels, or there might be encoding errors by the land registry.

Neighbourhood statistics with aggregated data

In table C.6 the different scales considered throughout this process are illustrated[20]. Although we first considered fine scale neighbourhoods (of 30m), including these different extents led to multicollinearity issues. Eventually, we accounted for the neighbourhood of 100 metres around AP. With regard to the average section size, larger extents covered entire sections, almost all cells of the country would have been considered as neighbourhood, the total area of the different land-uses compared to the total section size will thus provide similar results.

The proportion of a certain land-use was generated based on the sum of cells occupied

[18]From buying land till end of construction and building being referenced by the land registry there might be some decay.

[19]The minimum and maximum size of parcels actually sold between 2007 and 2011.

[20]For abbreviations see table C.8.

by a certain land-use in the 100m/600m[21] neighbourhood of AP cells. In a next step, the sums of cells covered in this land-use in a 100m neighbourhood around AP as well as the total cells belonging to neighbourhood have been aggregated at section scale[22].

	BuiltArea				Transport	Agri		Forest		Water		Horti	
	AP	BR	IND	BA	RESO	AGG	AGB	BW	FOR	RIV	WS	VY	GD
Neighbourhood around AP	Artifical Land Use					"Natural" Land Use							
NH100 (below 100m)													
NH600 (below 600m)													

Figure C.6: Fine and aggregated neighbourhoods

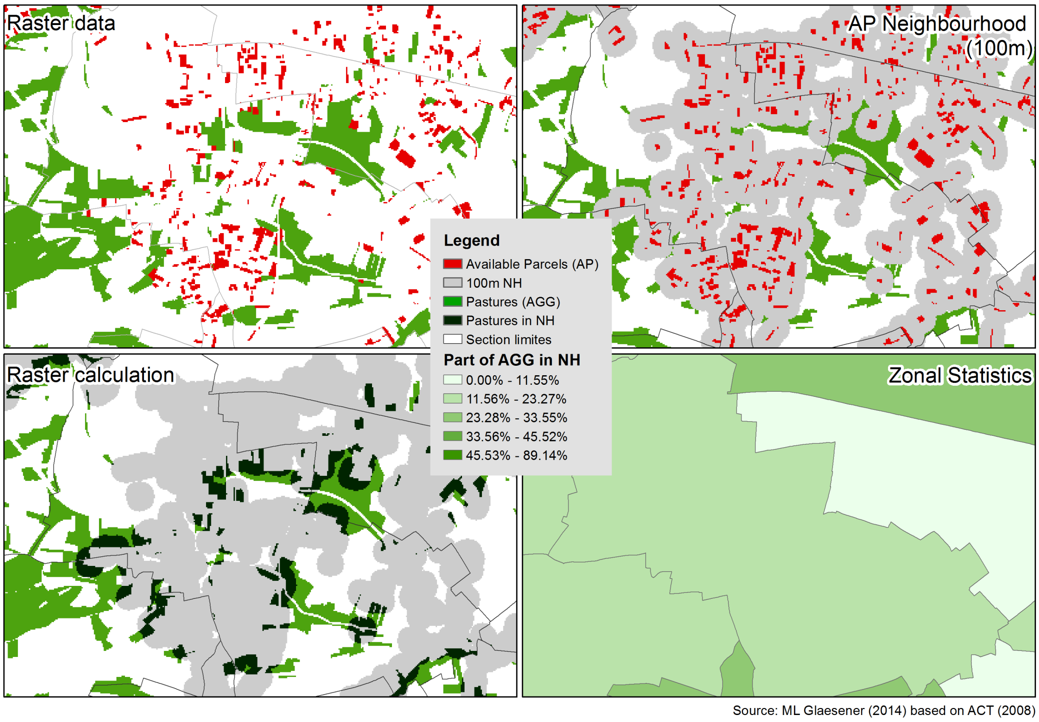


Figure C.7: Illustration of different steps for aggregated neighbourhood statistics

[21]Other thresholds have been considered but we eventually relied on the extents most commonly considered in literature (Geoghegan et al., 1997; Cavailhès et al., 2006)

[22]By "Zonal Statistics as table" tool ArcGIS10 (2010) to be able to consider them in the hedonic model. The average part of a land-use in a 100m radius 100m around the AP per section was than computed, as illustrated in "Zonal statistics" in figure C.7.

C.4 IGSS data: clean-up and geo-referencing

The IGSS[23] dataset was provided in June 2011 and covers the period of 2001 until 2008, referring always to the state of the 31^{st} of December of the year of reference. The dataset is based on the EUROMOD database, which is generated from different other datasets managed by the IGSS. The IGSS dataset contains information on all residents of the Grand Duchy who belong to the national health system, except the officials working at the European institutions. In order to include the IGSS data to our model, the data had to be aggregated at section scale (matching their "localities" to the sections). Frequently, the postal code zone and locality limits did not match (as illustrated in fig. C.8), some choices had in these cases to be made to located the observation in the appropriate section.

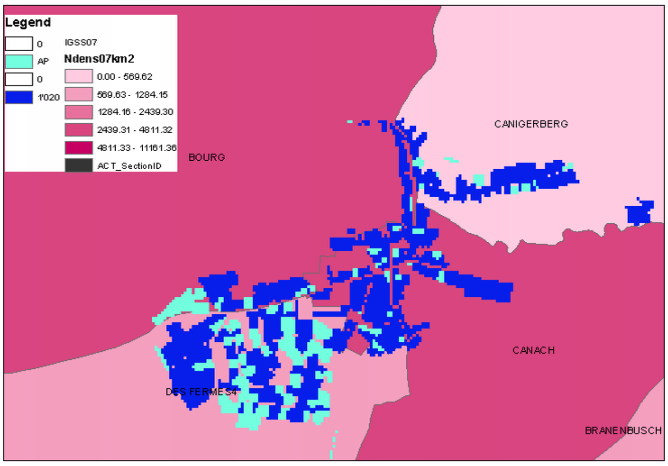


Figure C.8: Section/CP divergence

Some observations had to be deleted because of incomplete or ambiguous information. Hence, were not considered:

- entries without *ID*
- entries without *hshldID*
- entries with *hshldID*: 000000000000000000000&000000000000000000000
- *Unknown CP*: cross checked, official ZIP codes & internet

[23] Inspection Générale de la Sécurité Sociale.

C.4. IGSS data: clean-up and geo-referencing

In order to include the IGSS data to our model, we had to aggregate the data at section scale (matching their localities to our sections). Very often ZIP and locality did not match, we then kept the ZIP and located the observation in the corresponding section. First, in order to locate and convert the addresses, this data management process was supported and generalised by a python code. Still, this revealed to be quite difficult and time consuming due to many data entry errors. Often name of locality and ZIP did not match the official list of addresses.

Year	No ID	Number of obs	UnknownCP and LocliteVide
2001R	5	438,196	60+5+1
2002R	3	443,124	80+6
2003R	3	449,039	82+4
2004R	3	454,626	86+3
2005R	3	459,165	69+6+9
2006R	3	464,326	72+13+2
2007R	0	471,133	52+16+3
2008R	0	479,316	53+16+7

Table C.9: IGSS data after clean-up

According to the IGSS data in 2007, 471,133 residents were registered in Luxembourg[24]. The different available variables are illustrated in table C.10 and the explanatory variables aiming at controlling for the socio-economic context in the hedonic model have been detailed in part I.

[24] Only missing are the European civil servants.

Variable	Definition
General Information	
Idsect	Section ID
Hshld01-07	Persons per household
Total01-07	Total population
Employment status	
Farmer01-07	Farmer
SelfEmp01-07	Self-employed
Empl01-07	Employee
Pens01-07	Pensioner
Unemployed01-07	Unemployed
Student01-07	Student
Sick01-07	Sick or disabled
StatutNA01-07	Not available
Employment sector	
Privat01-07	Private
Public01-07	Public
SectorNA01-07	Not available
Employment category	
NotA01-07	Not available
BlueC01-07	Blue collar
WhiteC01-07	White collar
CivilSer01-07	Civil servant
NA	Not available
Marital Status	
UnDef	Not available
Celib01-07	Bachelor
MariE01-07	Married
Separe01-07	Separated
Veuf01-07	Widower
Divorce01-07	Divorced
Remarie01-07	Re-married
SprCor01-07	Married couple living apart
Type of insurance	
apac01-07	Primary Insured Active
appe01-07	Primary Insured Pensioner
aprm01-07	Primary Insured RMGiste
apvo01-07	voluntary Primary Insured
coac01-07	Co-insured active
cope01-07	Co-insured pensioner
corm01-07	Co-insured RMGiste
insurNA01-07	Not available
Gender	
M01-07	man
F01-07	women
Age class	
A1501-07	aged under 15
A163001-07	aged bet. 16-30
A314501-07	aged bet. 31-45
A466001-07	aged bet. 46-60
A617501-07	aged bet. 61-75
A7601-07	aged over 76
Nationality	
Lux01-07	Residents of Luxembourgish nationality
etrang01-07	Residents of non-Luxembourgish nationality
Income classes	
$i < 107$	people with income below 1SSM
$i >= 707$	people with income above 7SSM
$i1207$	people with income bet. 1-2 SSM
$i2307$	people with income bet. 2-3 SSM
$i3407$	people with income bet. 3-4 SSM
$i4507$	people with income bet. 4-5 SSM
5607	people with income bet. 5-6 SSM
$i6707$	people with income bet. 6-7 SSM

Table C.10: IGSS dataset

Appendix D

Notes to part II: Hedonic pricing method: Assumptions and identification

D.1 Main assumptions of the hedonic price function

Although hedonic analyses have been reported back to the 1920s[1], the method applied to the real estate context was introduced by Rosen (1974), based on the works of Griliches (1961) and Lancaster (1966)[2]. These papers focus on the demand side of the market, where utility is generated not by the good, but by the characteristics composing the good (Malpezzi, 2002, p.10). It is assumed that consumers have heterogeneous preferences for the different attributes composing the good. Thus hedonic prices can be defined as the implicit prices of the different attributes, revealed to economic agents from observed transaction prices and the specific amounts of attributes associated with it (Rosen, 1974). This method allows to estimate the marginal price consumers are willing-to-pay for the attributes composing the good and the marginal prices are considered to reflect "*both this pure price of land as space of given accessibility, and the value of local neighbourhood characteristics that are attached to the particular plot*" (Cheshire and Sheppard, 1998, p.361).

The hedonic approach is a revealed preferences analysis, assuming that consumers' preferences for different amenities are revealed by their purchasing behaviour, based on the actual price paid for properties (Nelson et al., 2004)[3].

[1] A detailed historic review of the hedonic pricing method has been put forward among others by Taylor (2008); Sopranzetti (2010)

[2] Short comparison of Lancaster (1966) and Rosen (1974) theory can be found in Kostov (2009, p.55). Malpezzi (2002, footnote p.11) points out the distinction between UK and US scholars, the former more often refer to Lancaster (1966) as fundamental reference, while the latter more frequently refer to Rosen (1974).

[3] In opposition to stated preferences where consumers are asked about their preferences and which relies on surveys directly asking for the individuals' preferences for economic goods and services (Ge-

The assumptions presented in this section essentially rely on the seminal paper of Rosen (1974)[4].

Since the goods are differentiated, the price on the market is not uniform and depends on the consumers' preferences for alternative packages. Residential land consumers can affect the price they pay for the parcel by the attributes they choose, however they are price-schedule-takers, thus they cannot affect the equilibrium price in the competitive market. The prices of residential land are demand determined and the supply is assumed to be fixed in the short run (Palmquist, 2006)[5]. All consumers are assumed to be perfectly and identically informed about the amount of attributes embodied in all the available packages. A wide variety and a sufficiently large amount of alternative packages is available, among which different choices can be made. Arbitrage is assumed impossible, thus the packages are indivisible and repackaging is impossible (or at least very expensive). Second hand markets are ignored, as residential land can be considered as pure consumption good. The sellers' identity is irrelevant to the purchase decision. Further, the consumer is assumed to purchase one good, additional land transactions by the same consumer would enter the utility function separately. With regard to these assumptions, residential land is considered a composite good and the equilibrium price of residential land (Y) is, as the result of an equilibrium of demand and supply, a function of the different characteristics composing the good (D.1).

The utility function (D.2) is maximized under a budget constraint (D.3), where y is the part of income spent on residential land (P_i) and all other consumed goods c. The linear hedonic function takes the following form[6]:

$$Y = \beta_0 + \beta_1 X + \epsilon \tag{D.1}$$

oghegan, 2002, p.92). In the stated preference context, the contingent valuation method is frequently used (Cavailhès et al., 2006). For an overview of stated preference methods for open space valuation we refer to McConnell and Walls (2005).

[4]Recent and more detailed presentations of the assumptions underlying the hedonic price method can be found among others in Cavailhès (2005), Bowen (2001), Palmquist (2006).

[5]Although Rosen (1974) shows that it can be easily modelled in the framework of perfect competition.

[6]In eq.D.1 a simplified formalisation is presented compared to the formalisation presented in chapter 3 eq.(3.2).

Consumers obtain utility (u) by consuming a good (i) composed of a bundle of characteristics, X, added to all other consumed goods, c. The consumers' preferences for alternative packages are heterogeneous, translated by the taste parameter α. The utility of a residential land parcel is a function of its attributes (Bowen, 2001).

$$U(\alpha, c, X) \tag{D.2}$$

Utility is maximised under lowest cost[7] and the utility function is maximized under a budget constraint (eq.D.3), where y is the part of income spent for land (Y) and all other consumed goods (c).

$$y = Y + c \tag{D.3}$$

The marginal price consumers are willing to pay for the attributes composing the parcel can thus be estimated via ordinary least squares estimation (OLS). Where Y is the transaction price of residential land and β_0 the intercept, the overall transaction price after all explanatory variables are accounted for. X is a vector of explanatory variables, in our case the structural and location-specific attributes of the land parcel, with β_1 representing the regression estimates. The error term (ϵ) represents all the factors not explicitly accounted for in the model.

By first order maximization conditions of the consumers' utility (D.2) and under the given budget constraint (D.3), this approach allows to estimate the marginal price (p_i) consumers pay for an attribute of the land parcel (eq.3.3 p.104) (Cavailhès, 2005).

