EFFECTS OF CROWDING, DENSITY AND DEPRIVATION ON RESIDENTIAL SATISFACTION

EILYA TORSHIZIAN

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR

THE DEGREE OF DOCTOR OF PHILOSOPHY IN ECONOMICS

THE UNIVERSITY OF AUCKLAND

MARCH 2017

Abstract

"Effects of crowding, density and deprivation on residential satisfaction"

This thesis assesses the impact of household crowding, population density and the proximity of rich and poor in the city on people's satisfaction with their residence (which I term 'residential satisfaction'). The relationships are examined at a fine geographic level accounting for a wide range of socioeconomic factors and by using appropriate spatial regression techniques.

The thesis is also motivated by a range of theoretical, empirical and technical questions:

Policy: Governments are trying to make economic development more sustainable. When it comes to the growth of cities the primary policy tool for increasing sustainability is residential intensification. However, the primary outcome of intensification is greater population density, higher rents (which leads to higher levels of crowding within dwellings) and, in many cases, the increased proximity of widely disparate income groups. It is possible therefore that in an attempt to become more sustainable, increasing intensification will lead to a reduction in the overall wellbeing of urban residents.

Theory: Urban intensification can raise rents without raising (and possibly lowering) residential satisfaction. Therefore the economists conventional measure of value, price, may not be adequate to judge the welfare consequences of intensification. The relative merits of judging policy outcomes on the basis of market price and individual measures of satisfaction are unclear. Rent (land price) is a collectively (market) generated objective measure of value accorded a site, while residential satisfaction is a subjective measure of the qualitative value of that site to the individual. The problem is that when applied to any given residence the two measures of 'value' may be weakly correlated. They may even be negatively correlated – reflecting the strong preference of New Zealanders for low density living. This thesis therefore argues for the value of using both price and residential satisfaction measures in developing policy around residential intensification.

I apply econometric methods to three rounds of the New Zealand General Social Survey, 2008, 2012, 2014 plus 2006 Census data covering the area administered by the Auckland Council and draw four main conclusions:

Firstly, after taking people's location choice into account, a higher level of population density in the resident's own suburb lowers residential satisfaction. At the same time, residents are more sensitive to the *perception* of crowding [in the dwelling] than with objective measures of crowding.

Secondly, a higher level of crowding in the person's household than expected, which is estimated based on their background, decreases residential satisfaction.

Thirdly, people respond to reference groups, which in the Auckland case are found to be those who live within a 15-minute walk of the resident. After taking into account the endogeneity of residential choice, people are more sensitive to their own household crowding levels compared to that of their neighbours.

Fourthly, residents are sensitive to the relative quality of neighbouring areas. In residential satisfaction terms, residents of (relatively) rich areas are negatively affected by the lower affluence of neighbouring areas, but residents of poorer areas are positively affected by the superior amenities of higher income neighbours. Therefore, while one group is sensitive to 'not-in-my-backyard' (NIMBY) effects, the other sees value in being located near high value areas.

In summary, this thesis demonstrates the worth of using residential satisfaction in urban economic studies. The results indicate that people apply expectations based on their social norms when evaluating their living environment. After taking these social norms into account, people value living in less crowded houses, in less dense suburbs and in proximity to affluent areas.

Keywords: Household crowding; Perceived crowding, Perceived density; Canadian National Occupation Standard, Room density; Neighbourhood satisfaction; Housing satisfaction; Population density; Residential density; Relative crowding; Social norms; Social comparison; Deprivation; Poverty trap; SPWT; Spatial weight matrix, Walking distance weights, K-nearest neighbour weights, spatial distance weighting matrix, Spatial econometrics.

JEL codes: C21; C26; I3; I31; O12; R12; R23; Z13.

To

MY FATHER

ACKNOWLEDGEMENT

I am truly and deeply indebted to so many people that there is no way to acknowledge them all, or even any of them properly.

I would like to express my deep and sincere gratitude to my supervisor, Professor Arthur Grimes, for his helpful and enthusiastic supervision.

I am grateful to the many people whose helpful insights and comments on earlier drafts have greatly improved the quality of this thesis, particularly my co-supervisor Dr Erwann Sbai, and my advisor Professor Sholeh Maani. Also, I would like to acknowledge Professor Ananish Chaudhuri and Chris Parker, for their support at different stages of my PhD studies. I am grateful for the valuable comments I received from two anonymous reviewers of this thesis.

I would like to thank the participants at the 2014 NZAE conference, especially Dr Adam Jaffe and Professor Jacques Poot; the attendees of the Economics Department's PhD workshop at the Auckland University, particularly Dr Alan Rogers and Dr John Hillas; the participants of the 2015 WEAI conference and Professor Philippa Howden-Chapman and Professor Karen Witten for their useful comments in the Resilience Urban Futures (RUF) 2013 meeting.

Funding from the New Zealand Centre for Sustainable Cities and the Chief Economist Unit of Auckland Council is gratefully acknowledged. I would also like to thank Statistics New Zealand for providing the data.

Finally, I wish to thank my parents, Mohammad and Farzaneh, for their love and encouragement, and my sisters, Elmira and Elnaz, without whom I would never have enjoyed so many opportunities. Any flaws that remain in this thesis are my own.

Disclaimer

Contents

List of Tables

List of Figures

1 Introduction

1.1 Overview

The benefits of agglomeration[1] have been an area of significant interest in the recent urban economics literature. Agglomeration benefits have led local governments around the world to plan for a higher intensification.[2] However, this competition amongst municipalities needs to be associated with a careful assessment of intensification's impact on residents, and particularly on their satisfaction with their living environment.

In order to achieve sustainable development, I must recognise and overcome the issues that may arise in the process of development. This can be done by having a better understanding of the structural issues of cities.[3]

The urban development process can be viewed through the lens of history. Major cities, including urban areas in New Zealand, have changed materially in terms of the planning of dwellings. For example, kitchens have changed in size. Helen Leach described this phenomenon in a book on New Zealand's kitchens in the 20th century (Leach, 2014), referring to the 1990s as a decade of 'small households and super-sized appliances'.

Population patterns across cities have changed, reflecting history and the level of their development. In the process of growth, the services sector grows much faster than other sectors of the economy.[4] This shift in the composition of the economy of cities is often associated with the concentration of jobs in the centre of cities and may lead to a higher level of density in city centres. However, living in city centres may not be a direct result of the location of jobs. In fact, jobs may have been created in city centres because of an exogenous[5] factor that has pushed households to live in the city centres. Therefore, it is important to recognise the factors that lead people to form clusters exogenously. This thesis employs spatial econometric techniques backed by rich datasets to reach unbiased estimates of the impacts that crowding, density and deprivation have on location and residential satisfaction.

---

[1] Agglomeration benefits are those derived from cities' agglomeration economies, which are the benefits caused by proximities of people and/or firms to each other.

[2] The direction of the relationship between density and agglomeration is not yet clear.

[3] Traditionally, urban economic studies need to address certain issues, such as location choice, spatial proximities and the definition of neighbourhood boundaries. However, the complexities associated with the methodology and the lack of data on small geographic scales often stand in the way of the causal inference.

[4] Some economists have argued that the shift from manufacturing to services is caused by the limitations on the expansion of a city's boundaries.

[5] By exogenous, I mean factors that have not been affected by other drivers of location choice. For example, rainfall is an exogenous factor in the causal system consisting of the process of farming and crop output.

Throughout the thesis, the outcome of interest is residential satisfaction, which is people's satisfaction with their residential environment. To the best of my knowledge, the lack of data has meant that this variable has rarely been considered as an outcome of regional studies. However, as I will discuss, it is a useful measure of the utility gained from the living environment.

Crowding and density are two challenging concepts of urban economics that have often been used as a basis for policy advice. Most regional plans aim at specific levels of intensification; in other words, levels of crowding and density. However, our knowledge of their impact on the residents' satisfaction is minimal. The measures of intensification may be defined at different geographic scales, such as very small suburbs or large neighbourhoods. They also vary depending on their construction method. For example, while both residential density and area density measure the level of density in a neighbourhood, their different definitions may lead to variations in their values. Hence, the first question to be answered is: What measure is appropriate to be used?

I have already explained the importance of the second half of this thesis's title: 'residential satisfaction'. The first part, 'crowding, density and deprivation', refers to the nature of the location of the properties and their residents. A novel element of this thesis is the way that it addresses the location choice. People choose to live in an environment and to have a specific level of crowding based on their social norms derived from their backgrounds. The impact of intensification needs to be assessed after taking the complexities associated with people's norms into account. People may or may not be happy with their living environment when they compare it with what they were expecting to provide themselves with. Therefore, a second question that will be addressed is: What impact does intensification have on the satisfaction of residents?

In addition to a comparison with the expected living conditions, people may compare themselves with the living environments of others. This study includes an analysis of social comparisons, and must therefore consider the relevant population that people consider as 'relevant others'. After understanding the comparison group, the second challenge is to define the best measure for social comparisons, constructed based on the characteristics of the residential environment. Potential comparison groups at different geographic scales include: neighbours living over different geographic scales, neighbours who are the symbols of success over different geographic scales, and neighbours who may be visited in a 15-minute walk (or some other time-defined walking distance), which is a walking distance definition for neighbourhood boundaries. The question to be asked is: Are people affected by their relative residential position in their evaluations of their living environment?

Another aspect of 'real estate characteristics' is the deprivation levels of residents' neighbourhoods. Based on their backgrounds and their other characteristics, people live in a neighbourhood with a specific level of average deprivation. People may compare themselves with the residents of their adjacent neighbourhood and may suffer from an envy effect when their adjacent neighbourhood is rich. However, residents of poor areas may also enjoy the amenities provided by adjacent rich neighbourhoods. On the other hand, residents of the rich areas may or may not be happy with being located next to a poor area. Hence, the question to ask is: How does the proximity to rich areas affect the satisfaction of the residents of poor areas, and vice versa?

The methodology of the study is based on spatial binary dependent variable modelling, as explained in Chapter 7. The endogeneity of the variables of interest has been taken into account by using a two-stage least squares estimator. The study aims to account for the complexities of the causal inference at the same time as keeping the methodology consistent across the chapters.

1.2 Motivation and contribution

With overall living standards having risen since the Second World War, economists have increasingly concentrated on the determinants of satisfaction, both with life as a whole and with different aspects of life (Dolan, Layard, & Metcalfe, 2011; Helliwell, Layard, & Sachs, 2012). Aspects of satisfaction include health, job and economic satisfaction, as well as residential satisfaction; that is, satisfaction with residential living conditions. This form of satisfaction has always played a prominent role in life.

Residential living conditions and residential satisfaction may be influenced by regulatory settings, including those related to planning restrictions. For instance, in order to contribute to more sustainable urban development, there is growing interest in increasing the intensification of urban areas by using policy instruments (Howley, Scott, & Redmond, 2009). However, the outcome of intensification may be a higher household crowding level (for example, the number of people per room), in addition to the (intended) increase in population density (number of people per unit of area). I examine how crowding and density affect people's levels of residential satisfaction.

The sample population of this study is the Auckland region, including the city of Auckland, which has consistently been ranked as one of the top five cities in the world according to the Mercer Quality of Living survey (Mercer, 2015). While the city is polycentric, approximately 33 per cent of its population live in the Central Auckland urban zone (Grimes & Liang, 2009; Johnston, Pousen, & Forrest, 2009; Maré, Coleman, & Pinkerton, 2011).[6]

Higher intensification has been cited as a trigger of regional development in many agglomeration studies (Glaeser, 2010). A large proportion of the regulations introduced by the Proposed Auckland Unitary Plan (PAUP) aim at intensification (Auckland Council, 2013; Grimes & Mitchell, 2015). However, our knowledge of the impact of intensification on outcomes, such as residential satisfaction, broader wellbeing or house prices, is very limited.

[6] This proportion reflects demand for locating in the CBD after accounting for the high price of housing in this area caused by regulations and other factors. For example, as Torshizian (2015a) concluded, resource consent regulations targeted at the goals of Resource Management Act 1991 (RMA) negatively affect regional development, which is positively correlated with the density of neighbourhoods.

In order to achieve a better understanding of the impact of intensification on residential satisfaction, good measures of crowding and density are needed. The measures can be categorised based on how subjective they are. For example, measures may be objective (for example, crowding measured as the number of people per room) or subjective (for example, based on qualitative survey information). Both types of measures are used in the present study.

1.3 Thesis outline

This thesis comprises four main essays, plus five introductory chapters, a technical chapter related to spatial effects estimation, and a concluding chapter.

Relevant topics to the overall concept of the thesis will be reviewed in the Chapter 2, which will also introduce Auckland's geographical position and look briefly at its housing market. I will then explain the measurement of residential satisfaction as the outcome of interest. In the next part of Chapter 3, I review the factors that most affect residential satisfaction. Neighbourhood effects, which are a major issue in urban economic study, are also examined.

Chapter 3 provides the statistical discussion relevant to the following chapters. This includes an overview of the data and sample used in the following chapters and an introduction to geographic scales. The specific variables of interest to each chapter will be presented in the descriptive statistics section of each chapter.

To my knowledge, this is the first comprehensive urban economic study on the satisfaction of residents with their living environment. Hence, Chapter 4 presents the previous studies' approach using prices as their outcome of interest. The use of subjective measures and their position in the economics literature will be reviewed. The chapter finishes with a review of the features of subjective measures and a comparison between the information that prices and residential satisfaction hold.

Chapter 5 presents the common methodology that the following chapters will use. I will present a general form of the utility function, followed by the common econometric approach amongst the following chapters. This chapter also discusses the appropriateness of the recategorisation of the five-category outcome of interest to a binary variable.

After a comprehensive review of the literature (in Chapter 2) and of the statistics and the methodology (in the subsequent chapters), I present four essays plus a technical chapter. The first essay (Chapter 6) compares the impact of three measures of crowding on residential satisfaction. The first measure is people per bedroom (sometimes called room density), which is a raw crowding measure. This measure is constructed based on objective characteristics of the observations. The second measure, the Canadian National Occupation Standard (CNOS), is a crowding index measure, which takes into account subjective issues such as the cultural understanding of crowding. The other measure is perceived crowding (PC), which measures crowding from the individual's point of view. The results of this chapter provide the basis for the analysis in the subsequent chapters.

Surprisingly, few empirical studies have analysed whether a higher crowding level contributes positively or negatively to residential satisfaction. This point is discussed in Chapter 7, which is also the second essay of this thesis. That chapter introduces a method to overcome the potential endogeneity of household crowding, and also discusses the differences between the expected level of crowding, derived from individuals' norms, and the unexpected crowding level, which together lead to the actual position of people. Chapter 7 discusses the impact of households' absolute residential positions, particularly the level of household crowding and the level of area density, in determining the satisfaction of residents after taking the endogenieties into account.

Chapter 8, which is the third essay of this thesis, discusses the role of people's relative residential position on their residential satisfaction. A number of measures of relative crowding and relative density will be developed by using different definitions for neighbourhood boundaries. By comparing the predictive power of the relative measures for residential satisfaction, this chapter finds the measures that best predict residential satisfaction.

Chapter 9 starts with an introduction to relevant deprivation studies and to the poverty trap literature. The chapter hypothesises that, for poor neighbourhoods, the positive amenity effect of locating near an affluent area outweighs any (negative) envy effect. The estimates in this chapter include all of the variables of interest from the previous chapters, together with the chapter's specific variable of interest, which is relative deprivation level.

In order to construct the relative measures discussed in chapters 8 and 9, a spatial tool is needed to define neighbourhood boundaries at different walking distances from an individual's location. Due to confidentiality issues, I had to access (the full file of) the regional-level data in Statistics New Zealand's datalab and was therefore unable to access geoprocessing software packages, such as ArcGIS. Another issue regarding the use of a geographical analysis software package combined with statistical software is the potential for compatibility issues and extra processing time. Therefore, Chapter 10 describes a program that I developed in Stata for the purpose of the constructing the spatial weight matrices used in the previous chapters. Chapter 10 begins with a brief introduction to the problem and to the spatial weight matrices, followed by a description of the package and some examples of the use of the package are presented. The code is provided in the chapter's appendix.

2 Common Literature

Chapter outline

Since this entire thesis falls within the context of urban economics, the common parts of the subsequent chapters' literature are discussed in this chapter in order to avoid repetition. The literature related to the main outcome variable in this study – that is, residential satisfaction – is presented in Section 2.2. Some variables, such as homeownership, are commonly used in the next chapters, mainly as control variables. The most relevant previous studies that have discussed these variables as their variable of interest are presented in Section 2.3.