D.2 Model specification and identification

Regression analysis using ordinary least squares estimation (OLS) has been common practice in the hedonic pricing context. For unbiased estimation results to be obtained, several assumptions need to be met: exogenous independent variables, constant variance across sample data and absence of correlation between the independent variables and the

[7]If two objects offer the same attributes but are sold at different prices; only the less expensive is considered by consumers.

error term (Wilhelmsson, 2002)[8]. Anselin and Lozano-Gracia (2009) present an overview of these methodological challenges in the hedonic pricing context. Firstly, with regard to the implementation of the hedonic pricing method and secondly with regard to the explanatory variables focussing on the problem of spatial scale and endogeneity.

D.2.1 Heteroskedasticity, non-linearity and functional form

A major assumption to be met in linear regression is the constant variance of the error term. In the hedonic pricing context this is violated when the variance of the error term differs between different types of properties (Fletcher et al., 2000). Different functional form specifications are proposed to deal with non constant variance of the residuals. To obtain heteroskedasticity robust estimates different tests[9] and corrections in the covariance matrices have been developed by White (1980), providing corrected standard errors. In R Core Team (2013) the "*sandwich*" package Zeileis (2004, 2006) provides means to estimate such heteroskedasticity robust standard errors according to White (1980). Heteroskedasticity might as well be a problem for spatial dependence identification Anselin et al. (2004).

Prices formed on a real estate market are non-linear (Rosen, 1974) as linearity would only occur if arbitrage were possible (Freeman, 1979, p.156). For example, the price of land does not vary linearly with regard to its size, this non-linearity is justified for instance by fixed transaction costs (Cavailhès et al., 2006, p.87). Or as put by Malpezzi (2002), the costs of adjustment ultimately yield to the non-linearities that are observed. This non-linearity has to be accounted for in the model specification through the functional form, for instance by logarithmic transformations (as done for accessibility or size measures).

In hedonic pricing models the dependent variable is generally the total transaction price (e.g.: land price, dwelling price or rent), although some authors consider the price per square meter instead (Ahlfeldt, 2011; Abelairas-Etxebarria and Inma, 2012). The

[8]An overview of regression analysis and related assumptions are further detailed in Verbeek (2008) and Wooldridge (2009).

[9]The most common test to identify non constant variance of OLS residuals in the hedonic pricing context is the Breusch-Pagan (Orford, 2000; Goffette-Nagot et al., 2011; Cavailhes and Thomas, 2012).

appropriate functional form for the hedonic pricing method has been largely discussed in hedonic literature (Fletcher et al., 2000; Anselin and Lozano-Gracia, 2009; Ahlfeldt, 2011; Kuminoff et al., 2010; Dubé et al., 2011). A detailed overview on different functional forms and how to test for them as well as a semi-parametric alternative are proposed by Angelin and Gencay (1996)[10]. McMillen (2010) discusses more flexible approaches (e.g.: series expansion and non-parametric estimators), but these will not be further considered in this thesis.

According to Anselin (1988b), the Semi-Log form performs best as the log of the dependent variable fosters a normal distribution and allows to cope for the problems of residual heteroskedasticity. However, as put by Kostov (2009), the functional form of the hedonic price function carries a great deal of uncertainty and the incorrect choice may be a source of residual cross-sectional autocorrelation. Dubé et al. (2011) claim that further research on the appropriate functional form has to be undertaken.

As highlighted by Can (1990), the Box-Cox transformation might provide the best empirically fitted parameters, hence it has been considered in several hedonic studies (Brown and Rosen, 1982; Halstead et al., 1997; Cheshire and Sheppard, 1995, 1998; Cavailhès, 2005; Cavailhes and Thomas, 2012). Halstead et al. (1997) implemented a hedonic price model to estimate the marginal value of landfill development and compared different functional forms (linear, Log-Log and Box-Cox). They found that results of the Log-Log and Box-Cox results are quite similar. Halstead et al. (1997) concluded that the choice of the functional form in hedonic pricing models is subject to the underlying data and that the hedonic models may require different functional forms depending on the context and available data. However, Box-Cox transformations make interpretation of estimates difficult, especially in presence of binary variables (Can, 1990) and are not implemented if considering spatial dependence (Kim et al., 2003; Bala et al., 2011).

With regard to the purpose of the regression functions to estimate the marginal prices, among others Can (1990) and Anselin and Lozano-Gracia (2009) advocate the

[10] Angelin and Gencay (1996) present the RESET test to test for the appropriate functional form, used among others by Abelairas-Etxebarria and Inma (2012) and discussed in Verbeek (2008).

use of simpler functional forms (e.g.: Linear or Semi-Log) as they perform better than more complex forms and with regard to the ease of results interpretation. In most applied work, this is the path taken (Ahlfeldt, 2011; Abelairas-Etxebarria and Inma, 2012; Kadish and Netusil, 2012). The aim of this thesis being to gain insights on the marginal effects of the geographical determinants of land values, the reasoning of Anselin and Lozano-Gracia (2009) of choosing a relatively simple form is followed. The Semi-Log form will be used for the main explanatory variables for which we assume a non-linear relationship with the land price (e.g.: size, time to CBD, density).

D.2.2 Endogeneity

In the hedonic pricing context, endogeneity is a major issue that occurs for instance if the consumer chooses the quantity and the price of an attribute (i.e parcel size) simultaneously. Endogenous variables are correlated with the error term and a basic OLS assumptions is hence not met[11] as a precondition to obtain unbiased estimation results is that all explanatory variables are independent of the error terms. The Hausman test is generally put forward to identify endogenous variables (Cavailhès, 2005) and instrumental variable regression is suggested to handle estimation errors related to endogenous variables (Irwin and Bockstael, 2001).

Open-space variables tend to be endogenous in hedonic price estimations, Cho and Roberts (2008) provide an overview of recent literature dealing with this source of endogeneity (Irwin and Bockstael, 2001; Smith et al., 2002; Geoghegan et al., 2003). Irwin and Bockstael (2001) investigate on the estimation and identification problems that might occur in estimating the value of open-space. They distinguish between privately held open-space, either developable and thus part of the residential land market or protected from development, and public open-spaces that are not developable. An endogeneity problem is identified with regard to the developable private open-space, as the implicit value of open-space amenities might be partly determined by its residential value (Irwin and Bockstael, 2001). Geoghegan et al. (2003) followed Irwin and Bockstael (2001) us-

[11] In linear regression it is assumed that the error terms are "*contemporaneously uncorrelated with the explanatory variables*" (Verbeek, 2008, p.129).

ing instrumental variables to address the potential endogeneity of the land-use measures. They chose variables correlated with these open-space variables but uncorrelated with the error term in equation as instruments. These variables are hypothesized to affect the relative costs and benefits of converting the land or maintaining it in an undeveloped use, not to be factors explaining housing prices, and quasi uncorrelated with the other explanatory. In addition to the soil attribute and slope variables proposed by Irwin and Bockstael (2001).

To control for endogeneity issues by instrumental variable estimation, the selection of appropriate instruments is a major challenge. It is necessary to find instruments for the characteristics price variables and the adjusted income variable to obtain consistent estimates. All of the characteristics' price variables and their prices may be correlated with the error terms, so the obvious choices for forming instruments are income and the socio-economic variables (Cheshire and Sheppard, 1995). However our datasets do not provide straight forward instruments such as households' income, generally put forward to control for the endogeneity related to the simultaneous choice of price and parcel size. Further, due to the aggregated scale of the data, similar instruments to those used by Irwin and Bockstael (2001) and Geoghegan et al. (2003) were not available.

D.3 Data aggregation and MAUP

Empirical studies are often confronted with problems related to the configuration of spatial units and different scales of aggregation: the *Modifiable Areal Unit Problem* (MAUP). The MAUP can be defined as the problem related to the sensitivity of analytical results to the definition of units for which data is collected (Fotheringham and Wong, 1991). Openshaw (1984) presented an overview on the MAUP and its implications for geographic research, that from a geographical perspective cannot be assumed away. The MAUP is twofold, on the one hand results may vary according to the configuration of the spatial units (zoning effect) and on the other hand to the level of aggregation of the data (scale effect). In literature, some attempts to identify problems related to differences in the level of aggregation between the dependent and independent variables, have been addressed among others by Goodman (1977); Shultz and King (2001). From an econometric standpoint, the MAUP can be viewed as an identification problem (Anselin, 1988b); there is insufficient information in the data to allow for the full specification of the simultaneous interaction over space. According to Anselin (1988b) the statistical measures for cross-sectional data are sensitive to the way in which the spatial units are organised; the level of aggregation affects the magnitude of several measures of association, such as spatial autocorrelation coefficients and parameters in a regression model. Fotheringham and Wong (1991) draw gloomy conclusions on the sensitivity of multivariate statistical analysis results to variations in scale and zoning systems. Since whatever aggregation method is applied, a smoothing effect tends to decrease the variation of a variable with increasing aggregation (Fotheringham and Wong, 1991). This loss of information yields to a change in the values for the various univariate, bivariate, and multivariate parameters (Reynolds and Amrhein, 1998).

D.4 Testing for spatial effects

The procedure for spatial dependence identification, essentially based on the works of Anselin (1988b); Le Sage (1999); Anselin (2002); Le Gallo (2002); Elhorst (2010). Elhorst (2010) draws an overview and decisional process scheme to proceed to the selection of the appropriate spatial regression model, they advocate the specific-to-general approach, starting with the non-spatial OLS estimation, test for different forms of spatial autocorrelation via Lagrange multiplier test (LM-test) and then proceed to the spatial models according to the test results. The LM-test was first proposed by Anselin (1988a), followed by the robust LM-test in Anselin et al. (1996) and is based on the residuals of the OLS model (similar to Moran's I[12]), following a chi-squared distribution with one degree of freedom (Elhorst, 2010). An advantage of the LM-tests[13] is that they do not require an alternative hypothesis, the OLS is the restricted model while the spatial models are the unrestricted alternatives (Conway et al., 2010).

[12] The most common test for spatial autocorrelation is the Moran's I test, for further details we refer to Anselin (1988b) and Anselin (1999).

[13] A more detailed review of the LM-test is provided among others by Le Gallo (2002).

As proposed by Anselin (2005), the LM-test and the robust LM-tests should guide the choice of the spatial model, depending on the significance of the LM-Error or LM-Lag result. These two refer to different spatial models, either the spatial error model, accounting for spatial autocorrelation among error terms, or the spatial autoregressive model, accounting for spatial dependence among dependent variables. If neither is significant, the OLS estimation is appropriate. In the case of one significant test result, the corresponding model is suggested to obtain unbiased estimation results. In the case where both are significant, the robust diagnostics should be considered in the same way. Elhorst (2010) extended this scheme to additional spatial models, either considering both, spatial lag and error (SAC[14]) or spatially autocorrelated explanatory variables (SDM[15]). McMillen (2003) claim that the rejection of the non-spatial model does not imply that the more elaborate, spatial model, is correct, as tests for spatial autocorrelation also detect misspecification of the functional form or heteroskedasticity. They warn that autocorrelation is often produced by model misspecification, focussing on issues related to the functional form. McMillen (2003) puts forward that further investigation is needed with regard to spatial autocorrelation induced by model misspecification.

In line with these claims, Anselin et al. (2004) highlight that tests for spatial effects, in particular Moran's I, assume a constant variance of the residuals and that results might be biased in case this assumption is not met. Further, they state that to date no theoretical results identify and prove the influence of heteroskedasticity on spatial dependence tests. Their findings suggest that the effect of heteroskedasticity on Moran's I and LM-tests depends on the spatial distribution of the heteroskedasticity, if it is itself spatially autocorrelated and if so, whether that correlation is positive or negative. However, their theoretical analysis suggests valid results for the tests in the case of spatially uncorrelated heteroskedastic residuals (Anselin et al., 2004). Further, Anselin

[14] Presented in chapter 3 section 3.2.

[15] The spatial Durbin model estimates a lag operator for the dependent variable as well as for all independent variables and was pointed out by Anselin (1988b) and further described in Anselin and Lozano-Gracia (2009). The spatial Durbin model estimates a lag operator for the dependent variable, as known from the SAR, as well as for all independent variables, thus considering the spatial relationship that might exist between independent variables (X) to the specification, via the spatial weight matrix (WX) (Viton, 2010).

and Lozano-Gracia (2009) highlight that a complexity of spatial econometric modelling is the difficulty to distinguish between spatial dependence and spatial heterogeneity in a cross-sectional setting. The specification tests and estimators developed for one spatial effect are affected by the presence of the other type.

Recently López et al. (2013) have developed an alternative specification test for spatial econometric models. The spatial Scan-test should, even in case of strong non-normal and heteroskedastic residuals, reliably detect spatial dependence. The Scan-test has power against a different alternative involving spatial instability of coefficients or spatial autocorrelation. An additional advantage of this test is that it does not require the definition of a spatial weight matrix and provides further information on clusters of high and low residuals. To date and in the framework of this thesis this test for spatial autocorrelation has not been considered, but we intend to consider it in future research.

Throughout this thesis, the procedure proposed by Elhorst (2010) is followed as it provides guidance to identifying the appropriate model accounting for issues related to spatial autocorrelation. There are, however, theoretical underpinnings to the different kinds of spatial autocorrelation in real estate data that might guide the spatial model choice beyond the different tests proposed here, in particular with regard to the above presented distrust in the traditional tests to identify spatial autocorrelation.

D.5 Further spatial models

The spatial econometric techniques used in the framework of this thesis have been presented in the empirical part. In this appendix we aim at providing some more detail on spatial estimation techniques not considered, mainly the spatial fixed effects and geographically weighted regression.

The spatial models presented in chapter 3 (SEM, SAR and SAC) are the most frequently considered spatial models in hedonic pricing literature. However as pointed out among others by Elhorst (2010), additional spatial model specifications using exogenous spatial weight matrices are available, an overview can be found for example in Wilhelmsson (2002); Elhorst (2010); Brady and Irwin (2011). Löchl and Axhausen (2010) draw a literature review on methods to consider spatial effects in hedonic pricing models (e.g.: spatial fixed effects, spatially adaptive filtering methods, expansion method, multi-level approach, spatial simultaneous autoregressive approach as well as geographically weighted regression models). It goes beyond the scope of this thesis to detail them all.

An other method often put forward to address the misspecification related to omitted spatial variables is the **spatial fixed-effects**[16] approach. It has been discussed and considered in several empirical implementations of the hedonic pricing method (Wilhelmsson, 2002; Cavailhès et al., 2009; Cavailhes and Thomas, 2012). This method should allow to control for spatial autocorrelation of the error term due to neighbourhood effects by introducing binary variables for the spatial units the transactions belong to. This underlies the assumption that the spatial extent of the unobserved heterogeneity or dependence fits this spatial extent of the units. It is however questionable, with regard to the complex nature of these neighbourhood effects, if such spatial effects manage to account for all residual autocorrelation (Anselin and Lozano-Gracia, 2009). Anselin and Arribas-Bel (2013) discuss more in detail and demonstrate analytically and by a series of simulations that spatial fixed-effects only correctly control spatial correlation if the omitted neighbourhood effect corresponds exactly to the extent of the considered spatial unit. They conclude that the spatial fixed effects only address a form of spatial heterogeneity, but do not necessarily manage to account for "true" spatial dependence (Anselin and Arribas-Bel, 2013, p.4).

Geographically weighted regression (GWR), incorporates geographical information using several distance-related weights (Löchl and Axhausen, 2010) and is considered to account for spatial heterogeneity in the valuation of he different price attributes. The GWR method explicitly allows spatially varying parameter estimates leading to independent spatial error terms. Rather than a single model, GWR estimates a separate model for each data point and weights observations by their distance to this point, thus allowing unique marginal price estimates at each location (Bitter et al., 2007). Löchl and Axhausen (2010) compare different spatial models (SAR, SEM and SARmix) to GWR

[16]Not to be confused with the fixed effects presented in the multilevel modelling framework.

and discuss the problem of real estate data unavailability in Switzerland[17]. The problem with the GWR model is that it does not solve the spatial dependence issues since the GWR residuals remain spatially autocorrelated and coefficients are correlated. Further, Löchl and Axhausen (2010) state that locally correlated GWR coefficient estimates remain a problem and that the GWR approach is helpful in situations where location related information or knowledge of local sub-markets are not available. Krause and Bitter (2012); Nilsson (2014) provide more detailed literature reviews on recent applications of the GWR method in the hedonic pricing context. In general, local regression models as GWR, show improvements over OLS models, however critics question the method in particular with regard to multicollinearity issues with the local parameters (Krause and Bitter, 2012).

[17] Combination of announced prices and stated preferences.

Appendix E

Notes to part II: Spatial relationships

The fundamental problem in analysing spatial correlation arises due to the "*lack of identification of the parameters of the complete covariance matrix*" (Anselin, 2002, p.256). Imposing a structure on this variance-covariance matrix by including a spatial weight matrix (SWM) to the hedonic model specification is necessary to account for the spatial ordering of the covariation between observations at different locations.

E.1 Finding the appropriate SWM...

Spatial econometricians generally turn to a spatial weights matrix (W) of dimension n x n[1], exogenously determined (De Graaff et al., 2001, p.260). A binary contiguity matrix as defined by Anselin (2002, p.257) specifies a neighbourhood set for each observation. For each row i a positive weight (w_{ij}) identifies j as a neighbour of i. Observations are not neighbours to themselves ($w_{ii} = 0$) and the diagonal elements are thus 0. Typically the spatial weight matrix is row-standardised (where each element is between 0 and 1), which allows more comparable estimation results and eases interpretation (Anselin, 2002). The specification of the appropriate weight matrix has been largely discussed in spatial econometric literature and according to Anselin (2002) there is little formal guidance for finding the most appropriate spatial weight matrix. Despite advances made in the last decades, this key issue remains largely unsolved (Abelairas-Etxebarria and Inma, 2012). Recently, some studies hedonic studies explicitly focussing on finding the appropriate spatial weight matrices have been conducted (Furtado and Van Oort, 2010; Brady and Irwin, 2011; Abelairas-Etxebarria and Inma, 2012; Seya et al., 2013).

Different ways of specifying the spatial relationships via spatial weight matrices have been put forward, mainly based on geographic weights (contiguity (Anselin, 2002)., dis-

[1] 6367 x 6367 in our case.

tance (Abelairas-Etxebarria and Inma, 2012)) or socio-economic weights (neighbourhood (Furtado and Van Oort, 2010; Furtado, 2011b)). Some spatial hedonic models rely on k-nearest neighbour matrices (e.g.: Kadish and Netusil, 2012; Christafore and Leguizamon, 2012), however such a specification would have required the exact transaction locations and could thus not be considered. Further some criticism has been raised among others by Anselin (2002), claiming that it lacks theoretical and empirical intuition and that the use of this kind of spatial weight matrix does not provide consistent results with regard to maximum likelihood estimation (Furtado and Van Oort, 2010). Abelairas-Etxebarria and Inma (2012) use inverse square distance matrix, which assigns more weight to closer plots, decreasing with distance. They found however, that neither the model nor the conclusions changed when using other types of weight matrices. Seya et al. (2013) present an automatic model selection algorithm for spatial econometric models using the trans-dimensional simulated annealing algorithm.

Some shortcomings of the use of spatial weight matrices in econometric modelling have been highlighted by Furtado and Van Oort (2010), claiming that contiguity or distance matrices are probably too abstract, as they rely on administrative boundaries that are not natural. Further, in the case of inverse distance matrices, the threshold is chosen by researchers and thus an arbitrary choice. While some claim a more theoretical basis for the definition of W (Harris et al., 2011; Corrado and Fingleton, 2012) others highlight that finding the ideal SWM is critical but that if substantial changes in estimation results are observed by changing the spatial weight matrix, this would rather be a sign of model misspecification (Le Sage and Pace, 2010; Krause and Bitter, 2012; Le Sage, 2014).

E.2 ... for observations at aggregated level

In the framework of this thesis, different spatial weight matrices have been tested to capture potential spill-over effects between developable land prices. A shortcoming of the dataset being the aggregated location of the observations at section scale, the distance weights between individual transactions could not be considered. Hence very local spill-over effects could thus not be controlled for as transactions in a same section are all considered neighbours. Further, some difficulties were encountered with isolated sections (fig.E.1), since not all of the 521 sections register transactions (actually only 321 register more than three observations). With regard to these limitations, we specified a contiguity matrix as well as inverse distance matrices (decreasing weight with sections centroid to neighbouring sections), presented in table E.1.

With regard to these dataset specificities some transformations were necessary to generate the matrices. We eventually considered a contiguity matrix, considering contiguity of the corners and edges of polygons, where in addition to the transactions located in neighbouring sections, observations located within a same section were considered as neighbours, except to themselves. In the isolated sections only transactions within the same section were considered as neighbours. To add more weight to nearby transactions located in the same section to those located in neighbouring sections we considered the inverse distance matrices. These spatial weights, thus the strength of the spatial relationship, decrease with distance. We considered two cut-off distances, 3,838 metres and 5,000metres. The first being the one suggested by ArcGIS10 (2010), to allow a minimum of one neighbouring section for every section above the critical threshold of observations. The second being chosen to test whether spatial autocorrelation might be measured at a larger extend, approximating the municipal level.

	Contx6367	ID38386367	ID50006367
Relationship	Contiguity	Inverse distance	Inverse distance
Distance cut-off	-	3,838m	5,000m
Number of regions	6,367	6,367	6,367
Number of non-zero links:	909,736	958,986	1,493,172
Percentage non-zero weights:	2.24	2.37	3.68
Average number of links:	142.88	150.62	234.52

Table E.1: Spatial weight matrices

No substantial differences in the overall information between the contiguity and the inverse distance matrix with 3,838m cut-off are observed from table E.1. The percentage of non-zero weights is around 2.3% while on average a transaction registers between 143

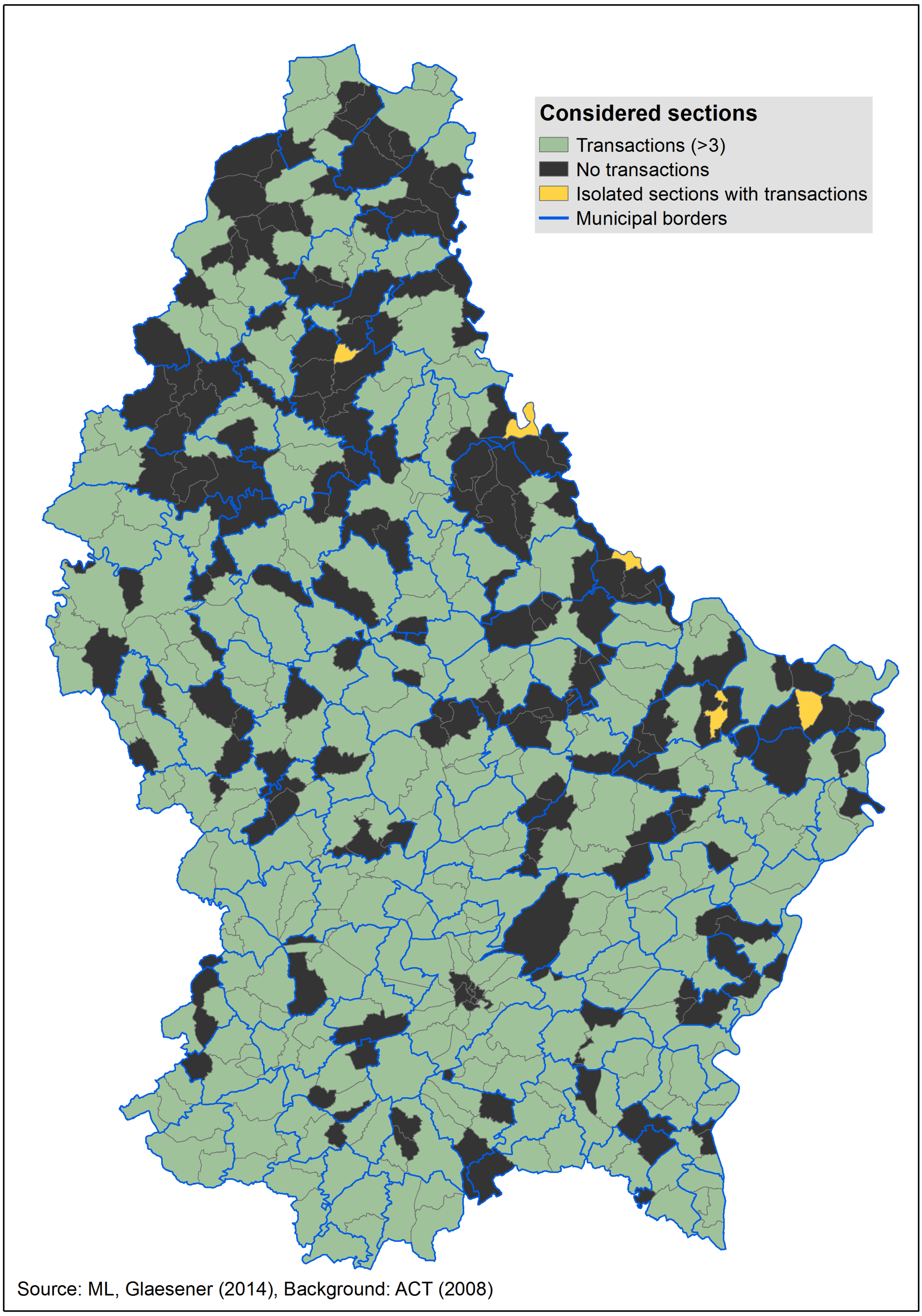


Figure E.1: Sections with transactions

and 151 neighbouring transactions. As expected with the 5km distance cut-off a larger number of transactions are considered neighbours, hence 3.7% of the 40,538,689 links are non-zero, with on average 235 neighbours per observation. These different spatial weight matrices have been used to test for spatial autocorrelation issues in our model and conclusions on the appropriate W in this case study have been presented in section 3.4.

Appendix F

Notes to chapter 4: QR results

	Global Model	Quantile regression D1	D2	D3	D4	D5	D6	D7	D8	D9
Intercept	12.280 ***	11.793 ***	12.073 ***	12.213 ***	12.331 ***	12.413 ***	12.439 ***	12.440 ***	12.734 ***	12.859 ***
Intercept	12.280***	11.810***	12.065***	12.164***	12.346***	12.410***	12.439***	12.375***	12.655***	12.916***
lnSize+	0.599***	0.728***	0.699***	0.668***	0.659***	0.640***	0.625***	0.603***	0.484***	0.369***
dVFA	0.291***	0.363***	0.254***	0.275***	0.308***	0.341***	0.362***	0.368***	0.212***	-0.078***
dTerrain	-0.397***	-0.201***	-0.217***	-0.279***	-0.311***	-0.346***	-0.370***	-0.412***	-0.482***	-0.616***
lntLUX^{+}_{ci}	-0.352***	-0.209***	-0.268***	-0.280***	-0.323***	-0.376***	-0.420***	-0.485***	-0.470***	-0.464***
lntLUX^{+}_{pi}	-0.084**	-0.275***	-0.255***	-0.234***	-0.188***	-0.128***	-0.076***	0.008	0.102**	0.091*
rAP	0.004**	0.012***	0.008***	0.005***	0.005***	0.005***	0.004***	0.002*	0.000	0.000
DIGreen	-0.168*	-0.027	-0.238**	-0.328***	-0.346***	-0.314***	-0.266***	-0.178**	-0.086	-0.164
mAPsh100	-0.071	-0.141	-0.119.	-0.103.	-0.176***	-0.138**	-0.091.	-0.016	-0.119	0.037
mAPsh1000	0.173***	0.243***	0.257***	0.249***	0.211***	0.163***	0.124***	0.129***	0.143**	0.107*
lnNetDens+	0.049***	0.040**	0.036.	0.046***	0.038***	0.037***	0.044***	0.055***	0.047*	0.036*
rBelow15y	-0.006**	-0.011***	-0.009***	-0.006***	-0.003*	-0.003**	-0.004***	-0.005**	-0.008*	-0.006 .
rHighIncome	0.026***	0.019*	0.018***	0.022***	0.021***	0.022***	0.025***	0.029***	0.044***	0.040***
txVarPop	0.000	0.000	0.001	0.000	0.000	-0.001***	-0.001***	-0.002***	-0.001	0.000
rUnemployment	0.003	-0.030***	-0.022***	-0.013*	-0.010**	-0.006	0.006	0.016**	0.022***	0.018 **
DI	0.002	-0.023	0.025	-0.011	-0.001	-0.017	-0.028.	-0.022	0.010	0.028
SSopp	0.001***	0.001**	0.000	0.000	0.001**	0.002***	0.003***	0.002***	0.001***	0.002 **

+ indicating grand-mean centred variables; Significance codes: 0 '***'; 0.001 '**'; 0.01 '*'; 0.05 '.'