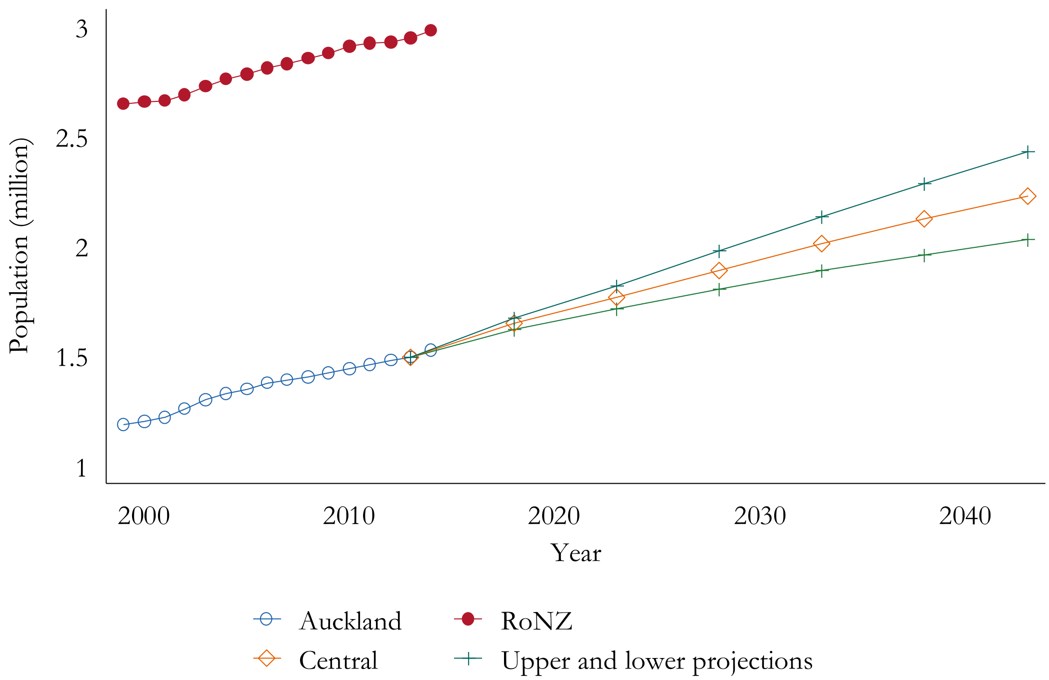
Neighbourhood effects are a principal concern of urban studies and are controlled for in this study. An example of neighbourhood effects would be the settlement of new owners in state houses, who may motivate their peers to behave similarly. Settlement of new residents may also lead to a change in the characteristics of their host neighbourhood by inducing a shift outward of dissimilar residents. These effects constitute a membership problem that is addressed by controlling for 'own group' attraction factors, such as ethnicity, income and country of birth. A discussion of neighbourhood effects is presented in Section 2.4, in which the definition of neighbourhood and the aggregate and individual effect of neighbourhood are discussed.

2.1 Auckland, the City of Sails

According to Mercer's Quality of Living survey, Auckland has been one of the most liveable cities for many years. In June 2014, Auckland was home to 1,527,100 people. As depicted in figure 1, the population of Auckland increased considerably in the years leading up to 2013 and is expected to rise in the next 30 years by almost 1 million people[7]. The increased population places greater strain on available land supply and so may affect crowding and density outcomes.

[7] RoNZ in the table shows the population for the rest of New Zealand.

Figure 1. Auckland's population, 2000–2043.



Produced by the Chief Economist Unit, Auckland Council
Source: Statistics New Zealand

Note: For data on predictions (base: 2013) see Appendix I.

A major goal of Auckland Council (AC) is to make Auckland the most liveable city in the world, which will attract new migrants and an associated increase in the demand for housing (Yeabsley & Stephenson, 2012).

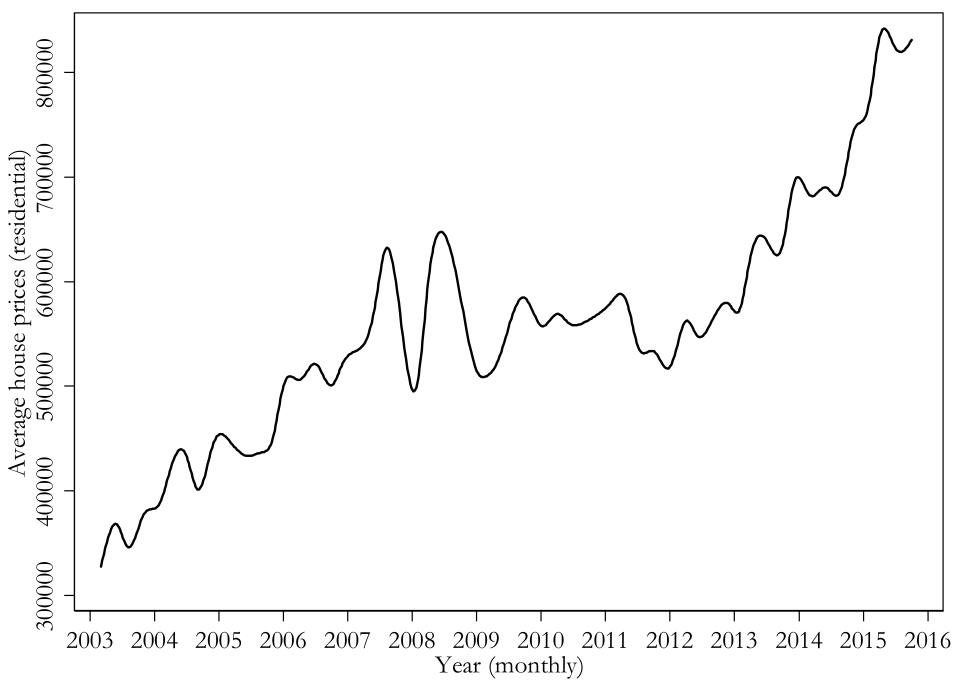
However, Auckland already faces a shortage in housing supply, which has led to household crowding (OECD, 2015). Regulations are considered one of the reasons for the lack of supply in Auckland. The 1991 Resource Management Act (RMA) affects Auckland's housing market by affecting new supply. The RMA's purposes in regulating the market are to avoid adverse effects on the environment, to manage development and land use effects, to boost efficiency in the use of natural and physical resources, to improve the quality of the environment, to prevent natural hazards, to control the subdivision of land, to control the emission of noise, and to maintain amenity values (MfE, 2010).

The Auckland Plan (AP) is the AC's instrument for guiding Auckland's future over the next 30 years (Auckland Council, 2013). The stated goals of the AP are to reduce transport and housing shortages, give children and young people a better start, create more jobs, and protect the environment. One aspect of the AP is the relaxation of Metropolitan Urban Limit (MUL), which imposes a limitation to the development beyond the urban boundaries. After controlling for a wide range of factors, Grimes and Liang (2007) concluded that the MUL leads to a boundary land value ratio of between 7.9 and 13.2; in other words, land in the urban area may be up to 13.2 times more expensive than land located just across the urban boundary (further discussed by Moshrefi, Kim, and Torshizian (2015a); Torshizian (2016b); Zheng (2013)).

Another factor affecting the shortage in housing supply is the geography of the city. Auckland is positioned on a narrow isthmus between the Waitemata and Manukau Harbours. The distribution of prices depicts the revealed preferences for locating near the city centre for Aucklanders (see figure 3), which means I would expect to see higher density in areas of high land and house prices.

As depicted in figure 2, Auckland has seen a remarkable increase in average house prices over the last few years, from almost $330,000 in 2003 to almost $800,000 in 2015.[8] In March of 2016, Arthur Grimes announced the need to increase Auckland's housing stock by 150,000 homes in order to reduce the level of prices by 40 per cent. Considering the size of Auckland's housing market, which only provides a maximum of 7500 new dwellings per year, the shortage is significant.

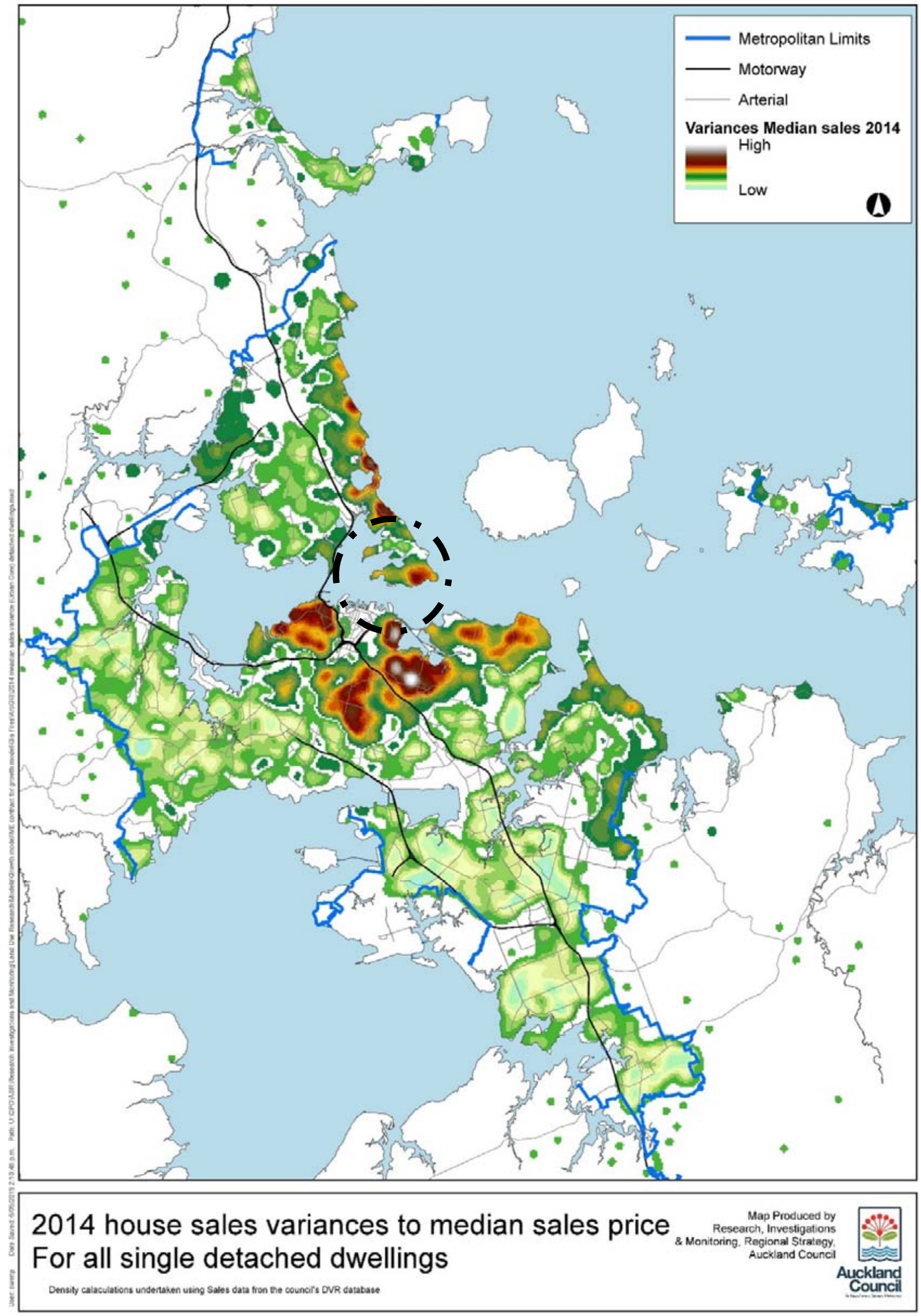
Figure 2. House prices in Auckland between 2003 and 2015.



Source: Torshizian (2016b).

[8] The decrease in the end of the period, in the end of 2015, is temporary; that is, the prices continued increasing after the beginning of the new year in 2016.

Figure 3. The distribution of prices for detached houses. The area inside the dashed circle is the CBD.

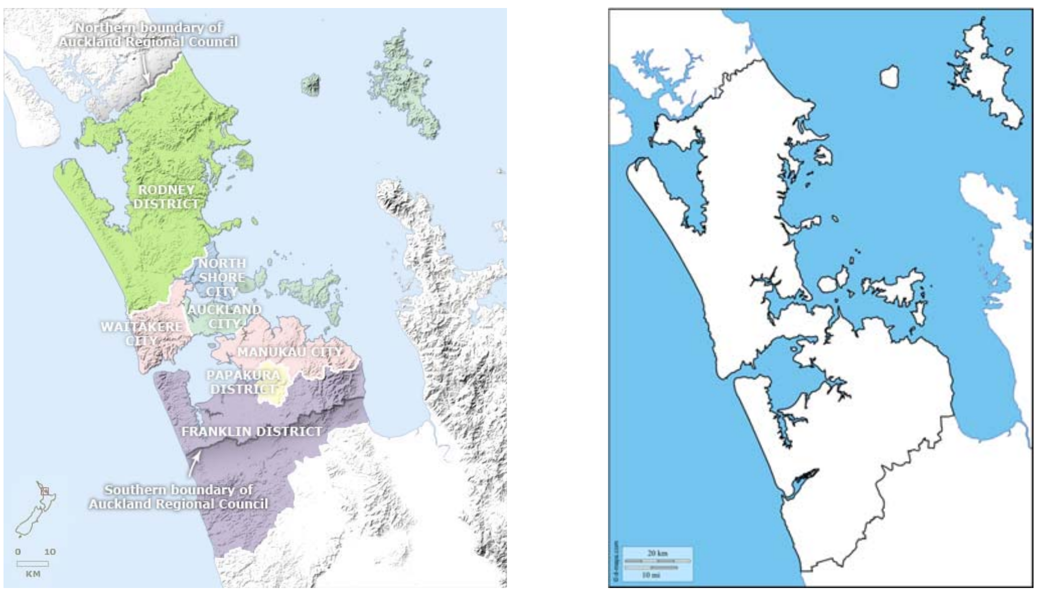


Note: For the purpose of illustration, the area demonstrated in this figure is the metropolitan area; that is, the figure does not illustrate the whole Auckland region. The range of colours represents the median sale prices amongst neighbourhoods.

Expansion of the city at the city fringe requires costly infrastructure upgrades. Therefore, the AC's main goal is to encourage greater intensification in parts of Auckland. Many of the building regulations in the AP aim at higher density in areas with relatively good transport connectivity. The population pattern in Auckland is also changing as the number of one- and two-person households increases (Statistics New Zealand, 2012a); this is a factor that needs to be taken into account in the intensification policies.

Until 2010, Auckland region consisted of one regional authority plus seven district authorities. The former Auckland councils are depicted in the left hand side of figure 4. However, in the late 2000s, the central government realised that having too many local authorities may have hindered Auckland's economic progress (Salmon & Shand, 2009), and started building up a stronger regional government. This led to an amalgamation under one local council, called the Auckland Council. Throughout this thesis, 'Auckland' refers to the area surrounded by the black line as depicted on the right hand side of figure 4. This area is covered by Auckland Council with an area of 4894 km².

Figure 4. Auckland city before and after amalgamation.



Source: McClure (2016); Torshizian (2017).

Summary

Auckland is one the world's most livable cities but has experienced a significant increase in its housing prices during the last few years. This can be attributed to an increase in demand that is not compensated by a matching increase in supply. The latter may be addressed by an increase in the supply of land, through an expansion of the city's urban area, or further intensification in the current urban area. Considering that Auckland Council has aimed to limit further expansion of the city in the next 30 years,[9] it is important to improve our understanding of the impact of intensification on residents' satisfaction – a topic that this dissertation addresses.

2.2 Residential satisfaction measurement

This section provides information about the measurement of this thesis's outcome of interest, which is residential satisfaction. The complexities involved in using subjective measures will be discussed in Sections 4.3 and 5.4.

Some previous studies have studied dwelling satisfaction and neighbourhood satisfaction separately, despite the fact that the two factors have a high correlation (Amérigo & Aragonés, 1997; Lu, 1999; Rodgers, 1982). Although the results of the previous studies have a different context and are usually in correlational terms, they confirm the importance of socioeconomic factors in people's evaluation of their living environment (for example, see Bonnes, Bonaiuto, and Ercolani, 1991). Rodgers (1982) provided a very useful understanding of the psychology of crowding, which will be reviewed in Section 6.2.1. In this study, I consider residential satisfaction as a variable that contains both housing and neighbourhood aspects.

Various studies have measured residential satisfaction (RS) in different ways. Some have measured it by asking respondents about their agreement or disagreement with certain statements on a Likert scale, for example, 1-strongly disagree… 5-strongly agree. In the present study, respondents answered the residential satisfaction question of 'How do you feel about where you are currently living?' using a Likert scale that ranged from '1 – very dissatisfied' to '5 – very satisfied', which is an ordered categorical base.

In addition to this five-level Likert item, respondents had a 'don't know' and a 'refused to answer' choice, which leads to more accurate qualitative measurement (the increased accuracy in results caused by including a 'don't know' category in qualitative evaluations is discussed by Fabling, Grimes, and Stevens (2012) and presented in Section 2.4.4).