Table F.1: OLS and QR estimation results

	S2SLS	Chernozhukov and Hansen IV Quantile Regression Results								
		D1	D2	D3	D4	D5	D6	D7	D8	D9
Intercept	9.231***	7.913***	9.749***	9.210***	10.012***	9.435***	9.952***	11.060***	9.574***	9.672***
lnSize$^+$	0.598***	0.730***	0.688***	0.669***	0.658***	0.645***	0.631***	0.601***	0.478***	0.354***
dVFA	0.291***	0.350***	0.250***	0.266***	0.314***	0.341***	0.367***	0.371***	0.210***	-0.090***
dTerrain	-0.394***	-0.154***	-0.219***	-0.253***	-0.316***	-0.341***	-0.372***	-0.413***	-0.472***	-0.621***
lntLUX_{ci}^+	-0.269***	-0.134 .	-0.206***	-0.214***	-0.255***	-0.298***	-0.364***	-0.448***	-0.405***	-0.361***
lntLUX_{pi}^+	-0.073*	-0.269***	-0.252***	-0.218***	-0.178***	-0.119***	-0.065**	0.020	0.122**	0.101*
DI	-0.001	-0.044	0.012	-0.020	-0.008	-0.028	-0.046*	-0.021	0.030	0.042
SSopp	0.001***	0.001*	0.000	0.001	0.001*	0.002***	0.003***	0.002***	0.001**	0.001*
rAP	0.004**	0.011***	0.007***	0.005***	0.005***	0.005***	0.004***	0.001	-0.001	0.001
DIGreen	-0.121	0.119	-0.223**	-0.290***	-0.318***	-0.297***	-0.215**	-0.138	-0.017	-0.192
mAPsh100	-0.074	0.002	-0.133 .	-0.115 .	-0.195***	-0.168**	-0.112 .	-0.051	-0.082	0.009
mAPsh1000	0.147***	0.199*	0.248***	0.226***	0.187***	0.155***	0.112***	0.099*	0.108 .	0.052
lnNetDens$^+$	0.045***	0.039 .	0.032	0.032*	0.031**	0.026**	0.039***	0.058***	0.045 .	0.030 .
rBelow15y	-0.005*	-0.012*	-0.007**	-0.005*	-0.002	-0.001	-0.002	-0.004 .	-0.005	-0.005
rHighIncome	0.018***	0.011	0.010	0.014**	0.017***	0.015***	0.019***	0.028***	0.038***	0.031***
txVarPop	0.000	0.001	0.000	0.000	0.000	-0.001**	-0.002***	-0.002**	-0.001	0.000
rUnemployment	0.007	-0.021 .	-0.021***	-0.009 .	-0.007	-0.002	0.008 .	0.018**	0.028**	0.025*
ρ	0.246**	0.300*	0.190*	0.240***	0.190**	0.240***	0.200*	0.110	0.240*	0.270*

$^+$ indicating grand-mean centred variables; Significance codes: 0 '***'; 0.001 '**'; 0.01 '*'; 0.05 '.'

Table F.2: IV Spatial AR and SIVQR results

Appendix G

Abbreviations and symbols

G.1 Abbreviations

AED	Administration de l'Enregistrement et des Domaines Administration of deeds
AGB	Farmland
AGG	Pastures, green agricultural land-use
AP	Available parcels
AP	Available plots
BA	other built area
BR	Built residential
BW	Brushwood
CBD	Central Business District
CEPS/INSTEAD	Centre for Population, Poverty and Public Policy Studies / International Networks for Studies in Technology, Environment, Alternatives, Development
CES	Conseil économique et social
CNPD	Data privacy commission
CRMM	Cross-regressive Multilevel Model
CTI	Centre des Technologies de l'Information de l'État
FOR	Forest
FRM	Fully random model
GD	Garden
GIS	Geographical Information System
GWR	Geographically Weighted Regression
High	Highway
HPM	Hedonic pricing method
ICC	Intraclass correlation coefficients
IGSS	Inspection Générale de la sécurité sociale
IND	Industrial land-use
IVQR	Instrumental Variables Quantile Regression model
MAUP	Modifiable areal unit problem
ML	Maximum likelihood estimation
MLM	Multilevel model
NDVI	Normalised difference vegetation index
NH	Neighbourhood
OLS	Ordinary least squares regression
ODS	Observatoire du développement spatial
PLUREL	Periurban Land Use Relationships project (Nilsson et al., 2013)
QR	Quantile Regression

G.1. Abbreviations

RAIL	Railway
RED	Real estate developer
RESO	Transport network
RIM	Random intercept model
RIV	River
RUG	Regional urban growth model (Nilsson et al., 2013)
S2SLS	Spatial two stages least squares estimation
SAC	Spatial simultaneous autoregressive model
SAR	Spatial Autoregressive Model/ Spatial Lag Model
SEM	Spatial Error Model
S.I.	Shannon index
SSM	Social minimum wage
STATEC	Institut national de la statistique et des études économiques du Grand-Duché du Luxembourg
SYVICOL	Syndicat des Villes et Communes luxembourgeoises
UM	Unconditional model
URG	Urban regional growth model
VY	Vineyards
WS	Watershed

G.2 List of symbols

α_i	Taste parameter
A_i	Accessibility measures to Luxembourg-city
β_{0jk}	Level-two intercept
β_0	Intercept in simple single level model
β_1	Regression estimates in simple single level model
c	All other consumption goods
δ_{00k}	Level-three intercept
E_i	"Natural" neighbourhood attributes
ϵ	Error term
γ_{000}	Overall intercept
$\gamma_{100}x_{ijk}$	Random slope at describing the overall coefficient of the attribute
i	Land transaction, composite good; level-one in the MLM environment
j	Section level
k	Municipal level
λ	Spatial autoregressive coefficient
N_i	Socio-economic neighbourhood composition
P_i	Land price
R_{ijk}	Level-one error term
ρ	Spatial lag operator
$\rho_\tau\ WY$	Tau defined spatial lag parameter
ρWx_{ijk}	Spatially lagged explanatory variables
S_i	Structural characteristics of residential land
σ^2	Variance between transactions within sections within municipalities
τ	Observed quantile of the price distribution
τ_0^2	Variance of the mean section price between sections within municipalities
τ_1^2	Variance for the random section slope
U	Utility function
U_{0jk}	Section random effect
U_{1jk}	Section random slope
X_i	Sum of structural and location-specific explanatory
φ_0^2	Variance of the mean price between municipalities compared to the overall mean
φ_1^2	Variance of municipal slope
φ_{01}	Covariance of random slope and intercepts at municipal level
V_{00k}	Municipal random effect
V_{10k}	Municipal random slope
y	Consumers' income
$Y/\ Y_{ijk}$	Total transaction price

Appendix H

Administrative and geological regions

Luxembourg is administratively divided into 116 municipalities sub-divided into 521 sections and on average municipalities include 4.53 sections. The mean size of a municipality is 22.35 km^2, while the mean size of a section is of 4.98 km^2. Urbanised sections correspond to a village and entourage or a neighbourhood in densely urbanised sections. In figure H.1 the different administrative levels (cantons (A), municipalities (A & B), sections (B)) have been illustrated.

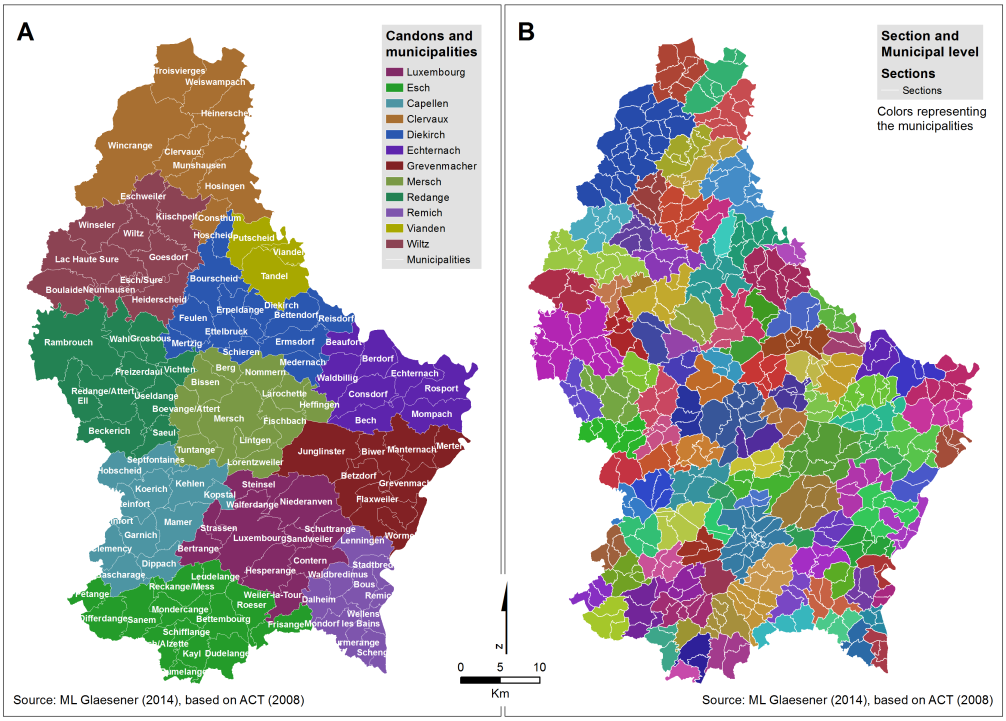
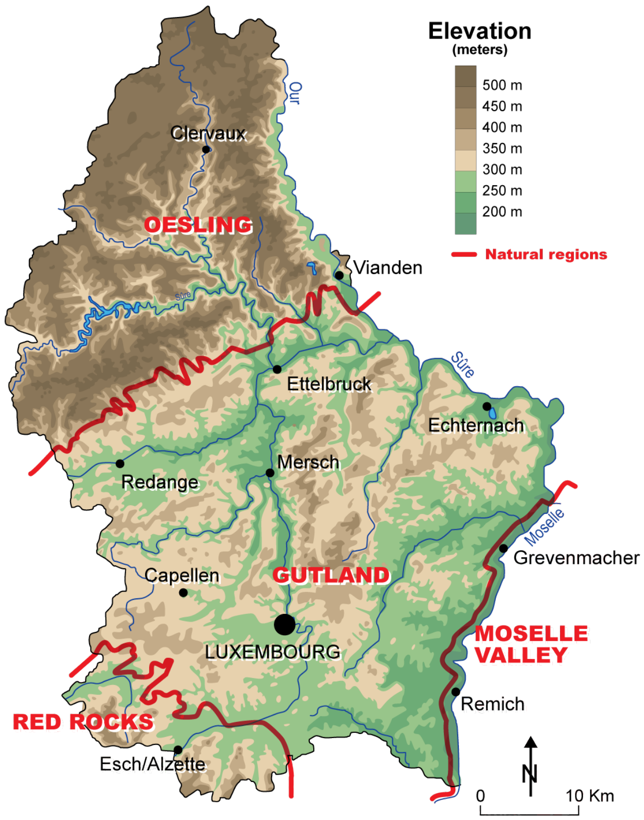


Figure H.1: Main administrative levels in Luxembourg

Throughout the thesis, mainly the municipal and section level will be referred to. While the municipal level is the second most important administrative level (below the national) we rely on the section scale to gain more detailed insights on the transactions location-specific context.

Map H.2 illustrates the geological borders of the four main natural regions. With the "Oesling" in the north, the southern part of the "Ardennes" characterised by hilly relief. The central and southern part, the "Gutland", is characterised by a flatter and more homogeneous geography, alternating between plateaus and planes. The south-eastern part of the country, the "Moselle Valley" where the picturesque landscapes of vineyards and agricultural activities prevail. In the south western part, the fourth main geological region, in the south-west, is characterised by its red rocks. The presence of iron ore deposits is at the origin of the past of mining and industrial activities in this region, also referred to as the former industrial south.



Source: Adapted by ML Glaesener (2014), based on MNHA LUXEMBOURG, http://www.prehistory.lu/lux/relief.htm (accessed 08.08.2014)

Figure H.2: Four main natural regions in Luxembourg

Acknowlegments

I would like to express my gratitude towards my supervisor Geoffrey Caruso, who competently guided me through this research in the past years, by improving and sharing new ideas while granting me the freedom to explore. Furthermore, I would like to thank my co-supervisors Dominique Peeters and Luisito Bertinelli for the constructive and encouraging discussions, suggestions and comments. I would like to thank Markus Hesse and Mohamed Hilal, for joining the jury and for taking interest in this thesis.

Thanks to different experts in the field of econometrics, especially Julie Le Gallo for her helpful comments on chapters 4 and 5. Philippe Kostov for suggestions on the implementation of the spatial quantile regression approach in an earlier stage of this thesis.

I would also like to acknowledge the support of the "Fonds National de la Recherche" in Luxembourg, funding this research since November 2009 within the AFR grant scheme (AFR grant number: PHD-09-095).

Big thanks to the whole team of the Institute for Geography and Spatial Planning, for the warm welcome back in 2009, for the opportunity and their trust during the Atlas project, giving me the chance to discover and contribute to an exceptional work environment. Special thanks to Tom Becker, for its numerous competent advice; Cyrille Medard de Chardon and Gilles Caspar, mainly for GIS related support and their help with different software tools, to improve the data management and geo-referencing processes; Mirjam Schindler for reading and commenting on parts of this thesis and for being a particular support, especially during the last year; Isabelle Pigeron-Piroth for sharing her experience and comforting words; to the former Masters' student Anissa supporting me in generating the public transport travel time data. Thanks to Carlos Marques and the SIU-team in Walferdange, who saved me and my data more than once.

Big thanks for the amazing support from besides my family, foremost to Yves, for his helpful comments and for taking care of our son in the last months of this thesis;

To Charles, who brightens-up every single day of my life and gave me the strength to go to the end. To my parents and my brothers, for days and nights of CEO-sitting and enduring me. Thanks also to my parents- and family-in-law who, despite very difficult times, were always available and full of comforting words. Special thanks to Anne, for improving my English and reducing grammar and spelling mistakes[1]. Thanks as well to all my friends from the "outside" for understanding and supporting me during all this time.

In more general, I would like to thank all former and actual residents of the Aquarium, for sharing their experiences and problems, in relation to their research or not, for the essential coffee and chocolate breaks and the numerous giggles, so helpful through all these years. I learned a lot from all of you. Eventually, very particular thanks to the *Mimies*: Anne, Annick, Conny, Lucie, and Sophie for being there, for reading early stage parts of this thesis and providing helpful comments on content and structure issues and most importantly on stress management.

[1] All remaining mistakes are under my responsibility

List of Figures

List of Tables

Bibliography

Abelairas-Etxebarria, P. and A. Inma (2012). Farmland prices and land-use changes in periurban protected natural areas. *Land Use Policy 29*(3), 674–683.

Acharya, G. and L. L. Bennett (2001). Valuing open space and land-use patterns in urban watersheds. *Journal of Real Estate Finance and Economics 22*(2-3), 221–237.

Administration du Cadastre et de la Topographie (2008a). Base de données topo-cartographique.

Administration du Cadastre et de la Topographie (2008b). Plan cadastral numérisé.

Affolderbach, J. and T. Becker (2011). Einzelhandel und Siedelungsentwicklung. In T. Chilla and C. Schulz (Eds.), *Raumordnung in Luxemburg / Aménagement du territoire au Luxembourg*, pp. 206–217. Binsfeld, Openscience.

Affolderbach, J. and C. Carr (2014). Blending scales of governance: Land use policies and practices in the small state of luxembourg. *Regional Studies online*, 12.

Ahlfeldt, G. M. (2011). If Alonso was right: Modeling accessibility and explaining the residential land gradiant. *Journal of regional science 51*(2), 318–338.

Aldenderfer, M. S. and R. K. Blashfield (1984). *Cluster analysis.* Number 07-044 in series on Quantitative Applications in the Social Sciences. Sage University Paper.

Allegrezza, S., F. Fehlen, D. Ferring, A. Heinz, F. Peltier, G. Thill, H. Willems, and P. Zahlen (2014). *La société luxembourgeoise dans le miroir du rencensement de la population.* Statec.

Alonso, W. (1964). *Location and land use.* Cambridge: Harvard University Press.

Anas, A., R. Arnott, and K. A. Small (1998, September). Urban spatial structure. *Journal of Economic Literature 36*(3), 1426–1464.

Anderson, S. T. and S. E. West (2006). Open space, residential property values, and spatial context. *Regional Science and Urban Economics 36*, 773–789.

Angelin, P. M. and R. Gencay (1996). Semiparametric estimation of a hedonic price function. *Journal of Applied Econometrics 11*, 633–648.

Anselin, L. (1988a). Lagrange multiplier test diagnostics for spatial dependence and spatial heterogeneity. *Geographical Analysis 20*(1), 1–17.

Anselin, L. (1988b). *Spatial Econometrics: Methods and Models.* Dordrecht: Kluwer Academic Publishers.

Anselin, L. (1999, April). Spatial econometrics. Bruton Center School of Social Sciences University of Taxas at Dallas.

Anselin, L. (2002). Under the hood: Issues in the specification and interpretation of spatial regression models. *Agricultural Economics 27*(3), 247–267.

Anselin, L. (2003). Spatial externalities, spatial multipliers, and spatial econometrics. *International Regional Science Review 26*(2), 153–166.

Anselin, L. (2005, March). Exploring spatial data with GeoDa: A workbook. online. http://www.csiss.org/clearinghouse/GeoDa/geodaworkbook.pdf.

Anselin, L. (2010). Thirty years of spatial econometrics. *Papers in Regional Science 89*(1), 3–25.

Anselin, L. and D. Arribas-Bel (2013). Spatial fixed effects and spatial dependence in a cross-section. *Papers in Regional Science 92*(1), 3–18. Wo.

Anselin, L., A. K. Bera, R. Florax, and M. J. Yoon (1996, February). Simple diagnostic tests for spatial dependence. *Regional Science and Urban Economics 26*(1), 77–104.

Anselin, L., Florax, and Rey (2004). *Advances in Spatial Econometrics*. Springer.

Anselin, L. and N. Lozano-Gracia (2009). Spatial hedonic models. In *Palgrave Handbook of Econometrics*, Volume 2, Chapter 26, pp. 1213–1250. Mills, Terence C. and Patterson, Kerry.

ArcGIS10 (2010). Conceptualitstation of spatial relationships. Technical report, ESRI.

AUS-EMAPS.com (2010). Address validation tool. Online tool for geocoding addresses. http://aus-emaps.com/bulk_geocoder.php, last accessed 20.03.2014.

Bala, A. P., D. Peeters, and I. Thomas (2011). Spatial issues on a hedonic estimation of rents in brussels. Technical report, Université Catholique de Louvain - Sustaincity.

Baranzini, A., J. Ramirez, C. Schaerer, and T. Philippe (2008). *Hedonic Methods in Housing Markets. Pricing Environmental Amenities and Segregation*. Springer.

Bartik, T. J. (1987). The estimation of demand parameters in hedonic price models. *Journal of Political Economy 95*, 81–88.

Bates, D., M. Maechler, B. Bolker, and S. Walker (2013). *lme4: Linear mixed-effects models using Eigen and S4*. The Comprehensive R Archive Network. R package version 1.0-5.

Becker, T. and M. Hesse (2011). Siedlungsentwicklung und Wohnungswesen. In T. Chilla and C. Schulz (Eds.), *Raumordnung in Luxemburg / Aménagement du territoire au Luxembourg*, pp. 79–95. Binsfeld, Openscience.

Beguin, H. (1979). *Méthodes d'analyse géographique quantitative*. Paris: Librairies Techniques.

Berger, F. (2004). Le patrimoine immobilier des ménages: qui possède quoi? *Population & Emploi 4*, 8.

Beron, K., J. C. Murdoch, and M. A. Thayer (1999). Hierarchical linear models with application to air pollution in the south coast air basin. *American Agricultural Economics Association 81*, 1123–1127.

Bhatta, B. (2010). *Analysis of urban growth and sprawl from remote sensing data*. Advances in geographic information science. Springer.

Bitter, C., G. F. Mulligan, and D. Sandy (2007). Incorporating spatial variation in housing attribute prices: A comparison of geographically weighted regression and the spatial expansion method. *Journal of Geographical Systems 9*, 7–27.

Bivand, R. (2010, October). Comparing estimation methods for spatial econometrics techniques using r. Discussion paper.

Bivand, R. (2013). *spdep: Spatial dependence: Weighting schemes, statistics and models.* The Comprehensive R Archive Network. R package version 0.5-68.

Blot, C. (2006, Mai). Peut-on parler de bulle sur le marché immobilier au Luxembourg. *Banque Centrale du Luxembourg - Working Paper 20*, 32.

Bockstael, N. and E. Irwin (2000). Economics and the land use environment link. In N. Edward Elgar Publishing (Ed.), *The International Yearbook of Environmental and Resource Economics 1999/2000*, pp. 1–54. Folmer, H. and Tietenberg, T.

Bolitzer, B. and N. Netusil (2000). The impact of open spaces on property values in Portland, Oregon. *Journal of Environmental Management 59*, 185–193.

Bousch, P. and A. Decoville (2012). Dynamiques de développement démographique et enjeux en matière d'aménagement du territoire. *Observatoire du Développement Spatial. Dossier thématique 1*, 62.

Bowen, W. M. (2001). Theoretical and empirical considerations regarding space in hedonic housing price model applications. *Growth and Change 32*(4), 466–490.

Brady, M. and E. Irwin (2011). Accounting for spatial effects in economic models of land use: Recent developments and challenges ahead. *Environmental & Resource Economics 48*, 487–509.

Brander, L. M. and M. J. Koetse (2011). The value of urban open space: Meta-analyses of contingent valuation and hedonic pricing results. *Journal of Environmental Management 92*, 2763–2773.

Brossard, T., J. Cavailhès, M. Hilal, D. Joly, F.-P. Tourneux, and P. Wavresky (2008). La valeur des paysages périurbains dans un marché immobilier en france. In F. Des Rosiers and M. Thériault (Eds.), *Information géographique et dynamiques urbaines 2*, Chapter 9, pp. 225–248. Lavoisier.

Brown, J. and H. Rosen (1982). On the estimation of structural hedonic price models. *Econometrica 50*, 765–768.

Brueckner, J. K. (1983). The economics of urban yard space: An "implicit-market" model for housing attributes. *Journal of Urban Economics 13*, 216–234.

Brueckner, J. K. (2000, April). Urban sprawl: Diagnosis and remedies. *International Regional Science Review 23*(2), 160–171.

Brueckner, J. K. (2001). Urban sprawl: Lessons from urban economics. *Brookings-Wharton Papers on Urban Affairs 1*, 65–96.

Brueckner, J. K. (2007). Urban growth boundaries: An effective second-best remedy for unpriced traffic congestion? *Journal of Housing Economics 16*, 263–273.

Brueckner, J. K. (2011). *Lectures on urban economics.* MIT Press.

Brueckner, J. K. and D. A. Fansler (1983). The economics of urban sprawl: Theory and evidence on the spatial sizes of cities. *The Review of Economics and Statistics 65*(3), 479–482.

Brueckner, J. K. and R. W. Helsley (2011). Sprawl and blight. *Journal of Urban Economics 69*, 205–213.

Brueckner, J. K. and A. G. Largey (2008). Social interaction and urban sprawl. *Journal of Urban Economics 64*, 18–34.

Brueckner, J. K., J.-F. Thisse, and Y. Zenou (1999). Why is central Paris rich and downtown Detroit poor? an amenity-based theory. *European Economic Review 43*, 91–107.

Brunauer, W. (2013). Modelling house prices using multilevel structured additive regression. *Statistical Modelling 13*(2), 95–123.

Buchinsky, M. (1997). Recent advances in quantile regression models, a practical guideline for empirical research. *The Journal of Human Resources 33*(1), 88–126.

Burchell, R. W., D. Listokin, and C. C. Galley (2000). Smarth growth: More than a ghost of urban policy past, less than a bold new horizon. *Housing Policy Debate 11*(4), 821–879.

Burchfield, M., H. Overman, D. Puga, and M. Turner (2006, May). Causes of sprawl: A portrait from space. *Quaterly Journal of Economics 121*(2), 587–633.

Can, A. (1990, July). The measurement of neighbourhood dynamics in urban house prices. *Economic Geography 66*(3), 254–272.

Carpentier, S. (2006). Structuration urbaine et typologie des communes luxembourgeoises pour l'étude de la mobilité quotidienne. *CEPS/INSTEAD Population et Territoire 9*, 20.

Carpentier, S. (2010). *La mobilité résidentielle transfrontalière entre le Luxembourg et ses régions voisines*. Forum Europa.

Carruthers, J., D. E. Clark, and R. Renner (2010). The benefits of environmental improvement: Estimates from space-time analysis. Working Papers and Research 2010-11, Marquette University, Center for Global and Economic Studies and Department of Economics.

Caruso, G. (2002). La diversité des formes de la periurbanisation en Europe. In P. Perrier-Cornet (Ed.), *Repenser les Campagnes*, Série bibliothèque des territoires, pp. 67–99. Editions de l'aube, Datar.

Caruso, G. (2005, Feburary). *Integrating urban economics and cellular automata to model periurbanisation, Spatial dynamics of residential choice in the presence of neighbourhood externalities*. Ph. D. thesis, Université catholique de Louvain, Faculté des Sciences.

Caruso, G. and J. Cavailhes (2010). Forme paysagères dans les modèles théoriques d'économie et de géographie urbaines. In W. Oueslati (Ed.), *Analyses économiques du paysage*, Chapter 3, pp. 63–85. Update Sciences & Technologies.

Caruso, G., D. Peeters, J. Cavailhès, and M. Rounsevell (2007). Spatial configurations in a periurban city: A cellular automata-based microeconomic model. *Regional Science and Urban Economics 37*, 542–567.

Caruso, G., G. Vuidel, J. Cavailhès, P. Frankhauser, D. Peeters, and I. Thomas (2011). Morphological similarities between dbm and a microeconomic model of sprawl. *Journal of Geographical Systems 13*(1), 31–48.

Cavailhes, J. and I. Thomas (2012). Are agricultural and developable land prices governed by the same spatial rules? The case of Belgium. *0*, 41.

Cavailhès, J. (2005). Le prix des attributs du logement. *Economie et Statistique 382*, 91–123.

Cavailhès, J., T. Brossard, J.-C. Foltête, M. Hilal, D. Joly, F.-P. Tourneux, C. Tritz, and P. Wavresky (2007). Les prix des paysages périurbains. *Economie rurale 297-298*, 71–84.

Cavailhès, J., T. Brossard, J.-C. Foltête, M. Hilal, D. Joly, F.-P. Tourneux, C. Tritz, and P. Wavresky (2009). GIS-based hedonic pricing of landscape. *Environemental and Resource Economics 44*, 571–590.

Cavailhès, J., T. Brossard, M. Hilal, D. Joly, F.-P. Tourneux, C. Tritz, and P. Wavresky (2006). *Les paysages périurbains et leurs prix.* Les Cahiers de la MSH Ledoux. Presses universitaires de Franche-Compté.

Cavailhès, J., P. Frankhauser, D. Peeters, and I. Thomas (2004). Where Alonso meets Sierpinski: An urban economic model of a fractal metropolitan area. *Environement and Planning A 36*, 1471–1498.

Cavailhès, J., D. Peeters, E. Sékeris, and J.-F. Thisse (2004). The periurban city: Why to live between the suburbs and the countryside. *Regional Science and Urban Economics 34*(6), 681–703.