The answer to the question of 'How do you feel about where you are currently living?' can be considered as the respondent's understanding of his or her satisfaction with his or her living environment as a whole, which includes the individual's 'neighbourhood satisfaction' and 'housing satisfaction', although it could alternatively be construed as an answer to just one of these aspects. However, before asking this question, the person administrating the questionnaire states, 'I am now going to ask you some questions about your house/flat'. This statement may imply that the respondent is only thinking about his or her housing satisfaction. However, for a better understanding of the respondent's understanding of the question asking

[9] Auckland Council's plan is published in its Auckland Plan, which introduces providing affordable housing as one of its fundamental goals. The Auckland Plan aims to provide a maximum of 30 per cent of the future developments in areas outside the current urban area.

'How do you feel about where you are currently living?', I need to look at the respondents' understanding of the survey as a whole.

The General Social Survey, which is used in the current study, consists of many modules;[10] namely, overall life satisfaction module, health module, knowledge and skill module, paid work module, economic standard of living module, housing module, physical environment module, safety and security module, support across households module, social connectedness module, leisure and recreation module, cultural and identity module, and human rights module. In each module, respondents answer a series of questions about different aspects of that general module. For example, in the knowledge and skill module, the respondents were asked 'Looking at Showcard 7, in general, how do you feel about your knowledge, skills and abilities?' and then asked about aspects of their knowledge and skills.

In the housing module, the respondents were asked about their problems with their dwelling and their neighbourhood. This implies that the questionnaire is not designed just for satisfaction with housing, but also satisfaction with neighbourhood; that is, with the respondent's living environment in general. The respondents' experience of living in their current place[11] is likely to capture the shortcomings that they may have in both their dwelling and their neighbourhood. I will discuss this issue further in Section 3.3.

In this study, the residential satisfaction variable is aggregated to a binary variable, such that the first four categories – '1 – strongly disagree', '2 – disagree', '3 – neither agree nor disagree, '4 – agree,' – are aggregated to one group and the other category – '5 – strongly agree' – to another group. There are several reasons for this aggregation. The first is the importance of satisfaction versus dissatisfaction from a policy-maker's point of view. Policy-makers often aim to ensure that people are satisfied with their situation; therefore, from a practical point of view, it makes sense to have a clear bifurcation between a satisfied versus a dis-satisfied group. Second, the distribution of this variable is such that a high proportion of people are satisfied/very satisfied. Therefore, I consider that the medium group ('neither agree nor disagree') and the 'satisfied' group indicate a *relatively* negative response. Third, as will be discussed in Section 5.3 (and in Appendix ii of Chapter 5), I found no significant differences in the slope of the coefficients of the variables of interest by taking into account a more general ordered design. Therefore, the aggregation does not lead to any material loss of precision.[12] For a comprehensive discussion on the aggregation of the RS variable, see Section 5.3.

The advantages of using a subjective measure as the outcome of interest are discussed in Chapter 4.

---

[10] Questionnaire for the 2012 survey: http://www.stats.govt.nz/~/media/Statistics/surveys-and-methods/completing-a-survey/faqs-about-our-surveys/nzgss/gss-questionnaire-2012.pdf.

[11] This is related to the 'retrospective' nature of the subjective measures, which is discussed in Section 4.3.

[12] In addition to the noted reasons, the ease of illustration of the results is another reason for the aggregation of the five categories into two groups.

Summary

Residential satisfaction is the outcome of interest in the current study. The variable is derived from the answers of the respondents to New Zealand General Social Survey on 'How do you feel about where you are currently living'. Originally, five categories were defined for the answers to this question. In the present study, however, I aggregated the five categories into two groups: satisfied and dissatisfied.

2.3 Factors affecting residential satisfaction

In New Zealand, Morrison (2011) considered the impact of cities on subjective well-being (SWB), happiness and quality of life (three dependent variables). He argued that the differences in outcomes amongst regions derived from differences in people's evaluations are influenced by their living environments. For example, people in denser cities may have a different definition of household crowding than those in rural areas. Morrison took a number of objective and subjective explanatory variables into account, including individual attributes, social capital and accessibility. However, his data precluded him from including density and crowding variables in his models. Morrison's results indicate that city effects are more significant for more material-based well-being outcomes, such as quality of life, than for less material-based well-being measures, such as happiness. This may be caused by residents of denser cities having a higher chance of employment than residents of less-dense cities. Morrison concludes that large-city dwellers are less satisfied than smaller-city dwellers.

2.3.1 Population density, perceived crowding and crowding

The importance of population density is evident for several outcomes associated with increasing returns to agglomeration, such as the rate of invention and growth (Carlino, Chatterjee, & Hunt, 2007; Glaeser & Kahn, 2001). Transit-oriented development studies and welfare studies have shown the benefits of a mixed-use urban development, such that people can walk from their places of work to the transportation system (see, for example, Frank, Kershaw, Chapman, Campbell, and Swinkels (2014)). Compact city development studies also seek a reduction in car usage by designing denser cities with mixed land use.[13] In development cost-benefit studies, the congestion associated with the wide spread of cities is counted as a cost. Based on energy economics studies, for example, sprawled development poses a high energy transmission cost on households and households living in sprawling areas are more likely to live in large, detached single-family houses than residents of compact areas (Ewing & Rong, 2008). However, when viewed across economic,

[13] The mixed land use design is an integration of residential, commercial, cultural and industrial uses, where facilities are located nearby people's homes.

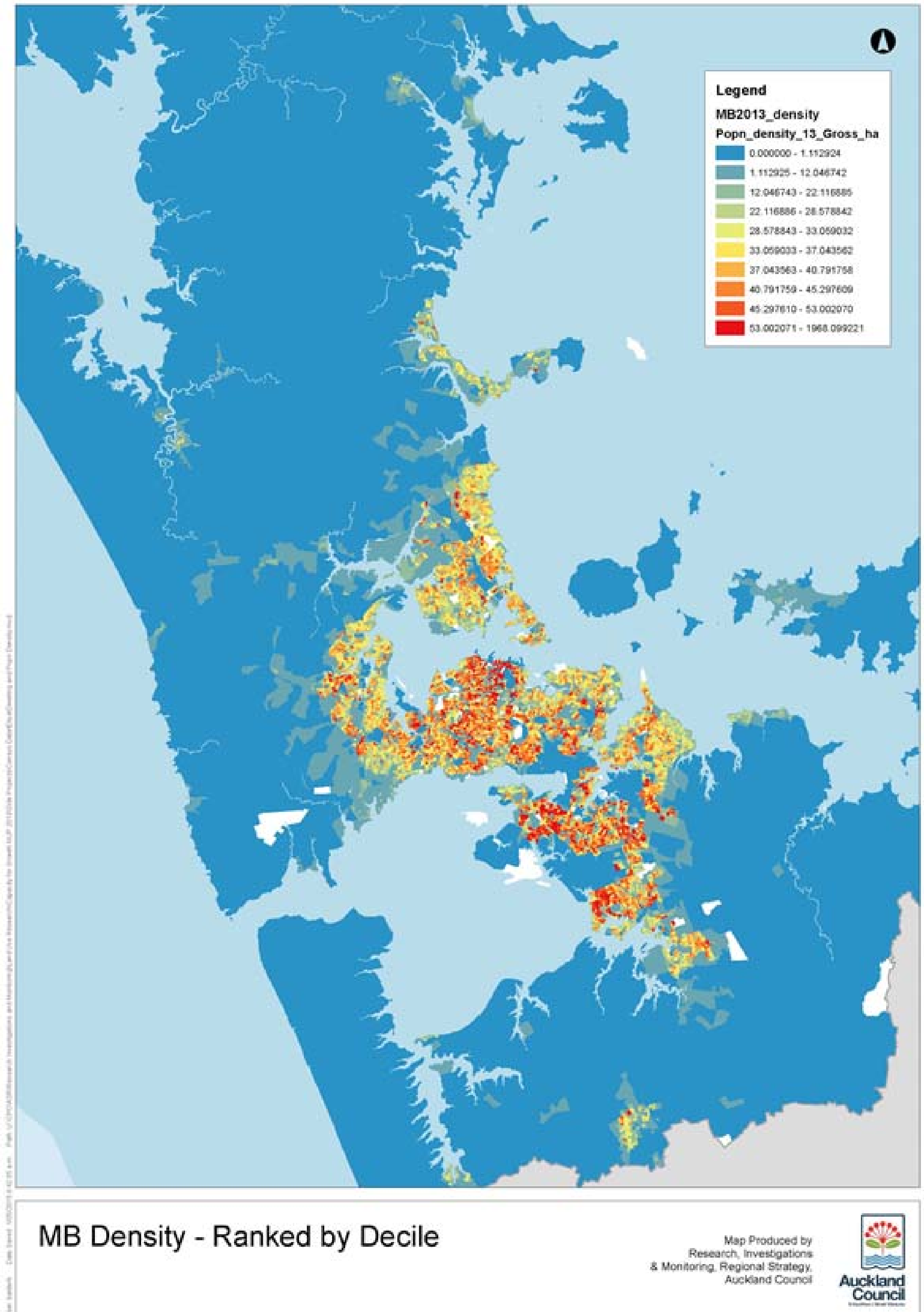
social and ecological sustainability criteria, a sustainable city is not necessarily a dense one (Boyko & Cooper, 2011)).

As Churchman (1999) noted, density may affect people's lives from different perspectives. Examples of aspects that may be affected by density include sustainable development, urban form, preferred building design, social issues and values, stress, cognitive and perceptual processes. Depending on the field and purpose of a study, different measures may be used. Thus, the definition of density needs to be clarified and comparisons across measures can be made. In addition, the difference between density and crowding measures needs to be explained.

Two conventionally used measures of density are population density (also known as area density), and residential density. As shown in equation (1), population density is defined as the ratio of the number of people living in an area to the unit of area. Thus, area density is an objective, quantitative measure.

$$\text{Area density} = \frac{\text{Population size}}{\text{unit of area}} \tag{1}$$

Figure 5. Population density calculated based on neighbourhood land area.

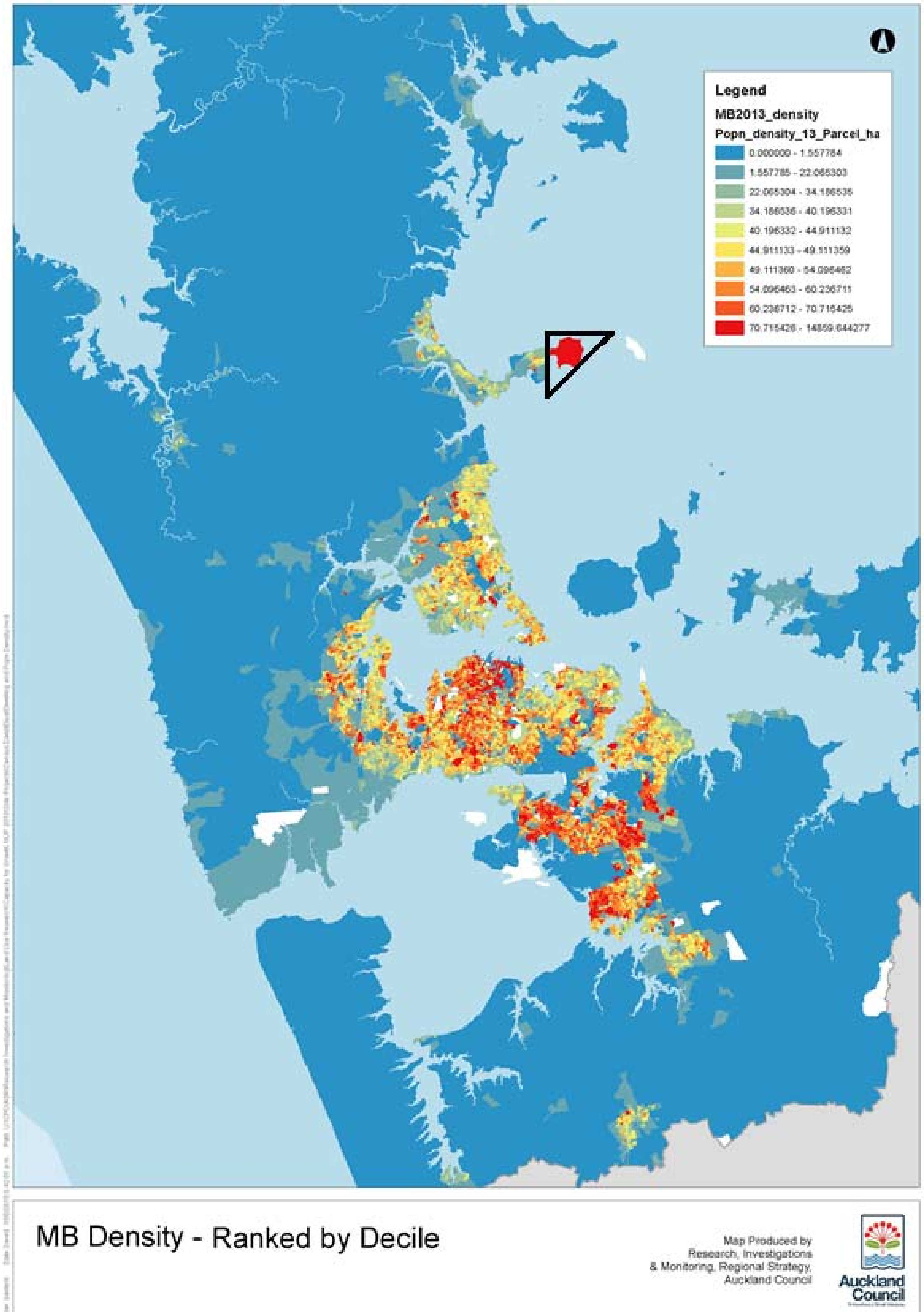


*** As demonstrated in the legend, the range of derived density values is categorised into deciles. For example, almost 45–53 people live in one hectare of an area shaded by dark orange, which is the colour of the ninth decile. (Source: RIMU, Auckland Council)**

The denominator of density measures can be varied, for example to include or exclude parks and roads. The distribution of population density in Auckland calculated, including all area types, is shown in Figure 5, while the density calculated excluding parks and roads[14] is depicted in Figure 6. A comparison between these figures shows a major difference in the area depicted by red on the top of Figure 6, where 'Shakespeare regional park' is located (This is the area shown by a dashed triangle). This highlights the large differences between the values of the density measure depending on the definitions used for its construction.

[14] The density measure is constructed based on the land available for construction (developable land), which excludes lands used for parks and roads.

Figure 6. Population density calculated based on the developable land area.



(Source: RIMU, Auckland Council)

If the numerator consists of the number of people, as it is in equation (1), it derives population density, also known as area density. If the numerator consists of the number of dwellings[15], it derives residential density, as it is shown in equation (2). The unit of area differs amongst regions based on the common standard in use, for example it can be km^2, mi^2.