CEPS/INSTEAD (2008). Suivi du développement territorial du Luxembourg à la lumière des objectifs de l'IVL. Technical report, CEPS/INSTEAD.

Chaix, B., J. Merlo, and P. Chauvin (2005). Comparison of a spatial approach with the multilevel approach for investigating place effects on health: the example of healthcare utilisation in France. *Journal of Epidemidology and Community Health 56*(6), 517–526.

Chasco, C. and J. Le Gallo (2012). Hierarchy and spatial autocorrelation effects in hedonic models. *Economics Bulletin 32*(2), 1474–1480.

Chasco, C. and J. Le Gallo (2013). The impact of objective and subjective measures of air quality and noise on house prices: A multilevel approach for downtown Madrid. *Economic Geography 89*(2), 127–148.

Chasco, C. and B. Sánchez Reyes (2012). Externalidades ambientales y precio de la vivienda en Madrid: un análisis con regresión cuantílica espacial. *Revista Galega de Economía 21*(2), 1–21.

Chen, V. Y.-J., W.-S. Deng, T.-C. Yang, and S. A. Matthews (2012). Geographically weighted quantile regression (GWQR): An application to U.S. mortality data. *Geographical Analysis 44*(2), 134–150.

Chernozhukov, V. and C. Hansen (2006). Instrumental quantile regression inference for structural and treatment effect models. *Journal of Econometrics 132*, 491–525.

Cheshire, P. and S. Sheppard (1995). On the price of land and the value of amenities. *Economica, New Series 62*(246), 247–267.

Cheshire, P. and S. Sheppard (1998). Estimating the demand for housing, land and neighbourhood characteristics. *Oxford Bulletin of Economics and Statistics 60*(3), 357–382.

Chilla, T. and C. Schulz (2011). *Raumordnung in Luxemburg / Aménagement du territoire au Luxembourg*. Openscience, Binsfeld.

Chilla, T. and C. Schulz (2012). L'aménagement du territoire au Luxembourg: défis, objectifs et instruments d'une politique émergente. *Revue Géographique de l'Est 52*, 18.

Chilla, T. and C. Schulz (2014). Spatial development in Luxembourg: mimetic evolution or emergence of a new planning culture? *European Planning Studies 22*, 20.

Cho, Seong-Hoon, P. N. C. and R. K. Roberts (2008). Spatial analysis of the amenity value of green open space. *Ecological Economics 66*, 403–416.

Christafore, D. and S. Leguizamon (2012). The influence of gay and lesbian coupled households on house prices in conservative and liberal neighborhoods. *Journal of Urban Economics 71*(2), 258–267.

Clapp, J. M., A. Nanda, and S. L. Ross (2008). Which school attributes matter? The influence of school district performance and demographic composition on property values. *Journal of Urban Economics 63*, 451–466.

Conseil Economique et Social (2013, octobre). L'accès à un logement abordable.

Conway, D., C. Q. Li, J. Wolch, C. Kahle, and M. Jerrett (2010). Spatial autocorrelation approach for examining the effects of urban greenspace on residential property values. *The Journal of Real Estate Finance and Economics 41*, 150–169.

Corrado, L. and B. Fingleton (2011). Multilevel modelling with spatial effects. Working Papers 1105, University of Strathclyde Business School, Department of Economics.

Corrado, L. and B. Fingleton (2012). Where is the economics in spatial econometrics? *Journal of Regional Science 52*(2), 210–239.

Day, B., I. Bateman, and I. Lake (2006). Estimting the demand for peace and quiet using property market data. *CSERGE Working Paper EDM 06-03*, 35.

De Bruyne, K. and J. Van Hove (2013). Explaining the spatial variation in housing prices: An economic geography approach. *Applied Economics 45*, 1673–1689.

De Graaff, T., R. Florax, and P. Nijkamp (2001). A general misspecisfiation test for spatial regression models: dependence, heterogeneity, and nonlinearity. *Journal of regional science 41*(2), 255–276.

Decoville, A. (2012). La centralité urbaine au Luxembourg : analyse et perspectives. Rapport de l'observatoire du développement spatial, CEPS/INSTEAD - Observatoire du Développement Spatial. last accessed 19.03.2014.

Des Rosiers, F., J. Dubé, and M. Thériault (2008). La modélisation hédonique des valeurs résidentielles: la mesure d'externalités urbaines au Québec. In F. Des Rosiers and M. Thériault (Eds.), *Information géographique et dynamiques urbaines 2*, Chapter 8, pp. 193–223. Paris: Lavoisier.

Des Rosiers, F. and M. Thériault (2006). Mass appraisal, hedonic price modeling and urban externalities: Understanding property value shaping processes. In *Advances in Mass Appraisal Methods Seminar*, pp. 40. Delft University of Technology.

Des Rosiers, F., M. Thériault, Y. Kestens, and P. Villeneuve (2002). Landscaping and house values: An empirical investigation. *Journal of Real Estate Research 23*(1/2), 139–161.

Des Rosiers, F., M. Thériault, Y. Kestens, and P. Villeneuve (2007). Landscaping attributes and property buyers' profile: Their joint effect on house prices. *Housing Studies 22*(6), 945–964.

Diederich, R. (2011a). Les documents fondamentaux de l'aménagement du territoire au Luxembourg: Le programme directeur et l'IVL. In T. Chilla and C. Schulz (Eds.), *Raumordnung in Luxemburg / Aménagement du territoire au Luxembourg*, pp. 59–76. Binsfeld, Openscience.

Diederich, R. (2011b). Objectifs de l'aménagement du territoire au Luxembourg. In T. Chilla and C. Schulz (Eds.), *Raumordnung in Luxemburg / Aménagement du territoire au Luxembourg*. Binsfeld, Openscience.

Diop, L. (2011a, October). Luxembourg: les marchés fonciers et immobiliers transfrontaliers à l'épreuve de la métropolisation. *CEPS Working Papers 48*, 11.

Diop, L. (2011b, March). Métropolisation et spécialisation sociale de l'espace: Quel effet de la frontière dans l'aire résidentielle transfrontalière de luxembourg. *CEPS Working Papers 5*, 17.

Djurdjevic, D., C. Eugster, and R. Haase (2008). Estimatinon of hedonic models using a multilevel approach: An application for the swiss rental market. *Swiss Journal of Economics and Statistics 144*(4), 679–701.

Dubé, J., F. Des Rosiers, and M. Thériault (2011). Segmentation spatiale et choix de la forme fonctionnelle en modélisation hédonique. *Revue d'Economie Régionale & Urbaine 1*, 9–37.

Duncan, C., K. Jones, and G. Moon (1998). Context, composition and heterogeneity: Using multilevel models in health research. *Social Science & Medicine 46*(1), 97–117.

Ebru, Çaglayan & Eban, A. (2009). Determinants of house prices in Istanbul: a quantile regression approach. *Quality & Quantity 45*(2), 305–317.

Echenique, M. H., A. J. Hargreaves, G. Mitchell, and A. Namedo (2012). Growing cities sustainably. *Journal of the American Planning Association 78*(2), 121–137.

Elhorst, J. P. (2010, March). Applied spatial econometrics: Raising the bar. *Spatial Econometric Analysis 5*(1), 9–28.

Elhorst, J. P. and A. S. Zeilstra (2007, November). Labour force participation rates at the regional and national levels of the European Union: An integrated analysis. *Papers in Regional Science 86*(4), 525–549.

Epple, D. (1987). Hedonic prices and implicit markets: Estimating demand and supply functions for differentiated products. *Journal of Political Economy 95*, 59–80.

Epstein, D. (2010). L'évolution de la vitesse de déplacement domicile-travail au Luxembourg de 1960 à nos jours. *CEPS Working Papers 17*, 19.

Espey, M. and K. Owusu-Edusei (2001). Neighborhood parks and residential property values in greenville, south carolina. *Journal of Agricultural and Applied Economics 33*(3), 487–492.

European Environment Agency (2010). Analysing and managing urban growth. http://www.eea.europa.eu/articles/analysing-and-managing-urban-growth, last accessed 16.05.2014.

EUROSTAT (2011, June). Population projections 2010-2060. online. News release.

Evans, A. W. (2004). *Economics & Land Use Planning*. Blackwell.

Ewing, R. H. (1994). Characteristics, causes, and effects of sprawl: A literature review. *Environmental and Urban Studies 21*(2), 1–15.

Ewing, R. H., R. Pendall, and D. Chen (2002). Measuring sprawl and its impact. Technical report, Smart Growth America.

Farmer, M. C. and C. A. Lipscomb (2010). Using quantile regression in hedonic analysis to reveal submarket competition. *Journal of Real Estate Research 32*(4), 435–460.

Fehlen, F., C. Burnotte-Boreux, and Y. Casteels (2003). *La société luxembourgeoise à travers le recensement de 2001*. STATEC.

Fingleton, B. and J. Le Gallo (2008). Estimating spatial models with endogenous variables, a spatial lag and spatially dependent disturbances: Finite sample properties. *Papers in Regional Science 87*(3), 319–339.

Finotto, F. (2011). Landscape assessment: The ecological profile. In C. Cassatella and A. Peano (Eds.), *Landscape indicators: Assessing and Monitoring Landscape Quality*, Chapter 4, pp. 47–75. Springer.

Fletcher, M., P. Gallimore, and J. Mangan (2000). Heteroscedasticity in hedonic house price models. *Journal of Property Research 17*(2), 93–108.

Fontes, G. G., R. F. Simões, A. M. De Oliveira, and C. Hermeto (2009). Urban attributes and wage disparities in brazil: A multilevel hierarchical model. *Regional Studies 44*(5), 596–607.

Fotheringham, A. S. and D. Wong (1991). The modifiable areal unit problem in multivariate statistical analysis. *Environment and Planning A 23*, 1025–1044.

Freeman, M. (1979). Hedonic prices, property values and measureing environmental benefits: A survey of the issues. *Scandinavian Journal of Economics 81*(2), 154–173.

Fujita, M. (1989). *Urban Economic Theory Land Use and City Size*. Cambridge University Press.

Fujita, M. and J. Thisse (2002). *Economics of Agglomeration: Cities, Industrial Location and Regional Growth*. Cambridge: Cambridge University Press.

Fujita, M. and J. Thisse (2013). *Economics of Agglomeration: Cities, Industrial Location and Regional Growth* (2nd ed.). Cambridge: Cambridge University Press.

Furtado, B. A. (2011a, Jan). Análise quantílica-espacial de determinantes de preços de imóveis urbanos com matriz de bairros: Evidências do mercado de belo horizonte. Discussion Papers 1570, Instituto de Pesquisa Econômica Aplicada - IPEA.

Furtado, B. A. (2011b). Neighbourhoods in urban economics: Incorporating cognitively perceived urban space in economic models. *Urban Studies 48*(13), 2827–2847.

Furtado, B. A. and F. Van Oort (2010, September). Neighborhood weight matrix in a spatial-quantile modelling environment: Evidence from Brazil. Conference papers, European Regional Science Association.

Fávero, L. P. L. (2011). Preços hedônicos no mercado imobiliário comercial de São Paulo: A abordagem da modelagem multínível com classificação cruzada. *Estudos Econômicos 41*(4), 777–810.

Galster, G., R. Henson, M. R. Ratcliffe, H. Wolman, S. Coleman, and J. Freihage (2001). Wrestling sprawl to the ground: Defining and measuring an elusive concept. *Housing Policy Debate 12*, 681–717.

Geniaux, G. and C. Napoléone (2008). Semi-parametric tools for spatial hedonic models: An introduction to mixed geographically weighted regression and geoadditive models. In A. Baranzini, J. Ramirez, C. Schaerer, and T. Philippe (Eds.), *Hedonic Methods in housing markets, Pricing Environmental amenities and segregation*, Chapter 5, pp. 101–127. Springer.

Geoghegan, J. (2002). The value of open spaces in residential land use. *Land Use Policy 19*, 91–98.

Geoghegan, J., L. Lynch, and S. Bucholtz (2003). Capitalisation of open spaces into housing values and the residential property tax revenue impacts of agricultural easement programs. *Agricultural and Resource Economics Review 32*(1), 33–45.

Geoghegan, J., L. Wainger, and N. Bockstael (1997). Spatial landscape indicies in a hedonic framework: An ecological economics analysis using GIS. *Ecological Economics 23*, 251–264.

Geraci, M. (2014). Linear quantile mixed models: The lqmm package for laplace quantile regression. *Journal of Statistical Software 0*(0), 0–0. In press.

Geraci, M. and M. Bottai (2013). Linear quantile mixed models. *Statistics and Computing published online*, 1–19.

Giuliano, G., P. Gorden, Q. Pan, and J. Park (2010). Accessiblity and residential land values : some tests with new measures. *Urban Studies 47*(14), 3103–3130.

Glaeser, E. L. (2008). *Cities, Agglomeration and Spatial Equilibrium.* Oxford University Press.

Glaeser, E. L. and M. E. Kahn (2004). Sprawl and urban growth. In J. V. Henderson and J. F. Thisse (Eds.), *Handbook of Regional and Urban Economics*, Volume 4 of *Handbook of Regional and Urban Economics*, Chapter 56, pp. 2481–2527. Elsevier.

Glaeser, E. L., J. Kolko, and A. Saiz (2001). Consumer city. *Journal of Economic Geography 1*(1), 27–50.

Goffette-Nagot, F., I. Reginster, and I. Thomas (2011). Spatial analysis of residential land prices in Belgium: Accessibility, linguistic border, and environmental amenities. *Regional Studies 45*(9), 1253–1268.

Goodman, A. C. (1977). A comparison of block group and census tract data in a hedonic housing price model. *Land Economics 53*(4), 483–487.

Goodman, A. C. and T. G. Thibodeau (1998). Housing market segmentation. *Journal of housing economics 7*, 121–143.

Gordon, P. and H. W. Richardson (2007). The debate on sprawl and compact cities: thoughts based on the congress of new urbanism charter. In H. S. Geyer (Ed.), *International Handbook of Urban Policy*. Edward Elgar Publishing.

Griliches, Z. (1961, December). Hedonic price indexes for automobiles: An econometric of quality change. In *The Price Statistics of the Federal Goverment*, NBER Chapters, pp. 173–196. National Bureau of Economic Research, Inc.

Hallin, M., Z. Lu, and K. Yu (2007). Local linear spatial quantile regresssion. Discussion paper 48, ECORE - International association for research and teaching Economics, Finance, Operations Research, Economatrics and Statstics.

Halstead, J. M., R. A. Bouvier, and B. E. Hansen (1997). On the issue of functional form choice in hedonic price functions: Further evidence. *Environmental Management 21*(5), 759–765.

Hansen, F. (2012). Caractéristiques du parc automobile de 1962 à 2012. *50 ANS STATEC 17*, 4.

Harris, R., J. Moffat, and V. Kravtsova (2011). In search of "W". *Spatial Economic Analysis 6*(3), 249–270.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica 47*(1), 153–161.

Heinz, A., G. Thill, and F. Peltier (2012a). La répartition territoriale de la population. In S. Allegrezza, D. Ferring, H. Willems, and P. Zahlen (Eds.), *RP 2011 - Premiers résultats*, Volume 3. STATEC.

Heinz, A., G. Thill, and F. Peltier (2012b). L'évolution de la population. In S. Allegrezza, D. Ferring, H. Willems, and P. Zahlen (Eds.), *RP 2011 - Premiers résultats*, Volume 2. STATEC.

Hesse, M. (2012). Suburbaner Raum. Annäherungen an Gegenstand, Inhalte und Bedeutungszuweisungen. In W. Schenk (Ed.), *Suburbane Räume als Kulturlandschaften*, pp. 13–24. ARL.

Hesse, M. (2014a). Distorted density: Where the views of developers and NGOs on sustainable urban development intersect. Working Paper.

Hesse, M. (2014b). On borrowed size, flawed urbanisation and emerging enclave spaces: The exceptional urbanism of Luxembourg, Luxembourg. *European Urban and Regional Studies 0*, 16.

Hill, P. W. and K. J. Rowe (1996). Multilevel modelling in school effectiveness research. *School Effectiveness and School Improvement 7*(1), 1–34.

Hosszú, S. (2009). Counterurbanization: A literature study. Working Paper No. 06/2009, Danish Institute of Rural Research and Development.

Hoyt, W. H. and S. S. Rosenthal (1997). Household location and Tiebout : Do families sort according to preferences for locational amenities? *Journal of Urban Economics 42*, 159–178.

Irwin, E. G. and N. Bockstael (2004). Land use externalities, open space preservation, and urban sprawl. *Regional Science and Urban Economics 34*, 705–725.

Irwin, E. G. and N. E. Bockstael (2001). The problem of identifying land use spillovers: Measuring the effects of open space on residential property values. *American Journal of Agricultural Economics 83*, 698–704.

Irwin, E. G. and N. E. Bockstael (2007, December). The evolution of urban sprawl: Evidence of spatial heterogeneity and increasing land fragmentation. *Proceedings of the National Academy of Sciences of the United States of America 104*, 20672–20677.

Irwin, E. G. and J. Geoghegan (2001). Theory, data, methods: developing spatially explicit economic models of land use change. *Agriculture, Ecosystems and Environment 85*, 7–23.

Jenness, J., B. Brost, and P. Beier (2013). Land facet corridor designer: Extension for ArcGIS. Technical report, Jenness Enterprises.

Jim, C. and W. Y. Chen (2009). Value of scenic views: Hedonic assessment of private housing in hong kong. *Landscape and Urban Planning 91*, 226–234.

Jones, K. (1991). *Multi-level models for geographical research* (CATMOG ed.), Volume 54. Geo Books.

Jones, K. and N. Bullen (1993). A multi-level analysis of the variations in domestic property prices.. *Urban Studies 30*(8), 1409–1426.

Jones, K. and N. Bullen (1994). Contextual models of urban house prices: A comparison of fixed- and random-coefficient models developed by expansion. *Economic Geography 70*(3), 252–272.

Kadish, J. and N. R. Netusil (2012). Valuing vegetation in an urban watershed. *Landscape and Urban Planning 104*, 59–65.

Kelejian, H. and I. Prucha (1998). A generalized spatial two stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. *Journal of Real Estate Finance and Economics 17*, 99–121.

Kestens, Y., M. Thériault, and F. Des Rosiers (2004). The impact of surrounding land use and vegetation on single-family house prices. *Environment and Planning B 31*, 539–567.

Kestens, Y., M. Thériault, and F. Des Rosiers (2006). Heterogeneity in hedonic modelling of house prices: looking at buyers' household profiles. *Journal of Geographical Systems 8*(1), 61–96.

Kiel, K. A. and J. E. Zabel (2008, June). Location, location, location: The 3L Approach to house price determination. *Journal of Housing Economics 17*(2), 175–190.

Kim, C. W., T. T. Phipps, and L. Anselin (2003). Measuring the benefits of air quality improvment: a spatial hedonic approach. *Journal of Environmental Economics and Management 45*, 24–39.

Kim, T.-H. and C. Muller (2004). Two-stage quantile regression when the first stage is based on quantile regression. *Econometrics Journal 7*, 218–231.

Klein, S. (2010). Les transports en commun au luxembourg en 2006-2007: Caractérisation du réseau et de l'offre publique de transports en commun. *Les cahiers du CEPS/INSTEAD GEODE 11*(11), 19.

Koenker, R. (2005). *Quantile Regression.* Econometric society monographs.

Koenker, R. and G. J. Bassett (1978). Regression quantiles. *Econometrica 46*, 33–50. @quantile regression.

Koenker, R. and K. F. Hallock (2001). Quantile regression. *Journal of Economic Perspectives 15*(4), 143–156.

Kohnen, M. (2011). Nachhaltige Planung und der Nationale Nachhalitgkeitsplan. In T. Chilla and C. Schulz (Eds.), *Raumordnung in Luxemburg / Aménagement du territoire au Luxembourg*, pp. 318–326. Openscience, Guy Binsfeld.

Kong, F., H. Yin, and N. Nakagoshi (2007). Using GIS and landscape metrics in the hedonic price modeling of amenity value of urban green space: A case study in Jinan-city, China. *Landscape and Urban Planning 79*, 240–252.

Koschinsky, J. and N. Lozano-Gracia (2012). The welfare benefit of a home's location: an empirical comparison of spatial and non-spatial model estimates. *Journal of Geographical Systems 14*, 319–356.

Kostov, P. (2009, March). A spatial quantile regression hedonic model of agricultural land prices. *Spatial Economic Analysis 4*(1), 53–72.

Kostov, P. (2013a). Choosing the right spatial weighting matrix in a quantile regression model. *ISRN Economics 2013*, 16.

Kostov, P. (2013b). Empirical likelihood estimation of the spatial quantile regression. *Journal of Geographical Systems 15*, 51–69.

Krause, A. L. and C. Bitter (2012). Spatial econometrics, land values and sustainability: Trends in real estate valuation research. *Cities 29*, 19–25.

Krugman, P. (1991). Increasing returns and economic geography. *Journal of Political Economy 99*(3), 483–499.

Kuminoff, N. V., C. F. Parmeter, and J. C. Pope (2010). Which hedonic models can we trust to recover the marginal willingness to pay for environmental amenities? *Journal of EnvironmentalEconomicsandManagement 60*, 145–160.

Kuznetsova, A., P. Bruun Brockhoff, and R. Haubo Bojesen Christensen (2013). *lmerTest: Tests for random and fixed effects for linear mixed effect models (lmer objects of lme4 package).* Technical University of Denmark, Department of Applied Mathematics and Computer Science. R package version 2.0-3.

Lancaster, K. (1966). A new approach to consumer theory. *The Journal of Political Economy 74*(2), 132–157.

Löchl, M. and K. W. Axhausen (2010). Modeling hedonic residential rents for land use and transport simulation while considering spatial effects. *The Journal of Transport and Land use 3*(2), 39–63.

Le Gallo, J. (2002). Économétrie spatiale: l'autocorrélation spatiale dans les modèles de régression linéaire. *Économie et Prévision 4*(155), 139–157.

Le Gallo, J. (2004). Hétérogénéité spatiale. *Économie et Prévision 1*(162), 151–172.

Le Gléau, J.-P., D. Pumain, and T. Saint-Julien (1996). Villes d'europe: à chaque pays sa définition. *Économie et Statistique 294-295*, 9–23.

Le Sage, J. P. (1999). *The Theory and Practice of Spatial Econometrics.* University of Toledo.

Le Sage, J. P. (2014). What regional scientists need to know about spatial econometrics. Working paper, Texas State University-San Marcos.

Le Sage, J. P. and K. R. Pace (2010). The biggest myth in spatial econometrics. Technical report, Available at SSRN: http://ssrn.com/abstract=1725503orhttp://dx.doi.org/10.2139/ssrn.1725503.

Lebel, A., Y. Kestens, C. Clary, S. Bisset, and S. V. Subramanian (2014). Geographic variability in the association between socioeconomic status and BMI in the USA and Canada. *PLoS ONE 9*(6), 11.

Leduc, K. (2009). Les structures d'accueil de la petite enfance au Luxembourg. In P. Bousch, T. Chilla, P. Gerber, O. Klein, C. Schulz, C. Sohn, and D. Wiktorin (Eds.), *Der Luxemburg Atlas/ Atlas du Luxembourg.* Emons.

Leishman, C. (2009). Spatial change and the structure of urban housing sub-markets. *Housing studies 24*, 563–585.

LeRoy, S. F. and J. Sonselie (1983). Paradise lost and regained: Transportation innovation, income, and residential location. *Journal of Urban Economics 13*, 67–89.

Liao, W.-C. and X. Wang (2012). Hedonic house prices and spatial quantile regression. *Journal of Housing Economics 21*(1), 16–27.

Licheron, J. (2013, September). Recent research results relating to the residential real estate market in the Grand Duchy of Luxembourg. Technical report, LuxReal CIGDL FORUM - From Macro to Micro.

Liu, S. and D. Hite (2013). Measuring the effect of green space on property value: An application of the hedonic spatial quantile regression. 2013 Annual Meeting, February 2-5, 2013, Orlando, Florida 143045, Southern Agricultural Economics Association.

Loi du 21 mai 1999 concernant l'aménagement du territoire (1999, Memorial A). Loi du 21 mai 1999 concernant l'aménagement du territoire. http://www.legilux.public.lu/leg/a/archives/1999/0061/a061.pdf, last accessed 22.03.2014.

Loi du 30 juillet 2002 déterminant différentes mesures fiscales destinées à encourager la mise sur le marché et l'acquisition de terrains à bâtir et d'immeubles d'habitation (2002, August 2002). Loi du 30 juillet 2002 déterminant différentes mesures fiscales destinées à encourager la mise sur le marché et l'acquisition de terrains à bâtir et d'immeubles d'habitation. Memorial A. http://www.legilux.public.lu/leg/a/archives/2002/0082/a082.pdf, last accessed 23.06.2014.

López, F. A., C. Chasco, and J. Le Gallo (2013). Exploring scan methods to test spatial structure with an application to housing prices in madrid. *Papers in R 0*, 30. Early View.

Luke, D. a. (2004). *Multilevel Modeling*, Volume 143. SAGE University Paper.

Lum, K. and A. E. Gelfand (2012). Spatial quantile multiple regression using the asymmetric Laplace process. *Bayesian Analysis 7*, 235–258.

Lutzenhiser, M. and N. Netusil (2001). The effect of open space on a home's sale price. *Contemporary Economic Policy 19*(3), 291–298.

Mahmoudi, P., D. Hatton MacDonald, N. D. Crossman, D. M. Summers, and J. van der Hoek (2013). Space matters: the importance of amenity in planning metropolitan growth. *Agricultural and resource economics 57*, 38–59.

Malpezzi, S. (1999, June). Estimates of the measurement and determinants of urban sprawl in U.S. metropolitan areas. Wisconsin-Madison CULER working papers 99-06, University of Wisconsin Center for Urban Land Economic Research.

Malpezzi, S. (2002, April). Hedonic pricing models: A selective and applied review. Online.

Malpezzi, S. and W.-K. Guo (2001). Measuring "Sprawl": Alternative measures of urban form in U.S. metropolitan areas. Working paper, University of Wisconsin Center for Urban Land Economic Research.

Maruani, T. and I. Amit-Cohen (2007). Open space planning models: A review of approaches and methods. *Landscape and Urban Planning 81*, 1–13.

McConnell, V. and M. Walls (2005). The value of open-space: evidence from studies of nonmarket benefits. Working paper, Resources for the Future.

McDonald, J. and D. P. McMillen (2011). *Urban Economics and Real Estate: Theory and Policy* (2 ed.). Wiley.

McMillen, D. (2003). Spatial autocorrelation or model misspecification? *International Regional Science Review 26*, 208–217.

McMillen, D. (2010). Issues in spatial data analysis. *Journal of Regional Science 50*, 119–141.

McMillen, D. P. (2013). *Quantile Regression for Spatial Data*. Springer.

MDDI (2010, November). *Plan national pour un développement durable.* Ministère du Développement durable et des Infrastructures.

Medard de Chardon, C. and G. Caruso (2010). Itinary function based on Google Maps. Online application.

Meijer, E. and J. Rouwendal (2006). Measuring welfare effects in models with random coefficients. *Journal of Applied Econometrics 21*(2), 227–244.

Melchiar, J. and K. Kaprová (2013). Revealing preferences of prague's homebuyers toward greenery amenities: The empirical evidence of distance-size effect. *Landscape and urban planning 109*, 56–66.

MIAT (2003). *Programme directeur de l'aménagement du territoire: Vers un développement durable du territoire*, Volume Partie A. Ministry of the Interior and Spatial Development,. http://www.dat.public.lu/strategies_territoriales/programme_directeur/detail_progr_direc/index.html, last accessed 22.03.2014.

MIAT (2004). IVL - an integrated transport and spatial development concept for Luxembourg: the idea, the analysis, the concept. Luxembourg.

Mieszkowski, P. and E. S. Mills (1993). The causes of metropolitan suburbanisation. *Journal of Economic Perspectives 7*(3), 135–147.

Mills, E. (1967). An aggregative model of resource allocation in a metropolitan area. *American Economic Review 57*, 197–210.