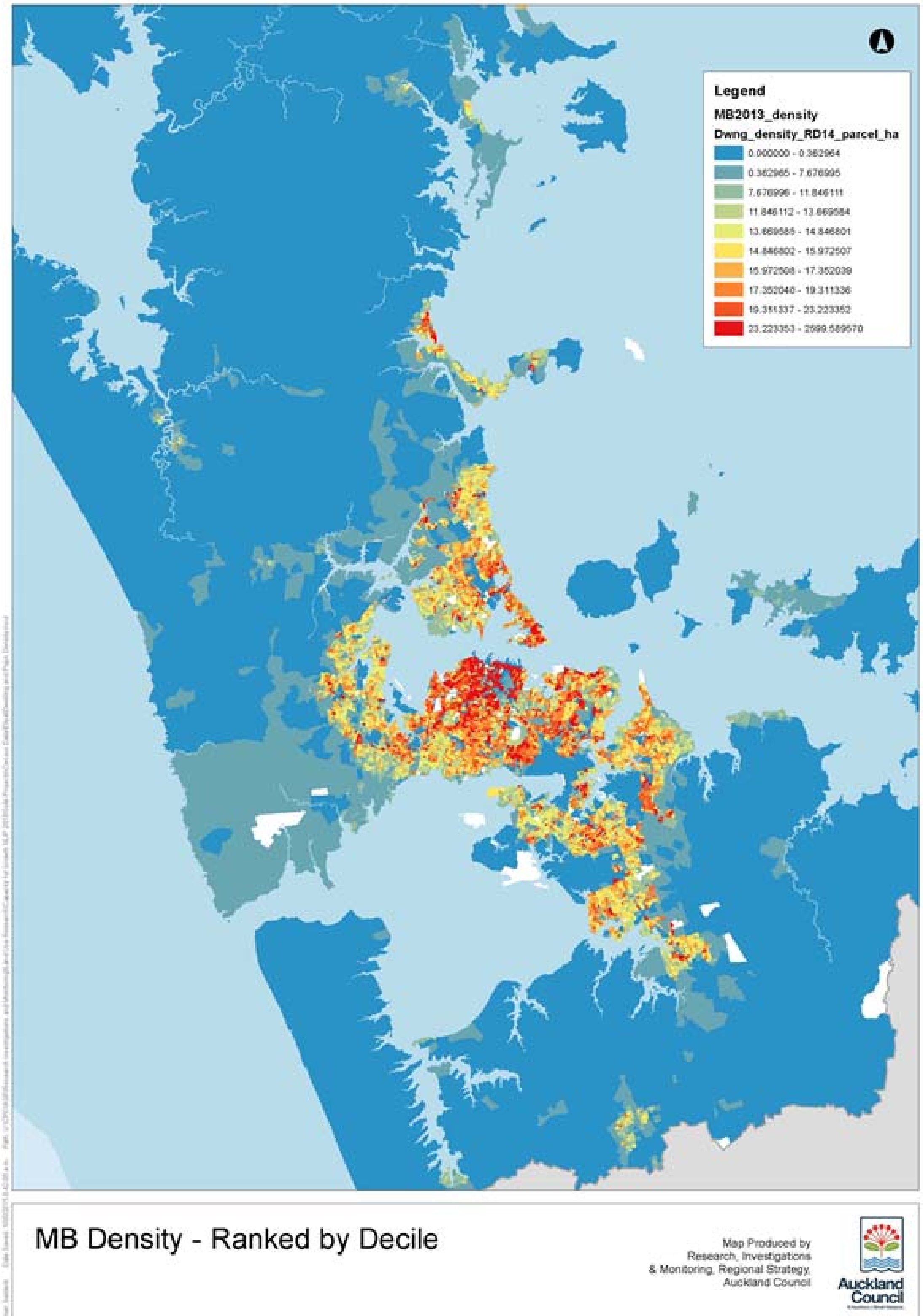
$$\text{Residential density} = \frac{\text{Number of dwellings in an area}}{\text{unit of area}} \tag{1}$$

A map of the distribution of residential density calculated based on the developable land measure is depicted in Figure 7. This map, compared to Figure 6, demonstrates a higher residential density in the central business district (CBD), emphasizing the importance of choosing the right measure depending on which concept is required.

In addition to the complexities of the definitions of density measures, the definition of neighbourhood needs to be clarified. Neighbourhoods may be defined flexibly based on a continuous data (e.g., Duranton & Overman, 2005), or based on administration boundaries (e.g., Krupka, 2007). This point will be discussed from a theoretical perspective in Section 2.4, and from a practical perspective in Section 3.2.

[15] A dwelling is any building or structure used by one or more households as a home in a permanent or temporary nature. This includes houses, apartments, motels hotels, prisons, tents, huts and motor homes. For example, an apartment building consists of many apartments (units), each of which is counted as a separate dwelling.

Figure 7. Residential (dwelling) density calculated based on the developable land area.



(Source: RIMU, Auckland Council.)

There are two main sources of concern with the definition of density. First, density measures suffer from aggregate measures' pitfalls; for example, a high population density of a city may be misleading if there is a large variation in the population densities of its neighbourhoods. Such a problem can occur on smaller scales as well; for example, household crowding may be high in a neighbourhood with a large variation in crowding among households living in that neighbourhood. A second problem with density measures stems from the form of amenity distribution in areas. For example, given equal density measures of two neighbourhoods, green spaces may be scattered in one of them but concentrated in another, which means that the latter residents do not have an equal opportunity to benefit from green spaces. The absence of these drawbacks in subjective measures may lead us to conclude that they are superior measures to the objective measures (however, taking the spatial location of dwellings and the number of neighbours they have into account does address the second problem to some extent).

In addition to the objective measures of household crowding, there are a few subjective measures of it. One subjective concept is perceived crowding, which is a gauge of individuals' evaluation of the crowding of their dwelling. A range of crowding measures have been defined at different quantitative levels; some are more quantitative/objective in nature and some are more qualitative/subjective. The factors taken into account in these measures depend on a population group's understanding of what amount of space is considered sufficient for an individual to satisfy his or her needs, especially his or her privacy. A more comprehensive discussion of the crowding measures is provided in Section 6.2.1, which compares various (objective and subjective) measures of crowding in terms of their predictive power for residential satisfaction.

2.3.2 Homeownership

Due to the importance of homeownership in urban economic literature, this section briefly highlights the relevant results from homeownership studies. For a more detailed review, see Appendix ii.

Homeowners are more involved in the loss and gains of their dwellings than renters (e.g., Hilber 2010). Becoming an owner affects both the absolute and relative situation of an individual and it may affect socioeconomic outcomes. For example, people who become owners feel more committed to their community. This affects social capital formation positively, which leads to higher life satisfaction (Bjørnskov, 2003; Roskruge, Grimes, McCann, & Poot, 2013).

According to Statistics New Zealand (2011), renters are twice as likely as owners to have major problems with their residential environment. There are a number of potential reasons for this. For example, given that homeowners have chosen and paid for their house, they are less likely to state major problems, or low satisfaction, with their dwelling. Another reason may be the variety of choices available to homeowners who plan for a long-term stay. Also, owners usually spend more on their dwelling in order to make it more enjoyable for them. As homeowners put a lot of time into finding an appropriate dwelling, and because homeownership is associated with a large fixed cost, they end up with a better choice of living environment than renters.

The effect of tenure status on different socioeconomic outcomes has been discussed in a wide range of economics papers. In a comprehensive study, Dietz and Haurin (2003) reviewed the association between homeownership and a range of individual socioeconomic outcomes: saving behaviour, childbearing decisions, relatives' gift-giving behaviour, wealth accumulation, portfolio choices, mobility, labour force behaviour, housing maintenance, urban form, social activities and crime, political activities, child outcomes, residents' health, demographics and perceptions.

In the United States, housing investment represents 61 per cent of the net wealth of both African-American and Hispanic households (McCarthy, Van Zandt, & Rohe, 2001). This large proportion makes people more sensitive to the socioeconomic conditions of their dwelling, as any change in it may lead to a gain or a loss. Consequently, owners may change their behaviours relative to their renter peers. For example, Helderman, Mulder and Ham (2004) showed a negative correlation between homeownership and mobility, which may be the result of high transaction costs for owners (DiPasquale & Glaeser, 1999). However, apart from high transaction costs, higher satisfaction among homeowners with their dwelling and neighbourhood has been presented as another reason for moving less than renters do (Kearns & Parkes, 2003), which stems from their location choice by the time they choose their dwelling.

As some studies have mentioned, homeownership significantly increases residential satisfaction as it provides more choices for occupants, such as the choice of neighbourhood and house quality (Diaz-Serrano, 2009). However, the extent of the satisfaction improvement depends on people's evaluation of the ownership value, which derives from housing affordability. Another issue about the measured influence of tenure status relates to the behavioural impact of assessing satisfaction after paying for a house. As Dekker et al. (2011) mentioned, people are not generally keen to state low satisfaction with a possession that they have paid a lot for, while renters are more honest in their assessments as they do not have a sense of ownership of their dwelling.

Summary

It is important to have a clear understanding of the definition of the variables of interest in the analysis. Two main variables of interest in this thesis are household crowding and density, both of which can be measured using different definitions. Area density and residential density are two measures of density that will be used in the current study. Perceived crowding is introduced as a subjective measure of household crowding. The other crowding measures will be introduced in Section 6.2.1.

Another important factor of residents' satisfaction is homeownership. Owners may report higher levels of residential satisfaction. A number of reasons have been identified for the higher satisfaction of owners; for example, owners put more effort into finding an appropriate place than renters, and owners have a wider range of choices available to them.

2.4 Neighbourhood effects

Neighbourhood effects have been a controversial issue within urban studies over the last two decades. The externalities involved in living in a neighbourhood with specific characteristics, such as poor or rich neighbours, are called neighbourhood effects. With the introduction of spatial effects in economics, causality issues become a major source of concern, with some studies challenging the possibility of deriving any reliable result in spatial studies. However, reliable results can still be obtained in the presence of appropriate assumptions and by employing appropriate econometric methods. In a study on the multiplier effect of the area effects, Galster and Hedman (2013) compared the results derived from multiple methods, including fixed-effects panel model, random-effects panel model, fixed-effects panel model with instrumental variables, fixed-effects panel model for a sub-sample, and ordinary least squares. While they found considerable sensitivity of results to the statistical approach taken, they also found significant neighbourhood effects, regardless of what approach was taken.

In order to understand the effect of neighbourhoods we should first have a clear definition of 'neighbourhood', which has been defined in various ways. In earlier urban economic studies, it was common to define neighbourhood as a geography that is limited by boundaries defined by administrative units of the government. However, this definition does not account for individuals' perceptions of their neighbourhoods' boundaries, so the results of these studies may be biased. A more recent approach favours bespoke neighbourhoods, which are defined based on the mutual characteristics of the residents. The latter definition is based on the assumption that neighbourhoods are broadly homogeneous regions. The appropriate definition of the neighbourhood depends on the purpose of a study. A study that seeks boundary-oriented results should follow the administrative boundary definitions, while a study on residents' behaviours may follow the bespoke definition. The current study's definition of neighbourhoods' boundaries is presented in Section 7.4.2.

The modifiable areal unit problem (MAUP) is a source of concern in testing statistical hypotheses when researchers aggregate a spatial phenomenon into districts. When selecting a measure, it is important that researchers avoid personal judgments. The measures of intensification may be defined at different geographic scales, such as very small suburbs or large neighbourhoods. These measures also vary depending on their construction method. For example, while both residential density and area density measure the level of density in a neighbourhood, their different definition may lead to different variations in their values. Hence, any measure used in the subsequent chapters will be tested with different definitions for neighbourhood boundaries and will be constructed by using different methods.

Durlauf (2004) introduced two main reasons for the recent interest of economists in studying the effect of neighbourhoods on their residents. One reason is based on recent advanced methods in economic theory (for more details see Manski, 2000) and the other is the importance of a comprehensive approach to the persistence of poor conditions amongst poor people over long time periods; in other words, poverty traps. This happens when the choices of peer groups affect other members of the group. The poverty effects derived from the social interactions amongst peer groups are influenced by three main factors. The first factor, which relates to the present study, is the psychological factor, whereby people often compare themselves with peer groups. The second and third factors are the dependence of behavioural costs on others' behaviours and the dependence of people's behaviours in the future on information derived from peers' past behaviours.

Empirically, many studies have analysed the impact of geographical inequalities based on health outcomes. The results demonstrate an increase in health inequalities amongst residents of different neighbourhoods in New Zealand between 1980 and 2001 (Pearce & Dorling, 2006) and in the UK and the US (Shaw, 1999; Singh & Siahpush, 2006). Geographical inequalities derive from different spatial effects. Types of spatial effects can be categorised into service-oriented effects such as educational attainment; physical effects such as housing and amenity quality; social effects such as socialization; environmental effects, including residents' attitudes; locational impacts such as labour and housing markets; and concentration effects such as poverty traps. For example, based on a study of eight European countries, higher levels of green spaces and lower levels of incivility in the living environment (such as graffiti and litter) are associated with higher levels of physical activity and, hence, lower levels of obesity (Ellaway, Macintyre, & Bonnefoy, 2005). In an individual-level study on geographical effects in the UK, Atkinson and Kintrea (2001) compared living in a socially deprived area with living in a non-deprived area. They concluded that living in a geographically concentrated deprived area is generally more problematic for residents. However, as noted by Cheshire, Nathan, and Overman (2014), the causal impact of residential segregation on the residents' outcomes, such as the residents' health status, is extremely difficult to identify. The reason for the complexity of the causal inference is the difficulty in controlling for the other factors that affect people's location choice, which depends widely on their life chances.[16]

The topology of a neighbourhood indicates who is communicating with whom, which is important to the analysis of social interactions. The direct connection between individuals derives a network with a specific topology, which depends on the shape of direct connections. Different shapes vary depending on the number of direct connections between individuals and the distribution of the population. The present study was inspired by the random graph theory (Erdős & Rényi, 1960) in order to cover the most plausible network structure derived from the probabilities of direct connections amongst individuals. Consequently, any independently distributed direct connection between agents i and j exists with a probability of p(I). If this probability has a high dependency on the population size (N), there would be a linear relationship between the number of direct connections and the population size. In this context, neighbourhoods comprise agents who are directly or indirectly connected to one another. Palmer (1985) discussed the

[16] Factors of location choice will be discussed further in Section 7.2.1.

relationship between neighbourhood size and direct connection probabilities. He argued that if p(I) exceeds the inverse of the population size (1/N), the number of connections is expected to be more than N. Consequently, the neighbourhood size and the topology of the neighbourhood change if the probability changes.[17]

The aggregate and individual effects of neighbourhood was mentioned by Manski (1993) and developed by Durlauf (2004). The neighbourhood effect may be of two types: firstly, its impact on individuals and consequently on neighbourhood outcomes; and secondly, its impact on neighbourhood membership. In order to study the influences, they made a distinction between neighbourhood effects derived from individuals' characteristics and those from members' mutual behaviour. The individuals' choices are then analysed by assuming that a person is more intensely affected by his/her beliefs about others' behaviours rather than by the other persons' actual behaviours. The prominent role of complementarity between choices of individuals is emphasised. The complementarity enables multiple equilibria to exist, which justifies differences in the aggregate behaviour of neighbourhoods with otherwise similar observations (for more details, see Durlauf (2004)). The second impact of neighbourhood effects will be considered in Section 2.4.4.

2.4.1 The average neighbourhood approach

In the present study, I have assumed that an individual's behaviour is affected by the average choice of neighbours. This is consistent with Brock and Durlauf (2002), who assumed the presence of the term Jm_nw_i in agent i's payoff function V_i, where m_n is the expected average choice in the neighbourhood, w_i is the individual i's choice derived from a set of possible behaviours (Ω_i), and J is the behavioral parameter; that is, J is the coefficient of the term m_nw_i in agent i's payoff function. Jm_nw_i also accounts for the presence of complementarity, as the second derivative of V_i with respect to m_n and w_i is equal to J,

$$\frac{\partial^2 V_i}{\partial m_n \, \partial w_i} = J, m_n = I^{-1}\Sigma_{i=1}^{I} E\left(w_i \middle| Y_n, X_j \forall j\right)$$

In this specification, for instance, an increase in the average crowding of a neighbourhood affects the residential satisfaction of a resident by J multiplies by his/her own dwelling's crowding. The effect of average choice of neighbours on the individual's behaviour leads to multiple equilibria derived from the heterogeneity caused by the variation of individuals' characteristics (X_j) across individuals and of neighbourhood characteristics (Y_n) across neighbourhoods. This average choice of neighbours approach is in line with the random graph theory approach in terms of justifying how individuals are connected to and affected by one another.

[17] The population size here is synonymous to the number of vertices (or the order of graph) in discrete mathematics literature.

Alternatively, Glaeser, Sacerdote and Scheinkman (1996) assumed that individuals are affected by their nearest neighbour rather than by their neighbours' average choice. Those authors assumed the presence of the term $J_{A(i)}w_iw_{i-1}$ in the agent's payoff function, where w_{i-1} is the behaviour of the neighbour who lives immediately to the left of the individual i. Although this assumption leads to the coverage of all the population of the neighbourhood by connecting neighbours to one another, it does not include a reciprocal relationship between neighbours; that is, individual i's behaviour does affect individual i-1's behaviour but not vice versa. This assumption limits the heterogeneity to three types of agents (A_i) randomly distributed amongst neighbours: the agent that always chooses independently a crowded dwelling, $w_i=1$, or a less crowded one, $w_i=-1$, and the agent that always chooses the same as his/her neighbour, $w_i= w_{i-1}$. This model derives a unique equilibrium, rather than the multiple equilibria in Brock and Durlauf's approach. By considering the variance of the normalised sample average,[18] it is shown that the model suggested by Glaeser, Sacerdote and Scheinkman has a higher variance of the average observed behaviours compared to the model suggested by Brock and Durlauf.

The lack of a reciprocal relationship between neighbours is a shortcoming of Glaeser et al.'s approach; accordingly, I have chosen to use the average approach, which incorporates reciprocity and complementarity.

2.4.2 Application of the average approach

In the current study, the following payoff function is assumed with respect to the impact of household crowding on residential satisfaction:

$$RS_i = (1-\alpha)\left(C_i + \beta C_i^2\right) + \alpha(C_i - \overline{C_n}) + \delta_i,\ \delta_i = \omega_i + f(X_i) \tag{2}$$

Where, RS_i is the residential satisfaction of individual i; C_i represents the household crowding of individual i; $\overline{C_n}$ is the average crowding of individual i's neighbourhood and δ_i is a gauge of the individual's impression of crowding. Thus, δ_i is a function of the individual's characteristics, X_i, and ω_i, which has a normal distribution, $N(0, \sigma_i^2)$. The impact of the second part of the equation, $(C_i - \overline{C_n})$, on residential satisfaction is gauged by the parameter α, which can be considered as a social effect. Thus, $1-\alpha$ is a measure of individual effects, $\left(C_i + \beta C_i^2\right)$.

2.4.3 Individual and household socio-demographic characteristics

[18] The distribution of the normalised sample average is:

$$(2n+1)^{-\frac{1}{2}}\sum w_i \sim wN\left(0, \frac{p(1-p)(2-\pi)}{\pi}\right), \qquad \text{where } p = \frac{p_1}{p_1+p_2} \text{ and } \pi = p_1 + p_2$$

Here, agent A is assumed to be distributed in the neighbourhood by the probability of p_A. Consequently, by assuming an increase in the per centage of type 2 agents, π approaches 0 and the variance approaches infinity.

Besides the density and crowding variables, socioeconomic variables also affect residential satisfaction (Amérigo & Aragonés, 1997; Chapman & Lombard, 2006). For example, Baum, Arthurson, and Rickson (2010) claimed that the level of income, the tenure status and the ethnicity of residents within a neighbourhood all affect neighbourhood satisfaction. In another study, Dekker, de Vos, Musterd, and van Kempen (2011) claimed that housing characteristics do not affect neighbourhood satisfaction (as mentioned in Section 1.1, this is in line with Rodgers, 1982). Dekker et al. found that the characteristics of individuals and their opinions about the living environment have a greater effect on neighbourhood satisfaction than the characteristics of the residential environment. People with higher income levels can afford more expensive houses and usually have fewer residential problems. Hence, household income is a key socioeconomic factor that may affect satisfaction with housing. It also affects the purchasing power of people and, consequently, their tenure status. In a hedonic pricing study, Eichholtz and Lindenthal (2014) investigated the dependence of house prices on age and other demographics, and concluded that demand for housing raises by higher levels of income, education, health status and age. Their results show that demand for housing may still grow in the presence of an aging population and a stagnant household number.

Homeownership raises the satisfaction of people with their dwellings significantly. In addition to the reasons discussed in Section 2.3.2, this positive impact relates to the behavioural impact of assessing satisfaction after paying for a house. As mentioned earlier, people are not keen to state low satisfaction with a possession that they have paid a lot for, while renters are more honest in their assesments as they do not have that ownership feeling to their house. According to Statistics New Zealand (2011), renters are twice as likely as owners to claim that they have major problems with their residential environment. This can be justified by renters' expectations of their current living environment. Renters are more sensitive to their living environment; that is, as soon as they do feel dissatisfied with their living environment, they have the opportunity to settle into another dwelling, which is not the case for owners. Homeownership also affects social capital formation (Roskruge et al., 2013), which leads to higher life satisfaction (Bjørnskov, 2003).

One of the factors that affects residential satisfaction is the extent to which individuals use their dwelling and neighbourhood environment. People who are more limited to staying within their living environment, such as the elderly, are more likely to have negative attitudes towards their neighbourhood (Dekker et al., 2011; Guest & Wierzbicki, 1999). Education level, as a socio-economic factor, usually leads to a more expanded social network and, consequently, a higher use of the urban environment. Ethnicity is another important factor in the use of the neighbourhood environment. People with the same ethnicity are generally more connected to each other. As a result, immigrant groups and ethnic minorities usually live in more concentrated areas and are therefore more limited to their own neighbourhood (Baum et al., 2010; Mare & Coleman, 2011; Wang & Maani, 2012).

With regard to individuals' characteristics, the effect of household's composition on individuals' satisfaction needs to be considered. For example, the probability of suffering from a small house is higher for a larger household than a smaller one. Consequently, parents without children are more likely to report higher residential satisfaction than those with children, especially since the latter have less mobility if they do not feel satisfied with their place (Brodsky, O'Campo, & Aronson, 1999). On the other hand, some studies have emphasised the intermediary role of children in generating social interactions (Guest & Wierzbicki, 1999). Another factor that affects satisfaction is age, which shows a positive correlation with satisfaction levels (see, for example, Lu, 1999). This might be a matter of time, as older people have probably lived in a house for a longer time and already adapted to probable problems (Dekker et al., 2011).

2.4.4 The identification of the causal relationship

The causal relationship amongst empirical studies of neighbourhood effects is a source of concern (Cheshire et al., 2014; Durlauf, 2004; Manski, 1993; Moffitt, 2001). The direction of causation has been the main source of concern in economic studies since 2000. As Moffitt (2001) mentioned, many neighbourhood effects studies suffer from identification bias or, in some cases, from the lack of a valid instrumental variable. He distinguished three main problems that usually arise in the estimation of the effect of group characteristics on individual characteristics: the simultaneity problem, the correlated unobserved problem and the endogenous membership problem (the first two are based on Manski, 1993). The simultaneity problem, which is also called the 'reflection problem', occurs if someone's behaviour affects another person's behaviour. Hence, these two people's behaviours should be regressed on one another, which means a simultaneous equations problem as developed further below.

If we assume n = 1, …, N neighbourhoods and, for simplicity, two individuals i = 1,2 living in each neighbourhood, then the individual's satisfaction is y_{in}, the individual's characteristics is x_{in} and the error term ε_{in} includes unobserved variables. So, the linear relationship is as follows:[19]

$$y_{1n} = \alpha_0 + \alpha_1 . x_{1n} + \alpha_2 . x_{2n} + \alpha_3 . y_{2n} + \varepsilon_{1n}, \qquad E(\varepsilon_{1n}|x_{1n}) = 0 \tag{3}$$

$$y_{2n} = \alpha_0 + \alpha_1 . x_{2n} + \alpha_2 . x_{1n} + \alpha_3 . y_{1n} + \varepsilon_{2n}, \qquad E(\varepsilon_{2n}|x_{2n}) = 0$$

In equation (3, α_2 represents the exogenous social interaction and α_3 shows the endogenous interaction. As a conventional linear model, the error term $(\varepsilon_{1n}, \varepsilon_{2n})$ is assumed to be independent of the explanatory variable (x_{1n}, x_{2n}). As in this simultaneous equation, it is not possible to exclude at least one exogenous variable, the identification of parameters is not possible. By comparing equations (2 and (3, α_3 is not a parameter of interest in this study; that is, we assume that while other residents' situations (x_{in}) affect each other's residential satisfaction, those other residents' own satisfaction does not impact on each other. As Moffit mentions, if α_3 equals zero and if ε_{1n} and ε_{2n} are independent, then identification is possible as the covariance of the residual values of different individuals in a group conditional on x_{1n} and x_{2n} is equal to

[19] This equation is very similar to Moffit's, except with some changes in notation to make it more compatible with this study's literature.

zero. Hence, the first concern is whether or not endogenous social interactions exist. Therefore, I should test the sensitivity of results to the possible effect of one's residential satisfaction on others' residential satisfaction. One way to do this is to use a spatial approach, randomly choosing subsets of the sample with a maximum one person per mesh-block (neighbourhood). It is then possible to compare the results of this limited estimation with the full model to see whether the results differ significantly; if they do not, endogenous social interactions can be considered to be immaterial. This solution is based on the assumption that social interactions, if they exist, are only within the mesh-block (MB). For more details on this approach, see Moffitt (2001).

The correlated unobserved problem arises if some characteristics of the group (neighbourhood) are not observed in the model whilst they are correlated with included variables. The problem occurs if we have a specific neighbourhood component in the error term (μ_n) that is correlated with x_{in}. Thus, E(x_{in} μ_n) is not equal to zero, which means inconsistency in the estimation of α_1 and α_2. As Moffit mentioned, this correlation might be the result of endogenous group membership (discussed later) or of omitted qualitative or quantitative variables in the model. To mitigate this problem, based on previous studies described in Section 2.3, I will include a large number of relevant variables within our rich dataset, in the form of control variables. Of course, there might be some variables that are not included in the model, but if they are not correlated with any of the variables of interest the unobserved neighbourhood effects problem is not a source of concern. One reason for taking this approach is the inability to include fixed effects in a cross-section model.

The endogenous group membership problem arises in the neighbourhood self-selection process. The severity of this problem is partly a matter of locational adjustment speed compared to the speed of social interactions formation. Neighbourhood selection is not a random process but is a process that depends on observed and unobserved individual characteristics. The unobserved individual characteristics are partly correlated with location choice and partly with the dependent variable, which might lead to endogeneity. In the model framework, it is assumed that individual i gains some utility from locating in region n. This utility is conditional on others' locational decisions; that is, people might prefer to live in a neighbourhood with neighbours who are similar to themselves. Therefore,

$U_{1n}=f(x_{1n},\varepsilon_{1n},x_{2n},\varepsilon_{2n})+\omega_{1n}$, where ω_{1n} is the intercept.

A reasonable location decision rule is that individual 1 chooses neighbourhood 'a' if the utility of locating in that neighbourhood is higher than the utility of locating in any other neighbourhood 'b' (Glaeser & Gottlieb, 2009). This utility function represents a relationship between error terms (ε_{1n},ε_{2n}), which is observable by individual 1 and 2 but not by econometricians, and individuals' characteristics (x_{1n},x_{2n}). This leads to the inconsistent estimation of parameters. This problem can be solved by using instrumental variables. However, as Van Ham and Feijten (2008) mentioned, it is difficult to identify a valid instrument because of the complex relationship between the behaviour of the housing market and neighbourhood characteristics. A conventional solution is to ignore the locational adjustment when its speed is slower than the speed of social interactions formation. In this study, I have used a proxy for the locational decision. In a study on the factors that affect Aucklanders' location choice, Mare and Coleman (2011) emphasised the effect of 'own group' attraction, such as ethnicity, income and the country of birth. Consequently, group membership is controlled for by including these variables as control variables. This point is discussed further in Section 7.2.1.

Other strategies may also be useful to overcome the identification problem. Since income affects both location choice and the residential environment, one strategy is to assess the effect of the variable of interest after splitting the sample population into higher and lower income groups and test whether the parameters of interest are stable across the split samples.

Another common solution to identification problems proposed by Brock and Durlauf (2004) is partial identification (Manski, 1990). This approach recognises that there is a trade-off between the credibility and the imposition of certain assumptions. For example, in the current study I could assert a strong assumption that individuals do not affect one another but this is not very realistic. To avoid such a credibility problem, it is better to adopt a number of weaker but more credible assumptions (Manski & Pepper, 2013). This approach does not derive the point-identified treatment effects but it does lead to a partial identification.

Summary

Neighbourhood effects are the externalities involved in living in a neighbourhood with specific characteristics. Considering the complexities involved in identifying neighbourhood effects, a careful understanding of these effects needs to account for a range of issues. A spatial analysis needs to provide a correct definition of neighbourhood boundaries. Special analyses often aggregate a measure at a neighbourhood level, which may lead to problems faced when using average measures. Therefore, the choice of the average measure and its impact on the results need to be justified carefully. Causal inference means that we need to identify the direction of causation. It is difficult to measure the causal impact of residential segregation on residents' satisfaction because of the other factors that affect individuals' location choice, and the estimation methodology needs to be of an appropriate form in order to overcome the endogeneity problems.

Appendix I. Data for the projected population for Auckland

Table 1. Projected population for Auckland, 2013 (base)–2043 (released Feb 2015)

Year at 30 June	Low series	Medium series	High series
2013	1,493,200	1,493,200	1,493,200
2018	1,622,500	1,646,500	1,670,100
2023	1,713,900	1,767,500	1,820,700
2028	1,803,700	1,890,900	1,977,200
2033	1,887,200	2,010,500	2,133,100
2038	1,961,700	2,123,000	2,284,600
2043	2,028,300	2,229,300	2,432,800

Note: The 2013 count is the estimated population count. Source: RIMU; Statistics New Zealand.

Appendix II. A review of homeownership studies

An important behavioural outcome of homeownership is the change in wealth accumulation behaviour. Households usually start to save some money for a deposit before purchasing a dwelling, to the extent that they reduce their consumption. The magnitude of the positive effect of higher house prices on savings is even larger for younger households (Sheiner, 1995). However, they may change their behaviour after purchase by increasing their non-housing consumption if house values (and hence their equity) are rising (Engelhardt, 1996a). Conversely, a higher house price may dissuade some people from becoming owners and lead them to a higher consumption level and a lower saving level (Yoshikawa & Ohtaka, 1989). Savings are also affected by the risk of earnings, and a higher risk of earning has been introduced as a reason for the decrease in homeownership among younger people after 1980 (Fisher & Gervais, 2011).

Prior to becoming owners, people usually try to find stable jobs in order to save more for the upcoming costs of homeownership. After becoming an owner, they may increase their labour supply in order to cover mortgage payments. This is usually followed by a higher job seeking rate among men and, more interestingly, married women during their families' ownership period. Consequently, some studies have attributed higher rates of unemployment amongst industrialised countries in the 1990s to higher rates of homeownership (Oswald, 1996). The new owners cannot tolerate long job-seeking durations due to their monthly mortgage payments; this short-term job seeking may not affect unemployment rates significantly (Green & Hendershott, 2001). Also, some studies have shown a higher level of house work as a result of feeling greater responsibility toward a new home. The reduced mobility of homeowners has been argued to be a source of reduced flexibility among them in response to a labour market supply shock, which may lead to lower earnings (Winkler, 2011). This is in line with a study by De Graaff and Van Leuvensteijn

(2011) that introduced the higher transaction costs of homeowners as a reason for their higher probability of becoming unemployed.

It is indicated that homeowners generally take more care of their dwellings than renters, perhaps as a result of making more money by renting their property at higher prices. Also, homeowners may initially be keen to take better care of their residential environment than renters, and this could be the reason why they put more effort into owning a home. Owners paying higher attention to their homes could be also a matter of spending more time in their own homes and, probably as a result of the less mobility of owners, a longer period of living in their current residence. However, this reasoning may be criticised by adaption effect; that is, as owners live in their dwelling for longer periods, they may adapt to plausible problems and not find it necessary to spend money on home maintenance. A number of papers have discussed the insignificance of homeownership effect on home maintenance after controlling for housing and neighbourhood structures (Dıaz & Luengo-Prado, 2010). For example, Ioannides (2002) found that one neighbour's home maintenance has a strong effect on other neighbours, which supports the importance of including neighbourhood effects in socioeconomic urban studies.

Homeownership also affects children's outcomes. Compared to renters, homeowners are more involved in the society, and hence they know better schools to educate their children. Furthermore, homeowners have wider social networks in their neighbourhood than renters, which lets them to have more control over their children's friendships. The reduced mobility of homeowners enables their children to settle well into neighbourhoods and, in terms of relationships and activities, make better choices in their life, which can have outcomes such as higher educational achievements. For example, Boehm and Schlottmann (1999) claimed that children of homeowners are more likely to become homeowners in the future, although this may be justified by the financial support of their family. It has been reported that the higher education of parents and the poor financial backgrounds of potential home-buyers are positively correlated with the reliance of first-time home-buyers on their relatives' financial supports (Mayer & Engelhardt, 1996). Also, home-owning parents may generally put greater effort into different aspects of their life, as they have put more efforts into becoming homeowners. Thus, their children will probably enjoy more attention and more financial support in the future. One may worry about the neighbourhood effects on child outcomes. In a recent study, Gibbons, Silva and Weinhardt (2013), controlling for a wide range of spatial effects, found no direct neighbourhood effect on children educational outcomes, but some effects on behavioural outcomes, such as attitudes to school.

The effect of homeownership on health and on the elderly is controversial. Homeowners may be healthier than renters as they know more about their neighbourhood and the quality of the provided health services. The greater location choice among owners compared to renters enables them to locate closer to the health services. Also, as discussed earlier, they take greater care of their dwellings, which may also lead to a better health status. One drawback, as claimed by Nettleton and Burrows (1998), is that mortgages may be a source of stress for owners and may decrease their health status. Elderly people usually have lower monthly incomes than people in some other age groups and are therefore less likely to be eligible for a mortgage. Also, if they have already owned a house, they are more likely to be keen to use reverse mortgages in order to compensate for their income shortages (Hancock, 1998). Both of these effects of homeownership on health and the elderly need further investigation in terms of causality.

The urban form is believed to be affected by homeownership as well. Increased urban sprawl may be stimulated as a result of the tendency of owners to buy bigger houses with bigger sections, which may also affect the communication amongst residents. However, the causality of this argument needs to be examined carefully. A number of studies have assessed the effect of different taxation policies on urban forms. Some of these policies may push homeowners to move away from the city centres, while some may encourage them to buy a more expensive house and other policies may discourage them from buying more than one house. However, the role of homeownership per se on urban form has not been addressed clearly. The urban form may be affected by ethnic groups, as well as discrimination. Van Ham and Feijten (2008) concluded that people in a neighbourhood with a higher proportion of ethnic minorities are more likely to shift away from their current residency environment. In a similar study on Aucklanders' location choice, after controlling for identification issues and by including spatial effects, Mare and Coleman (2011) found that own-group characteristics, including ethnicity, income and the place of birth, have a significant effect on settling choices; for example, ethnic minorities are more likely to settle into neighbourhoods with higher proportions of the same ethnic group. Furthermore, Immergluck (1998) claimed that the homeownership of ethnic minorities leads to a reduction in area segregation. Consequently, these ethnic minorities usually strengthen areas integration, which affects the urban form.