Mills, David, E. (1981). Growth, speculation and sprawl in a monocentric city. *Journal of Urban Economics 10*, 201–226.

Ministère du Développement durable et des Infrastructures (Ed.) (2012). *Stratégie globale pour une mobilité durable pour les résidents et les frontaliers.* Le Gouvernement du Grand-Duché de Luxembourg.

Möller, J. (2008). Regional variations in the price of building land: A spatial econometrics approach for west Germany. *The Annals of Regional Science 43*, 113–132.

Morenoff, J. (2003). Neighborhood mechanisms and the spatial dynamics of birth weight. *Journal of Sociology 108*(5), 976–1017.

Mueller, J. and J. Loomis (2013). Does the estimated impact of wildfires vary ith the housing price distribution? A quantile regression approach. Working paper, Northern Arizona University.

Muth, R. (1969). *Cities and Housing: the spatial pattern of urban residential land use.* Chicago: University of Chicago Press.

Nelson, N., E. Kramer, J. Dorfman, and B. Bumback (2004). Estimating the economic benefit of landscape pattern: An hedonic analysis of spatial landscape indices and a comparison of build-out scenarios for the protection of ecosystem functions. *Institute of Ecology, University of Georgia. Working paper*, 26.

Öner, z. (2013, Dec). RETAIL CITY: Does accessibility to shops explain place attractiveness? Working Paper Series in Economics and Institutions of Innovation 335, Royal Institute of Technology, CESIS.

Neuman, M. (2005). The compact city fallacy. *Journal of Planning Education and Research 25*, 11–26.

Nilsson, K., S. Paulit, S. Bell, C. Aalbers, and T. S. Lielsen (2013). *Peri-urban futures: Scenarios and models for land-use chane in Europe.* Springer.

Nilsson, P. (2014). Natural amenities in urban space: A geographically weighted regression approach. *Landscape and Urban Planning 121*(0), 45 – 54.

Observatoire de l'Habitat (2007, Septembre). Le potentiel foncier constructible pour l'habitat au luxembourg en 2004. *Note de l'Observatoire de l'Habitat 9*, 6.

Observatoire de l'Habitat (2007-2014a). Prix annoncés à la location. Online. http://observatoire.ceps.lu/publication_indicprix.cfm?cat=2location, last accessed 18.03.2014.

Observatoire de l'Habitat (2007-2014b). Prix annoncés à la vente. Online. http://observatoire.ceps.lu/publication_indicprix.cfm?pageKw=publiindicprix, last accessed 18.03.2014.

Observatoire de l'Habitat (2007-2014c). Prix de vente des appartements. Online. http://observatoire.ceps.lu/index.cfm?pageKw=publipe, last accessed 18.03.2014.

Observatoire de l'Habitat (2009, May). Préférences et satisfaction résidentielle des habitants du Luxembourg - un aperçu. *Note de l'Observatoire de l'Habitat 12*, 4.

Observatoire de l'Habitat (2009 - 2012). Type de propriétaires. Base disponibilité foncière 2007 et 2010. Technical report, Observatoire de l'Habitat.

Observatoire de l'Habitat (2010). *Le potentiel foncier et la consommation foncière dans les zones d'habitat au Grand-Duché du Luxembourg*. Observatoire de l'Habitat.

Observatoire de l'Habitat (2011, June). Synthèse des travaux 2010.

Observatoire de l'Habitat (2013, December). Note méthodologique sur la statistique des prix de vente des appartements. Technical report, Publication commune du STATEC et de l'Observatoire de l'habitat du ministère du Logement.

OECD (2007). *OECD Territorial Reviews: Luxembourg*. Organisation for Economic Co-operation and Development.

Ooi, J. T. and T. T. Le (2013). The spillover effects of infill developments on local housing prices. *Regional Science and Urban Economics 43*, 850–861.

Openshaw, S. (1984). *The Modifiable Areal Unit Problem* (CATMOG ed.), Volume 38. Geo Books.

Orford, S. (2000). Modeling spatial structures in local housing market dynamics: A multilevel perspective. *Urban Studies 37*(9), 1643–1671.

Orford, S. (2002). Valuing locational externalities : A GIS and multilevel modelling approach. *Environment and Planning B: Planning and Design 29*, 105–127.

Palmquist, B. R. (1984). Estimating the demand for the characteristics of housing. *The Review of Economics and Statistics 66*, 394–404.

Palmquist, B. R. (2003). Property value models. *Forthgoming in : Handbook of Environmental Economics 2*, 85.

Palmquist, R. B. (2006, June). Property value models. In K. G. Mäler and J. R. Vincent (Eds.), *Handbook of Environmental Economics*, Volume 2 of *Handbook of Environmental Economics*, Chapter 16, pp. 763–819. Elsevier.

Petit, S. (2007). Le processus de dépendance automobile au Grand-Duché du Luxembourg. *Population & Territoire 11*, 1–12.

Putnam, R. D. (2000). *Bowling Alone: The Collapse and Revival of American Community.* Simon & Schuster.

R Core Team (2013). *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing.

Ravetz, J., C. Fertner, and T. S. Nielsen (2013). The dynamics of peri-urbanization. In K. Nilsson, S. Paulit, S. Bell, C. Aalbers, and T. S. Lielsen (Eds.), *Peri-urban futures: Scenarios and models for land-use change in Europe*, pp. 13–44. Springer.

Reynolds, H. and C. Amrhein (1998). *Some effects of spatial aggregation on multivariate regression parameters*, Chapter 8, pp. 85–106. Kluwer Academic Publishers.

Ricardo, D. (1821). *On the Principles of political economy and taxation.* London: John Murray.

Rosen, H. S. (1974). Hedonic prices and implicit markets: Product differentiation in pure competition. *The Journal of Political Economy 82*, 34–55.

Sander, H., S. Polasky, and R. Haight (2010). The value of urban tree cover: A hedonic property price model in Ramsey and Dakota Counties, Minnesota, USA. *Ecological Economics 69*, 1646–1656.

Schindler, M. and G. Caruso (2014). Urban compactness and the trade-off between air pollution emisson and exposure: Lessons form a spatially explicit theoretical model. *Computers, Environement and Urban Systems 45*, 13–23.

Schreurs, E., S. Van Passel, L. Peeters, and T. Thewys (2013). Analyzing the impact of soil contamination on farmland values: A spatial hedonic approach using quantile regression. In *VII World Conference of the Spatial Econometrics Association, Washington D.C., 10/07/2013 - 12/07/2013.*

Seya, H., Y. Yamagata, and M. Tsutsumi (2013). Automatic selection of a spatial weight matrix in spatial econometrics: Application to a spatial hedonic approach. *Regional Science and Urban Economics 43*, 429–444.

Shin, W.-J., J. Saginor, and S. Van Zandt (2011). Evaluating subdivision characteristics on single-family housing value using hierarchical linear modeling. *Journal of Real Estate Research 33*(3), 317–348.

Shultz, S. D. and D. A. King (2001). The use of census data for hedonic price estimates of open-space amenities and land use. *Journal of Real Estate Finance and Economics 22*(2/3), 239–252.

Siedentop, S. and S. Fina (2012). Who sprawls most? exploring the patterns of urban growth across 26 European countries. *Environement and Planning A 44*, 2765–2784.

Sieverts, T. (2001). *Zwischenstadt: Zwischen Ort und Welt, Raum und Zeit, Stadt und Land* (3 ed.). vieweg.

Sieverts, T. (2006, May). Neither city nor countryside. Online. http://www.goethe.de/kue/arc/dos/dos/sls/zup/en1488502.htm.

Sirmans, G. S. and D. A. Macpherson (2003). The composition of hedonic pricing models: a review of literature. *National association of realtors 0*, 83.

Sirmans, G. S., D. A. Macpherson, and E. N. Zietz (2005). The composition of hedonic pricing models. *Journal of Real Estate Literature 13*, 3–43.

Smarth Growth Network (2014). Smarth growth online. http://www.smartgrowth.org/why.php, last accessed 31.03.2014.

Smith, V., C. Poulos, and K. H. (2002). Treating open space as an urban amenity. *Resource and Energy Economics. 24*, 107–129.

Snijders, T. A. B. and R. J. Bosker (1999). *Multilevel Analysis*. SAGE Publications.

Sohn, C. (2006, October). Villes et agglomérations au Grand-Duché de Luxembourg. Proposition d'une nomenclature des unités urbaines. *POPULATION & TERRITOIRE CEPS/INSTEAD 10*, 1–16.

Song, Y. and G.-J. Knaap (2004). Measuring urban form is Portland winning the war on sprawl? *Journal of the American Planning Association 70*, 210–225.

Sopranzetti, B. J. (2010). Hedonic regression analysis in real estate markets: A primer. In C.-F. Lee and J. Lee (Eds.), *Handbook of Quantitative Finance and Risk Management*. Spring.

StataCorp (2009). Stata statistical software: Release 11.

STATEC (2014, 04). Dossier indice des prix à la consommation et inflation. Technical report, STATEC. last accessed 21.05.2014.

Su, L. and Z. Yang (2007). Instrumental variable quantile estimation of spatial autoregressive models. *SMU Economics & Statistics Working Paper Series 05-2007*, 35.

Sudhira, H., T. Ramachandra, and K. Jagadish (2004). Urban sprawl: metrics, dynamics and modelling using GIS. *International Journal of Applied Earth Observation and Geoinformation 5*, 29–39.

Taylor, L. (2008). Theoretical foundations and empirical developments in hedonic modelling. In A. Baranzini, J. Ramirez, C. Schaerer, and T. Philippe (Eds.), *Hedonic Methods in housing markets, Pricing Environmental amenities and segregation*, Chapter 1, pp. 15–37. Springer.

Thériault, M., F. Des Rosiers, and F. Joerin (2005). Modelling accessibility to urban services using fuzzy logic. *Journal of Property Investment and Finance 23*(1), 22–54.

Thériault, M., F. Des Rosiers, and M.-H. Vandersmissen (1999). GIS-based simulation of accessibility to enhance hedonic modeling and property value appraisal: An application to the Quebec City metropolitan area. Technical report, Université Laval.

Tiebout, C. M. (1956). A pure theory of local expenditures. *The Journal of Political Economy 64*(5), 416–424.

Treg, C. (2010). *A multilevel property hedonic approach to valuing parks and open space*. Ph. D. thesis, University of Vermont.

Turner, M. (2005). Landscape preferences and patterns of residential development. *Journal of Urban Economics 57*, 19–54.

Uematsu, H. and A. K. Mishra (2012). The impact of natural amenity on farmland values: A quantile regression approach. In *Southern Agricultural Economics Association Annual Meeting, Birmingham, AL, February 4-7, 2010*.

UN (2011). *Population Distribution, Urbaniszation, Internal Migration and Development: An International Perspective*. United Nations Department of Economic and Social Affairs Population Division.

Uyar, B. and K. H. Brown (2007). Neighborhood affluence, school-achievment scores, and housing prices: Cross-classified hierarchies and HLM. *Journal of housing research 16*, 97–116.

Van Rijswijck, A. and C. Wagner (2011). Der kommunale Flächennuzungsplan. In T. Chilla and C. Schulz (Eds.), *Raumordnung in Luxemburg / Aménagement du territoire au Luxembourg*, pp. 160–170. Openscience, Guy Binsfeld.

Vanoutrive, T. and A. Parenti (2009). On proximity and hierarchy: exploring and modelling space using multilevel modelling and spatial econometrics. In *49th European congress of the Regional Science Association International*. ERSA Congress 2009.

Vaughan, L., S. Griffiths, M. Haklay, and C. E. Jones (2009). Do the suburbs exist? Discovering complexity and specificity in suburban built form. *Transaction of the Institute of British Geographers 34*, 475–488.

Verbeek, M. (2008). *A guide to modern econometrics*. Wiley.

Viton, P. A. (2010). *Notes on spatial econometric models*. Ohio State University.

Von Thünen, J. H. (1826). *Der isolierte Staat in Beziehung auf Landwirtschaft und Nationalökonomie*. Verlag von Gustave Fischer.

Ward, M. D. and K. S. Gleditsch (2008). *Spatial regression models*. Quantitative Applications in the Social Sciences. SA.

Whang, Y.-J. (2006). Smoothed empirical likelihood methods for quantile regression models. *Econometric Theory 22*, 173–205.

White, H. (1980, 05). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica 48*(4), 817–838.

Wilhelmsson, M. (2002). Spatial models in real estate economics. *Housing, Theory and Society 19*, 92–101.

Wooldridge, J. M. (2009). *Introductory econometrics, A modern approach*. South-Western CENGAGE Learning.

Wu, J. and A. Plantinga (2003). The influence of public open space on urban spatial structure. *Journal of Environmental Economics and Management 46*(2), 288–309.

Wu, W. and G. Dong (2013). Valuing the "green" amenities in a spatial context. *Journal of Regional Science 00*, 1–17.

Yang, C. H. and M. Fujita (1983, January). Urban spatial structure with open space. *Environment and Planning A 15*(1), 67–84.

Youssoufi, S. (2011, December). *Satisfaction résidentielle et configurations spatiales en milieu périurbain.* Docteur en géographie, Université de Franche-Comté.

Zahlen, P. (2012a). 50 ans de migrations. *50 ANS STATEC 12*, 6.

Zahlen, P. (2012b). 50 ans d'évolution démographique. *50 ANS STATEC 11*, 7.

Zahlen, P. (2012c). L'évolution économique globale du Luxembourg sur la longue durée. *50 ANS STATEC 22*, 8.

Zahlen, P. (2012d). Regards sur les bâtiments achevés. *REGARD-STATEC 2*, 4.

Zeileis, A. (2004). Econometric computing with HC and HAC covariance matrix estimators. *Journal of Statistical Software 11*(10), 1–17.

Zeileis, A. (2006). Object-oriented computation of sandwich estimators. *Journal of Statistical Software 16*(9), 1–16.

Zeileis, A. and T. Hothorn (2002). Diagnostic checking in regression relationships. *R News 2*(3), 7–10.

Zietz, J., Z. E. Norman, and S. G. Sirmans (2008). Determinants of house prices: A quantile regression approach. *Journal of Real Estate Finance and Economics 37*, 317–333.