From a portfolio choice point of view, house ownership is both an investment good and a consumption good, which includes a random risk, the risk of a decrease in house rents and also the risk of a decrease in house prices. By assuming that the risk of investing in the house is greater than optimal of the portfolio, Brueckner (1997) claimed that people aiming to decrease their investment in housing should decrease their consumption of housing and enter the rental housing market. However, Engelhardt (1996b) concluded that people are less likely to change their saving behaviours as long as they are enjoying real gains of housing capital. Also, as Sinai and Souleles (2005) mentioned, becoming an owner changes the housing costs risks faced by renters to asset risks faced by owners. They found that the prices of houses and the probability of becoming a homeowner are positively influenced by net rent risk.

Most studies agree that the house price growth is similar amongst low- and high-value houses (Pollakowski, Stegman, & Rohe, 1991), and that the price growth of medium value houses is usually lower. However, neighbourhood composition affects price growth significantly. Based on Grimes and Aitken (2006), a lowering socioeconomic trend leads to a lower or even negative growth of house prices in a neighbourhood. Also, as Mare and Coleman (2011) mentioned, similar characteristics affect location choices (in line with the findings of Flint and Rowlands (2003)). Consequently, a higher proportion of minorities in a neighbourhood may affect the growth of house prices negatively (Quercia, McCarthy, Ryznar, & Can Talen, 2000). Over time, economic conditions influence the growth rate of houses with different prices differently (Mayer, 1993). For example, an economic boom usually results in a faster growth of higher value houses, although with greater volatility, while a more stable economy with a falling interest rate and increased employment usually leads to an increase in the price of lower-value houses (Coleman & Scobie, 2009; Smith & Ho, 1996).

Given that state housing schemes aim to cover people in need of a place to live, many such settlements have been in socially deprived areas. Thus, social aspects of homeownership play an important role in this study. There are two main motivations for homeowners to become involved in political activities (Engelhardt, Eriksen, Gale, & Mills, 2010) such as voting, and social activities, such as connections between neighbours. The first is financial interest, as the political system and social environment may affect the price of their properties. Second, the lower mobility of homeowners makes them more patient about undesirable features of their residential environment; they try to avoid undesirable conditions by, for example, participating in elections (Cox, 1982). After controlling for identification problems, Roskruge, Grimes, McCann, and Poot (2012) found that a higher social infrastructure investment results in the accumulation of social capital. They also found a positive correlation between social infrastructure investment and the variety of social activities for people who participate in social activities, while the correlation between the decision to participate in social activities and social infrastructure investment is negative.

As a result of more social activities, homeowners are more involved in most social events. They are more concerned about their residential environment than renters; for example, by paying attention to recycling (Daneshvary, Daneshvary, & Schwer, 1998). This is agreement amongst studies that homeowners are more likely to vote. However, Gilderbloom and Markham (1995) disagreed with Kingston, Thompson and Eichar (1984) about the tendency of homeowners to vote in support the conservative parties. In a more recent study, Ansell (2013) concluded that homeowners are in favour of the party whose policies are more likely to cause an increase in house prices.

The greater social involvement of homeowners means that they are generally aware of safety issues and are therefore less likely to experience crime (Glaeser & Sacerdote, 1996). However, homeowners have more location choices. Thus, they usually choose to not live in neighbourhoods with relatively high crime rates, which are less likely to produce positive financial returns.

3 Common Descriptive Statistics

Chapter outline

Except for Chapter 4, which only uses the New Zealand General Social Survey (NZGSS) dataset, all of the chapters in this thesis utilise the NZGSS matched with the 2006 Census. Thus, many of the variables in these chapters are the same, especially those derived from the NZGSS. To avoid repeating the common sampling issues, the common context is discussed in the current chapter. The data collection process of the NZGSS dataset is explained in Section 3.1.

Due to computational limitations – given the complicated spatial analysis in Chapters 7, 8 and 9 – the sample is limited to Auckland city. To maintain consistency, the sample used in Chapter 6 is also limited to Auckland. However, to give a sense of the differences between the results derived from the entire sample and the Auckland sample, the results of Chapter 6 are also derived based on the New Zealand sample and presented in the appendix to Chapter 6 (see Appendix i). The common descriptive statistics based on the Auckland sample are presented in Section 3.1.1.

3.1 Data sources

The first dataset I will use is the New Zealand General Social Survey (NZGSS) by Statistics New Zealand. I have used the full file of the NZGSS, i.e. the CURFs provided by SNZ do not contain all the variables that I have used in the current study. The NZGSS was carried out in three series, in April 2008, October 2010 and April 2012.[20] The sample size for each cross-section was approximately 8500, a number that Statistics New Zealand (SNZ) selected to represent the total population. The data were collected by a face-to-face interview in which the interviewer used a laptop. Each household was interviewed once and the average interview time per household was 45 minutes. The survey consists of two separate parts: the household questionnaire (3.1.1) and the personal questionnaire (3.1.2).

The other dataset I used is the New Zealand Census 2006 (also conducted by SNZ), which is an official count of the population and dwellings in New Zealand. This dataset contains data at the mesh-block level, which is the smallest geographical level at which the data is processed by SNZ. I matched the NZGSS with this dataset in order to determine the urban area in which the individuals are located. I explain the matching process in Section 7.4.

3.1.1 The household questionnaire

[20] The questionnaire for these surveys is available at the following links:

- 2008: http://www.stats.govt.nz/~/media/Statistics/browse-categories/people-and-communities/families/general-social-survey/GSSQuestionnaire2008.pdf.
- 2010: http://www.stats.govt.nz/~/media/Statistics/browse-categories/people-and-communities/families/general-social-survey/GSSQuestionnaire2010.pdf.
- 2012: http://www.stats.govt.nz/~/media/Statistics/surveys-and-methods/completing-a-survey/faqs-about-our-surveys/nzgss/gss-questionnaire-2012.pdf.

The first questions of the survey are about members of the household and include sex, age and ethnicity. Based on some eligibility rules, the household members, who are in scope and are available during the interview period, answer these questions. The computer selects one of the members to answer the next part, which is the personal questionnaire.

3.1.2 The personal questionnaire

This questionnaire consists of 14 different topic modules. In this study, in addition to the core personal data, satisfaction with living environment, health, skills and abilities are considered. Residential satisfaction was described earlier in Section 2.2. Health perceptions are derived from a question that asks, 'In general, would you say your health is excellent, very good, good, fair or poor?' and skills perceptions are derived from a question that asks, 'How do you feel about your knowledge, skills and abilities?' Respondents are questioned about their feelings regarding their living environment as at the time of the interview and can choose from among five responses ranging from 'very satisfied' to 'very dissatisfied'. The core personal data are mostly categorical. For example, household income is measured based on the New Zealand Standard Classification of Income Bands 2009, which has 16 categories.

3.1.3 Resampling by using replicate weights

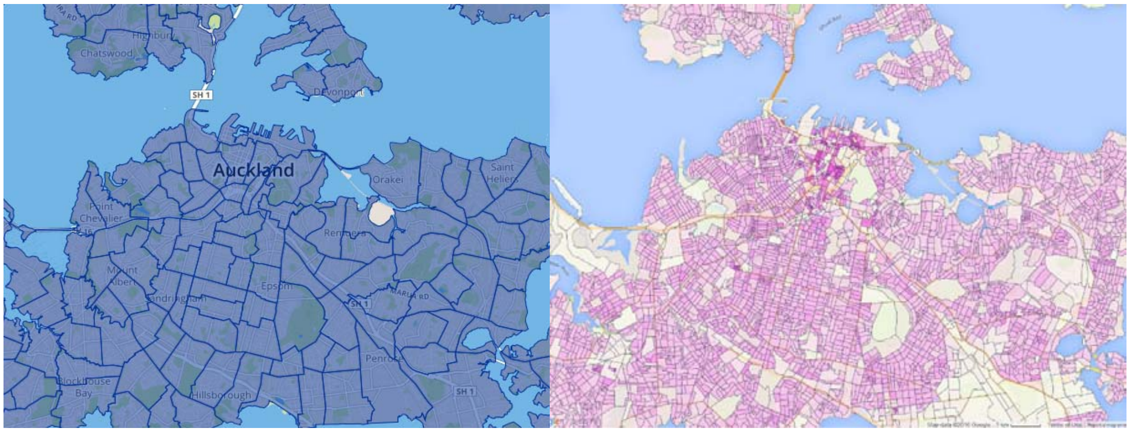
Different methods of resampling may serve to increase the precision of estimation using survey designs by deriving more robust standard errors, proportions, odds ratios and regression coefficients. To do this, a random set of observations is left out at each time of estimation. Replicating this process leads to the estimation of the bias of a statistic; this method is called jackknifing. The NZGSS provides us replicate weights produced by the delete-a-group jackknife method (Kott, 2001). In the dataset, 100 groups are derived by using primary sampling units (PSUs) randomly sorted into each stratum. This strategy results in 100 replicate samples in each, of which one of the groups is omitted and weights are adjusted accordingly. Using these weights in our estimation leads to estimates that tend asymptotically to true values.

3.2 The geographic scales

In order to have a reliable definition of the neighbourhood, we should rely on the resident's perception of their neighbourhood boundaries that can be derived from face-to-face interviews (Witten, McCreanor, & Kearns, 2003) 7.4.2. A typical definition of neighbourhood consists of a certain number of nearest neighbours. This definition is very similar to the mesh-block's definition as an area consisting of dwellings that are located close to each other. Mesh-block is defined by SNZ as a very small neighbourhood, which consists of a few houses located next to each other, with an average population of 50 people.[21] Because I assume that the reference group consists of people who live shoulder-to-shoulder with an individual, the mesh-block design suits our purpose well.

[21] Mesh-blocks were first defined in 1976. Based on the 2006 Census, New Zealand has a total of 41,376 mesh-blocks.

Figure 8. Area units and mesh-blocks.



Source: Torshizian (2017).

In the context of New Zealand, larger neighbourhoods are defined as area units, which approximate city suburbs (as opposed to administrative boundaries). Most people consider themselves as members of their suburb or area unit. Hence, to take perceptions into account, area units may also be considered as a good definition of neighbourhoods. For illustrative purposes, the area units located in central areas of Auckland are depicted on the left-hand side of figure 8. On the right-hand side, the mesh-blocks located in the same area are depicted. As will be discussed in Sections 7.2.1 and 5.4.2, however, individuals' perceptions of their neighbourhoods' boundaries depend heavily on their characteristics, such as their education and income levels. On a larger scale, I consider the boundaries of territorial authorities (TLAs) – that is, local councils – which are administratively determined community of interest that share a common local government. In other words, people with similar ideas and thoughts about a certain topic are considered to locate in the same territory. Similar people may tend to congregate in particular neighbourhoods, and it is an empirical matter as to which neighbourhood definition is most appropriate in this context.

Summary

The datasets used in the current study are three cross-sections of the New Zealand General Social Survey from 2008 to 2012 and Census 2006, both provided by Statistics New Zealand. This provides our study with 8500 observations. In order to derive robust results, I will use the replicate weights produced by the delete-a-group jackknife method. In order to find the correct definition of neighbourhood boundaries, I consider three conventional definitions of boundaries at the mesh-block level, the area-unit level and the territorial authority level.

Mesh-blocks are very small neighbourhoods that consist of a few houses located next to each other, with an average population of 50 people. Area units are aggregates of mesh-blocks and are the suburbs that people usually consider themselves members of. Territorial authorities are very large areas determining communities of interest that share a common local government.

3.3 Descriptive statistics

All variables in Table 2, descriptive statistics, are categorical variables. For example, 'Gender' is a binary variable with 0 for females and 1 for males (in NZGSS, this variable is called 'cordv10', as shown in the brackets). The mean of this variable is equal to 0.485, indicating that 48.5 per cent of respondents are male and the rest are female. On the right-most column, the equality of the mean of each variable between three cross-sections of data; – 2008, 2010 and 2012 – is tested. The significance level of less than 0.001 for the 'Age' variable, which is an indicator of respondents' age, statistically violates the null hypothesis that there is no change in the mean of this variable across the three surveys; in other words, it indicates that the percentage of people distributed in different age groups changes amongst these three cross-sections of data.

To avoid a very long list of variables in the regression tables, some of the variables with similar effects have been recategorised. In the 'Group' column of table 2, the group of each category is shown amongst different variables. For example, for 'Homeownership status', group 'i' includes categories 'Owned, not defined', 'Owned, mortgage' and 'Owned, no mortgage', and so on.[22] Two criteria are used for grouping some categories. The first is the diversity of the category; for example, MELAA ethnicity, which consists of Middle Eastern, Latin American and African ethnicities, is a highly diverse category and is therefore grouped with 'Other'. The second criterion is the results derived from different categories; two categories are grouped if they show very similar patterns in regression models.

In a comparison between three cross-sections, for almost all residential issues, a lower fraction of respondents report problems with their neighbourhoods in 2010. For most housing problems, however, this difference is not statistically significant at the 1 per cent level. Three variables in this table are based on a five-level Likert scale: health perception, ability perception and residential satisfaction. Health and ability perception variables are of interest here to address the potential 'over-optimism' bias, as is discussed in the next section.

[22] The homeownership variable has many categories, including:

- Dwelling not owned by usual resident(s), rental arrangements not further defined.
- Dwelling not owned by usual resident(s), who make rent payments.
- Dwelling not owned by usual resident(s), who do not make rent payments

Considering that the estimated slope for these are not significantly different, the aggregate impact of them reported as the coefficient for renting.

Table 2. Descriptive statistics of variables used from NZGSS.

Variables	Description	Group	Mean	Mean difference significance
Gender [cordv10]	0 = ‘female’, 1 = ‘male’	-	0.484	0.591
Partner [cordv1]	0 = ‘Non-partnered’, 1 = ‘Partnered’	-	0.597	0.413
Age [cordv9]	= 15–19	i	0.093	0.691
	= 20–24	ii	0.103	
	= 25–29	iii	0.095	
	= 30–34	iii	0.089	
	= 35–39	iv	0.092	
	= 40–44	iv	0.097	
	= 45–49	iv	0.094	
	= 50–54	iv	0.078	
	= 55–59	v	0.068	
	= 60–64	vi	0.056	
	= 65–69	vii	0.046	
	= 70–74	viii	0.034	
	= 75–79	ix	0.024	
	= 80–84	ix	0.019	
	= >85	ix	0.014	
Length of living in New Zealand [cordv7]	< 4 years	i	0.103	<0.001
	= 4–10 years	ii	0.139	
	= 10–25 years	iii	0.107	
	>25 years	iv	0.651	
Ethnicity [cordv18]	European	i	0.607	0.221
	Maori	ii	0.044	
	European/Maori	ii	0.079	
	Pacific	iii	0.188	
	Asian	iv	0.028	

	MELAA	v	0.037	
	Other	v	0.026	
Education [cordv15]	= No qualification	i	0.136	0.237
	= Certificates	ii	0.479	
	= Degree	iii	0.386	
Personal income[23] [cordv12]	= Zero	i	0.094	0.111
	= $1–$5,000	ii	0.059	
	= $5,001–$10,000	iii	0.049	
	= $10,001–$15,000	iii	0.083	
	= $15,001–$20,000	iii	0.09	
	= $20,001–$25,000	iii	0.065	
	= $25,001–$30,000	iii	0.057	
	= $30,001–$35,000	iii	0.053	
	= $35,001–$40,000	iii	0.056	
	= $40,001–$45,000	iii	0.083	
	= $45,001–$50,000	iii	0.068	
	= $50,001–$70,000	iv	0.086	
	= $70,001–$100,000	v	0.083	
	= $100,001–$150,000	v	0.047	
	= $150,001 or more	vi	0.026	
Household income [cordv13]	= Zero	i	0.005	0.046
	= $1–$5,000	ii	0.003	
	= $5,001–$10,000	iii	0.006	
	= $10,001–$15,000	iii	0.018	
	= $15,001–$20,000	iii	0.029	
	= $20,001–$25,000	iii	0.026	

[23] Strictly speaking, we should inflation adjust the thresholds for personal and household income, but this is not a practical approach owing to the questionnaire design; due to the categorical design of the income variable, we are not aware of individuals' absolute value of income. CPI inflation from the quarter one of 2008 to the second quarter of 2012 was approximately 12 per cent.

	= $25,001–$30,000	iii	0.037	
	= $30,001–$35,000	iii	0.031	
	= $35,001–$40,000	iii	0.032	
	= $40,001–$45,000	iii	0.058	
	= $45,001–$50,000	iii	0.058	
	= $50,001–$70,000	iv	0.093	
	= $70,001–$100,000	v	0.189	
	= $100,001–$150,000	v	0.218	
	= $150,001 or more	vi	0.196	
Household size [cordv6]	= One person	i	0.096	0.557
	= Two people	ii	0.258	
	= Three people	iii	0.208	
	= Four people	iv	0.241	
	= Five people	v	0.117	
	= Six people	vi	0.052	
	= Seven people	vii	0.017	
	= Eight people	viii	0.011	
Health perception [heaq01]	= Very dissatisfied	i	0.021	0.033
	= Dissatisfied	ii	0.092	
	= No feeling either way	iii	0.243	
	= Satisfied	iv	0.372	
	= Very satisfied	v	0.271	
Ability perception (includes individual's feeling about his/her knowledge, skills and abilities) [kasq01]	= Very dissatisfied	i	0.002	0.011
	= Dissatisfied	ii	0.042	
	= No feeling either way	iii	0.074	
	= Satisfied	iv	0.631	
	= Very satisfied	v	0.25	
Residential Satisfaction [houq01]	= Very dissatisfied	i	0.007	0.398

	= Dissatisfied	ii	0.066	
	= No feeling either way	ii	0.068	
	= Satisfied	ii	0.492	
	= Very satisfied	ii	0.367	
Urban area[24] [ua]	= Main urban	i	0.942	0.327
	= Secondary urban	ii	0.017	
	= Minor urban	iii	0.012	
	= Rural	iv	0.029	
Free time (Individual's feeling about having enough free time) [leiq01]	= Too much free time	i	0.421	0.01
	= The right amount of free time	ii	0.479	
	= Not enough free time	iii	0.1	
Socialising (Frequency of meeting friends) [socq07]	= Every day	i	0.087	0.311
	= Around 3–6 times a week	ii	0.149	
	= Around 1–2 times a week	ii	0.373	
	= Around once a fortnight	iii	0.213	
	= At least once in the last four weeks	iv	0.178	
Local political involvement [humq03]	0 = 'No', 1 = 'Yes'	-	0.382	0.583
Recycling (How much does the household recycle) [phyq06]	= None	i	0.011	0.478
	= A little	i	0.02	
	= Some	ii	0.11	
	= Most	iii	0.556	
	= All	iv	0.303	
Council services (Individual's feeling about the quality of council services) [phyq05]	= Very dissatisfied	i	0.025	<0.001
	= Dissatisfied	ii	0.099	
	= No feeling either way	iii	0.138	
	= Satisfied	iv	0.603	

[24] The definition is based on the 1996 Census. Based on this, definition of urban area types differ based on how strong their economic ties are, how active they are from cultural and recreational point of view, how well they offer services to businesses, how easy it is to access their transportation network, and their prospective development.

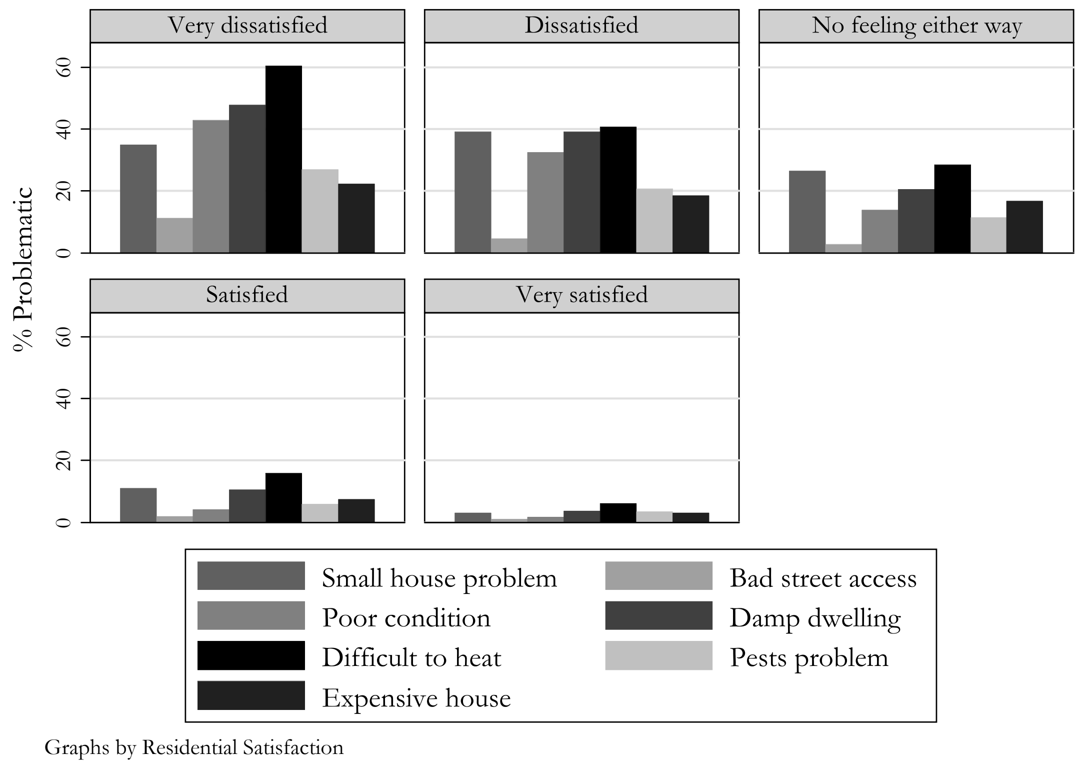
	= Very satisfied	v	0.135	
Green space access (Access to local green spaces including bushes, forests, nature reserves) [phyq13]	= Never want or need to go	i	0.018	0.01
	= None of them	i	0.01	
	= A few of them	i	0.056	
	= Some of them	i	0.152	
	= Most of them	ii	0.437	
	= All of them	iii	0.328	
Green space state [phyq14]	= Not been to	i	0.001	0.042
	= Very dissatisfied	ii	0.002	
	= Dissatisfied	iii	0.03	
	= No feeling either way	iv	0.092	
	= Satisfied	v	0.644	
	= Very satisfied	v	0.221	
Coastline access (Access to local lakes, rivers, harbours and oceans) [phyq11]	= Never want or need to go	i	0.015	0.02
	= None of them	i	0.017	
	= A few of them	i	0.055	
	= Some of them	i	0.15	
	= Most of them	ii	0.42	
	= All of them	iii	0.343	
Coastline state (Household's feeling about the state of coastlines) [phyq12]	= Not been to	i	0.01	0.131
	= Very dissatisfied	ii	0.007	
	= Dissatisfied	iii	0.081	
	= No feeling either way	iv	0.113	
	= Satisfied	iv	0.632	
	= Very satisfied	v	0.157	
Water use (How often does the household try to minimise water use?) [phyq09]	= Never	i	0.085	0.336
	= A little of the time	ii	0.107	
	= Some of the time	iii	0.267	

	= Most of the time	iv	0.398	
	= All of the time	iv	0.143	
Energy use (How often does the household try to minimise energy use?) [phyq07]	= Never	i	0.023	0.056
	= A little of the time	ii	0.086	
	= Some of the time	ii	0.299	
	= Most of the time	iii	0.466	
	= All of the time	iii	0.126	
Facilities access (including shops, schools, post shops, libraries and medical services) [phyq01]	= Never want or need to go to any of them	i	0.003	0.634
	= None of them	ii	0.006	
	= A few of them	ii	0.025	
	= Some of them	iii	0.058	
	= Most of them	iii	0.302	
	= All of them	iv	0.606	
Year	= 2008	-	0.339	
	= 2010	-	0.332	
	= 2012	-	0.329	
Homeownership status [cordv16]	= Owned, not defined	i	0.001	0.175
	= Owned, mortgage	i	0.336	
	= Owned, no mortgage	i	0.179	
	= Not owned, not defined	ii	0.001	
	= Not owned, rent	ii	0.313	
	= Not owned, no rent	ii	0.014	
	= Family trust, not defined	iii	0.002	
	= Family trust, mortgage	iii	0.08	
	= Family trust, no mortgage	iii	0.073	
			2008 / 2010 / 2012	
Small-house perception [houq03_11]	Being too small is a major problem with the person's house/flat? 0 = 'No', 1 = 'Yes'		0.112 / 0.104 / 0.101	0.696
Bad street access [houq03_12]	Being hard to get to from the street is a major problem with the person's house/flat? 0 = 'No', 1 = 'Yes'		0.023 / 0.024 / 0.011	0.037
Poor condition [houq03_13]	Being in poor condition is a major problem with the person's house/flat? 0 = 'No', 1 = 'Yes'		0.068 / 0.052 / 0.057	0.292

Damp dwelling [houq03_14]	Being damp is a major problem with the person's house/flat? 0 = 'No', 1 = 'Yes'		0.098	0.109	0.105	0.716
Difficult to heat [houq03_15]	Being too cold, or difficult to heat/keep warm is a major problem with the person's house/flat? 0 = 'No', 1 = 'Yes'		0.143	0.144	0.151	0.85
Having pests [houq03_16]	Having pests such as mice or insects is a major problem with the person's house/flat? 0 = 'No', 1 = 'Yes'		0.065	0.066	0.053	0.316
Expensive house [houq03_17]	Being too expensive is a major problem with the person's house/flat? 0 = 'No', 1 = 'Yes'		0.082	0.064	0.073	0.167
Far from work [houq04_11]	Being too far from work is a major problem with the person 's street/neighbourhood? 0 = 'No', 1 = 'Yes'		0.067	0.049	0.03	<0.001
Far from facilities [houq04_12]	Being too far from other things that he/she wants to get to is a major problem with the person's street/neighbourhood? 0 = 'No', 1 = 'Yes'		0.041	0.037	0.018	0.02
Unsafe neighbourhood [houq04_13]	Being unsafe is a major problem with the person's street/neighbourhood? 0 = 'No', 1 = 'Yes'		0.052	0.032	0.035	0.061
Noise and vibration [houq04_14]	Noise or vibration is a major problem with the person's street/neighbourhood? 0 = 'No', 1 = 'Yes'		0.14	0.123	0.102	0.045
Air pollution [houq04_14]	Air pollution from traffic fumes, industry or other smoke is a major problem with the person's street/neighbourhood? 0 = 'No', 1 = 'Yes'		0.052	0.034	0.029	0.013
Household crowding (Canadian index) [houdv2]	= One bedroom needed	i	0.003	0.073	0.081	<0.001
	= No bedrooms needed	ii	0.135	0.264	0.242	
	= One bedroom spare	iii	0.374	0.343	0.329	
	= Two or more bedrooms spare	iv	0.488	0.32	0.347	
Observations	5715					

Figure 9 illustrates the percentages of various housing problems across each residential satisfaction category. The most important problem amongst all satisfaction groups is the problem with heating the dwelling. For instance, the first histogram shows that approximately 60 per cent of those who reported they were dissatisfied with their residential situation said the house was difficult to heat, while 6 per cent of those very satisfied with their residential situation found their house difficult to heat. Insufficient space, poor condition and dampness of the dwelling were the next three most common problems amongst almost all satisfaction groups, except for the 'very satisfied' group, who stated that having pests in the house was the second-most frequent problem. Overall, the more satisfied people reported fewer problems with their dwellings.

Figure 9. The proportion of various problems with dwellings across different residential satisfaction levels.

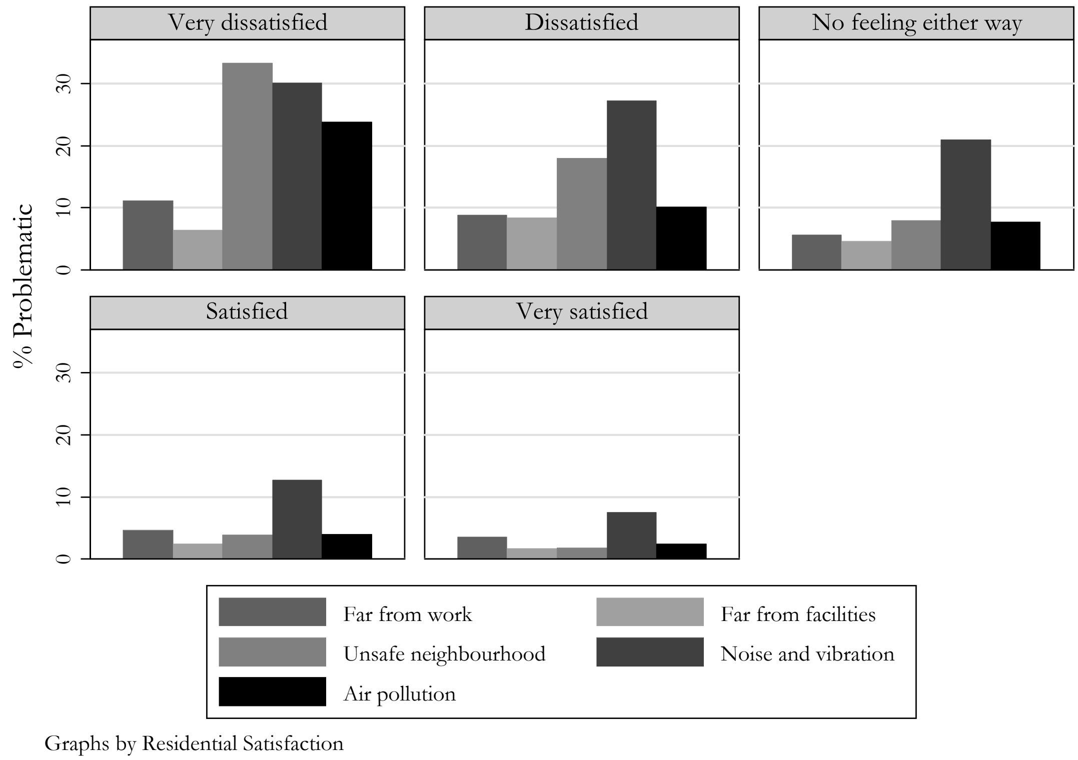


Source: Torshizian (2017).

Figure 10 illustrates the proportion of various problems with neighbourhoods amongst different residential satisfaction levels. For most groups, noise or vibration was reported as the most important problem. For the most dissatisfied people, the safety of the neighbourhood is the most prevalent problem. For the most satisfied people, the distance to their workplace is the second-most significant problem and safety is the least problematic issue. In the current study, the neighbourhood challenges are hypothesised to have a negative correlation with residential satisfaction. In order to control for the variations in residential satisfaction that have been caused by the residential environment's problems, all of the neighbourhood problems are included as control variables in the equations in the following chapters.

Figures 9 and 10 together indicate that our interpretation of the residential satisfaction variable as being determined both by house-specific and neighbourhood-specific factors is likely to hold. In each figure, I see that the prevalence of problems (with the house and the neighbourhood respectively) decline along with the degree of residential satisfaction. Thus, I consider it appropriate to interpret the residential satisfaction variable as reflecting factors at both the house and the neighbourhood levels.

Figure 10. The proportion of various problems with neighbourhoods across different residential satisfaction levels.



Source: Torshizian (2017).

Morrison (1994) used published and unpublished census tabulations, together with spatial analysis, to analyse the housing occupancy and the changing size of New Zealand households over a 40-year period. He highlighted the need for closer cooperation between researchers and census-takers in order to identify the problems with the datasets and also the needs of researchers for a better understanding of crowding. Baker, Goodyear, Telfar Barnard, and Howden-Chapman (2006) provided a comprehensive description of household crowding data using census data from 1991 to 2006. Based on their analysis using Census 2006, 10.4 per cent of New Zealand's population was exposed to household crowding. This number is slightly higher than the 10.1 per cent in 2001 but much lower than the 1991 figure of 11.9 per cent. Based on their analysis, Auckland's household crowding has not changed between 1991 and 2006.

Table 3 illustrates the percentage of crowded households with 1 or more bedroom deficit for 1991 to 2006 as reported by Baker, Goodyear, Telfar Barnard and Howden-Chapman (2006). The prevalence of one or more bedroom deficit in the Auckland region, which is the case of study in the current thesis, dropped from 6 per cent in 1991 to 5.2 per cent in 2006.[25] This percentage changed from 17 per cent in 1991 to 16 per cent in 2006 for Auckland metropolitan area. Looking at suburbs of Auckland metropolitan area,

[25] Their measure of household crowding is Canadian National Occupation Standard, which is introduced in Section 6.2.1.

derives a 9.4 percentage household crowding for 2006 in the central areas of Auckland (Waitemata), while in the same year the percentage of crowding in Manukau is equal to 21.9.

Table 3. Auckland's crowding level over time.

	1991	1996	2001	2006
Auckland region	6.0%	5.7%	4.9%	5.2%
Auckland city	17%	17.7%	16.4%	16.5%
Waitemata	9.7%	9.4%	9.1%	9.4%
Manukau	21.1%	21.3%	21.2%	21.9%
Waikato	3.2%	3.1%	2.7%	3.1%
Canterbury	1.5%	1.4%	1.0%	1.4%

Source: Baker et al. (2006).

The percentages alone are not informative enough because they are aggregates of many suburbs. For a better understanding of the percentages we need to have a careful geographic analysis, which will be discussed in Chapter 7. A policy may aim to decrease the percentage of the households in Manukau with a bedroom deficit of one or more to the percentage in central areas (9.4 per cent). We need to have a good understanding of residents' choices and the proximities of the suburbs with specific features in order to have a good benchmark for judging the appropriate percentage of crowded households in an area. As will be discussed in the next section, this study recognises the likelihood of being residentially satisfied as an appropriate benchmark for importance of the differences between residents' crowding levels.

In a cross-country study, Goodyear and Fabian (2012) compared New Zealand's household crowding levels and patterns with similar countries, including Australia, the United Kingdom, Canada, and the United States. Their results indicate that household crowding in New Zealand is higher than Australia, while New Zealand's crowding is almost half that of the United States. They found that crowded households have similar characteristics internationally. For example, crowding rates are higher for indigenous populations. Statistics New Zealand (2012b) examined crowding amongst ethnic groups in New Zealand from 1991 to 2006 and the results indicate higher crowding levels for some ethnic groups; for example, one-quarter of Maori people lived in crowded households in 2006.

Summary

Table 2 describes the wide range of variables that are used in the current study. The most important problems that residents have with their dwellings are heating the dwelling and having a small house. Amongst neighbourhood problems, having noise or vibration is reported as the most important problem, followed by safety issues in the neighbourhood. All problems with dwelling and neighbourhoods are included in the analyses of the following chapters.

Looking at the changes in crowding levels in Auckland's metropolitan area, I found much higher crowding levels in southern suburbs compared to the central area. This finding highlights the importance of residents' choice of residential location and of the distribution of amenities and facilities across the city. I also highlighted the importance of having a valid benchmark in understanding the appropriate level of crowding in a specific area.

4 The Outcome Of Interest

Chapter outline

Two consecutive papers were published in the *Journal of Economic Literature* (2009) on 'the wealth of cities' (Glaeser & Gottlieb, 2009) and on 'the quest for a measure of social welfare' (Fleurbaey, 2009). Glaeser and Gottlieb (2009) examined the relationship between productivity and density with a focus on house prices. Fleurbaey (2009) responded to the need for economists to find the right measure for assessing social welfare. Despite the recent attention that economists have paid to alternative measures of welfare, urban economists have only rarely considered outcomes other than prices and rents. In this section I review the literature of these measurement methods.

Residential satisfaction, which is the current study's outcome of interest, has rarely been used in previous urban economic studies. This section explains the pros and cons of considering this subjective measure as the outcome of this study.

4.1 Hedonic demand theory

A popular outcome in urban economics is the price of houses. For example, Cheshire and Sheppard (2004) assessed the impact that a change in the quality of schools has on house prices. They argued that the quality of schools, as a local public good, has been capitalised into the prices of houses. Theirs is one of the many urban economic studies to use the hedonic pricing method that was initially developed by Rosen (1974).

Hedonic demand theory is a revealed preference method of estimating the prices of houses,[26] which are derived from demand and supply of housing. The theory assumes that the price of a house is affected by a particular combination of the features of houses. To the best of my knowledge, Samuelson (1938) was the first economic paper on the revealed preference method. He made a strong assumption that consumers' choices reveal their preferences.

A critical point about the revealed preference method is its limitation in monitoring the consumers' ability to costlessly choose between alternatives. For example, a consumer may be dissatisfied with its household crowding level but, due to the heterogeneous nature of the housing stock and the costs of moving, may not be able to choose a better house. The dissatisfaction of the consumer with his or her crowding level will not be captured using a revealed preference method (Kőszegi & Rabin, 2007).

The hedonic demand theory has a few other limitations. People may not know about, or be able to observe, all aspects of their purchase. Therefore, the price may not account for all of the important factors. For example, buyers may not have prior knowledge about the level of pollution (or similar externalities) in the

[26] Revealed preferences interpretation of house prices is that if house A, with specific characteristics, is more expensive than house B, house A's features are more desirable than house B's characteristics.

neighbourhood of their new home purchase. Generally, prices do not accurately reflect the factors that buyers have not experienced before or cannot observe. For example, a person's willingness to pay may not show that person's unawareness of environmental qualities. Therefore, a buyer, who has not experienced environmental issues in his or her previous residence may not account for environmental qualities in his or her new neighbourhood. Another important issue with most hedonic pricing studies is that the wide range of data required for a causal study of prices is often unavailable.

This thesis is one of the few studies to consider residential satisfaction as its outcome of interest. The outcome of the study can be used for a comparison with the previous studies that have used the hedonic pricing method to verify the impact of unrevealed preferences on our understanding of the impact that different features of houses and neighbourhoods have on the welfare of residents. The following section provides an introduction to the use of subjective measurement in economics literature. I then discuss the merits of using the subjective measures.

4.2 Subjective measures and the economics literature

The dependent variable used throughout this thesis is residential satisfaction. Section 2.2 described this measure briefly. Residential satisfaction is a self-reported measure of happiness. The emergence of subjective measurement in the economics literature is briefly discussed in this section.

As Layard (2003) noted, economists in the 18th century were looking for a measure of utility or happiness that was measurable and comparable across people. They assumed that the marginal utility of income is higher for poor people than for rich people and, therefore, income should be re-distributed unless there will be a high efficiency cost. The utilitarian[27] philosophers and economists believed that public policy should aim to maximise society's sum of happiness. Robbins (1932) challenged the utilitarian approach by arguing that, in order to predict a person's behaviour, economists do not need to measure that person's level of happiness or compare his or her level of happiness with others. In other words, Robbins stated that economics only allocates scarce resources to fixed preferences, which are formed outside the scope of economics.[28]

Factors other than prices are involved in people's understanding of their environment. Rationality, in neoclassics terms, is defined as the behaviour with the best outcome for an individual with regard to the costs and benefits of an action. A person may still be rational but does not have the necessary tools, including information about his or her available choices, to make his or her desired choice. This situation is called 'bounded rationality'. However, neoclassics economists may not have reflected this situation in

[27] Based on the Utilitarianism theory, the best moral action is the one that maximises utility, which is usually defined in terms of well-being. Hence, the fulfilment of desire and happiness are the main meanings of utility.

[28] Robbins' statement was an agenda for positive economics, which is a branch of economics that focuses on the cause and effect relationships and is objective and fact-based. Normative economics, on the other hand, is subject to value judgement as it concerns what the outcome of the economic policy ought to be. For details, see Friedman (1953).

their utility maximisations. The rationality approach of neoclassics does not account for a range of behavioural issues, such as compromises and moral attachments to a group. Hence, Sen (1977) motivated incorporating behaviours motivated by feelings other than egoism in neoclassics' definition of rationality.[29]

Economists suggested a measure of national welfare similar to GDP but adjusted for leisure and pollution. Starting with Easterlin (1974), the use of subjective well-being measures in the economics literature become popular. Drawing on Easterlin's seminal study, Layard (2003) concludes that GDP is not an appropriate measure as it has not been informative enough since the Second World War, when the happiness of nations almost stagnated. Layard concluded that, for a better understanding of welfare, economists need to use psychology as well as economics and draw more on subjective wellbeing measures.

Since subjective quality of life is a multidimensional concept, it is measured using three broad measures. Happiness measures capture the aspects of well-being derived from experiencing pleasant emotions. Satisfaction, or cognitive well-being, measures assess the level of satisfaction with conditions of life (Lucas, Diener, & Suh, 1996). Quality of life measures capture the standard of living in material terms. Veenhoven (1994) noted three specifications of satisfaction measures. First, they are not necessarily in a constant state as people can change their view of life. Second, in the same situations, people are not satisfied to the same extent and their level of satisfaction depends on their evaluation of their situation especially relative to others. Third, satisfaction measures do not reflect solely internal factors; that is, subjective measures may depend on factors with a broader environment.

On the relationship between income and satisfaction, the 'Easterlin paradox' suggests no significant relationship between a society's economic development and its average level of happiness (Easterlin, 1974). Dowling and Yap (2007) noted three facts about the relationship between life satisfaction[30] and income: first, people with relatively higher income levels are generally more satisfied; second, people's income increases over time; and third, happiness is constant over time. In conclusion, material expectations change with the same proportion that income changes. This suggests that a higher welfare level caused by higher levels of consumption may vanish over time (Frey & Stutzer, 2002).

In fact, satisfaction is produced by changes in consumption rather than the level of consumption; that is, a constant consumption level results in diminishing increase of satisfaction (Helliwell et al., 2012). Subjective measures may reflect features that affect utility other than the person's current consumption level, including relative comparisons with others, and adaptation to past consumption. These issues were highlighted by Duesenberry's (1949) relative income hypothesis, which suggests that an individual's attitude to consumption and saving depends on his or her income in relation to others than by abstract standard of living; the percentage of income consumed by an individual depends on his or her percentile position within

[29] For further details on the relationship between well-being and consumption see the 8th chapter of Luigino and Porta (2007)'s handbook on the economics of happiness.

[30] Throughout this thesis, happiness and life satisfaction have been used interchangably.

the income distribution. This idea of relativity will be the main focus of the third essay of this thesis, presented in Chapter 8. Secondly, the present consumption is not influenced merely by present levels of absolute and relative income, but also by levels of consumption attained in previous period. It is difficult for a family to reduce a level of consumption once attained. The aggregate ratio of consumption to income is assumed to depend on the level of present income relative to past peak income.

Glaeser, Gottlieb, and Ziv (2014) asked, 'Why are the residents of some cities persistently less happy? Given that they are, why do people choose to live in unhappy cities?' They recognised the existence of differences in self-reported well-being across cities in the United States. However, this finding is in contradiction with the concept of spatial equilibrium that people cannot improve their overall utility levels by migrating within the US; that is, there is no arbitrage opportunity across space. Therefore, the authors concluded that subjective well-being is not the same as the economists' concept of utility. Their results imply that subjective well-being is one of many arguments of the utility function. As will be discussed in Section 5.1, the approach towards the relationship between utility and subjective well-being in the current thesis is similar to that of Glaeser et al. (2014).

4.3 Features of subjective measures

Diener (2009) discussed the reasons for choosing a subjective measure over an objective one. An advantage of the subjective indicators, and the so-called self-reported measures of happiness, is that people's own evaluations of their lives are reflected in their satisfaction indicators. The subjective measures can be of great use to researchers as they are equally, or arguably more directly, related to people's well-being compared to objective measures (Deiner, Lucas, & Oishi, 2002; Diener, Sandvik, Seidlitz, & Diener, 1993; Helliwell et al., 2012; Pavot, Diener, Colvin, & Sandvik, 1991). Well-being measures provide us with people's understanding of their own situation. Using this direct personal judgement, which is the most democratic of the well-being measures, provides policy makers with an opportunity to base policies on people's understanding of a good life rather than the policy makers' judgement of it.[31]

Helliwell et al. (2012) claimed that subjective measures are valid, robust and reliable. They argued that aiming for objective measures, such as GDP and prices, needs to be complemented by aiming for subjective measures, such as happiness. They looked at four dimensions of sustainable development: ending extreme poverty, environmental sustainability, social inclusion and good governance. While aiming for an objective

[31] Due to differences in backgrounds, people's understanding of the same matter is different. As people's experiences and evaluations are captured by their subjective assessments, these measures provide broader information than economic and social indicators. As an example, Auckland Council's chief economist noted that:

'The most fundamental challenge is how to balance what communities judge is best for them with what is best for Auckland and the country. There is a need to relax regulatory controls on height and density, while retaining provisions for quality urban design. That will allow more dwellings per unit of land, more choices, and better housing affordability' (Parker, 2015).

However, the tradeoff between housing intensification and residential preferences needs to be verified first, a topic addressed by this dissertation.

measure may lead to an extreme focus on specific aspects of sustainable development, aiming for subjective measures provides a more comprehensive result.[32]

In 2010, the UK's prime minister announced that the well-being measures would be used for policy advice. Bache and Reardon (2013) referred to the success of well-being measures in entering the UK's policy environment. However, they also reminded us of the need to account for the complexities of the subjective measures, especially the importance of accounting for spatial proximities.[33] I will discuss these issues in the following chapters.

The retrospective nature of subjective measures may lead to an alignment between people's responses to subjective questions and their behaviour. In other words, people may try to avoid cognitive dissonance between their attitudes and behaviours. Sloan and Morrison (2015) studied the post-move satisfaction of people moving within New Zealand. They use Statistics New Zealand's 2007 Dynamics and Motivations for Migration survey, which provided them with approximately 5000 observations. A contribution of their study on the impact of moving is to use subjective measures as their outcome of interest. They explored the subjective responses to moving in detail and concluded that moving to a large urban area is associated with an increase in satisfaction of movers in some domains such as employment, but a decrease in the movers' housing satisfaction. Sloan (2013) discussed that people may rationalise their outcomes by looking at their past and retrospectively lowering their anticipated outcomes. Therefore, in Sloan's example of the post-migration satisfaction, one's anticipated outcome of migration may be reduced such that it matches his or her realised outcomes; that is, the individual likes to decrease his or her experienced dissonance. The retrospective inaccuracies of subjective measures are mainly caused by a learning process that is highly associated with a person's backgrounds and age. This issue is discussed further in Chapter 7.

Based on mainstream economics, the price of a good is where demand meets supply. Demand for a good consists of the ability and willingness to pay. Individuals maximise their consumption of housing and non-housing goods subject to their budget constraints. A difference between housing and non-housing goods is the finances that banks provide for households' consumption of owner-occupied housing, in terms of mortgages, while households do not receive similar finances for their non-housing consumption. This difference between the ability to pay and the willingness to pay for housing consumption may be captured by using subjective measures.[34]

[32] As an example, there is often a battle between proponents of intensification, who are usually planners, and the fans of relaxing legislative containments, who are usually from the real estate industry. At the centre of this debate is people's desirable living environment. Aiming at this subjective outcome may help policy makers achieve their sustainable planning goals without supporting either planners or the real estate industry.

[33] On a related point, see Morrison and Weijers (2012).

[34] The finances provided on housing may change consumption patterns as people may consider housing more as an investment good than a consumption good.

Morrison (2014b) identified the need to use subjective measures in regional studies:

> 'Subjective measures are particularly useful in areas where the distribution of outcomes is not easily identified using other, especially market, criteria. The effect of investment in public infrastructure or the provision of green space or in fostering community networks or in redeveloping neighbourhoods can be captured in responses to questions on wellbeing, preferably over time. These subjective measures, which have been shown to be highly correlated with clinical and other assessments of well-being, are likely to be of particular interest in regional science because of the way changes to places result from, or generate, a range of positive or negative externalities.'

Also, a better understanding of the impact of different factors on well-being measures provides us with a complementary understanding of the impacts of economic trade-offs that result from a policy. One example of such a study is that of Tella and MacCulloch (2007), who analysed the relationship between life satisfaction, unemployment rate and inflation rate of Europeans. The authors highlighted the importance of the impact of a change in each of the rates, namely unemployment rate and inflation rate, on people's happiness.[35] Use of subjective measures helps policy makers to evaluate the economic trade-offs that were previously hard to decide on.

4.4 Are subjective measures superior to objective ones?

As discussed above, prices are the most common outcome of interest in the urban economics literature. House prices are the sum of the value of the improvements and the value of land. According to Hansen (1982), land prices consist of the land's agricultural value, the cost of conversion from agricultural land to residential land, the value of the accessibility of the land and the expected future rental value.

Table 4, presented in the summary of this section, briefly illustrates the features of the subjective measures. Reflecting people's understanding of their current situation is one of the features of the subjective measures. Because prices of assets, including house prices, reflect current and future returns to those assets, including rents, they are not necessarily a useful indicator of how people feel about their current situation. Based on Hansen's (1982) decomposition of the price measure, considering that prices contain information on the expected future rental value, prices are an amalgamation of perceptions about current and future conditions. By contrast, a subjective measure tells us about how people feel regarding their current situation and are therefore a better indicator of the present.

Because of the lack of information, most hedonic studies do not account for expectations. For estimating expectations, I need to account for a wide range of the characteristics of both the buyer and the seller. In

[35] Their results suggest that the impact of a 1 per cent increase in unemployment on people's happiness is twice as negative as the impact of a 1 per cent increase in inflation rate (on happiness).

addition, a careful study of prices must account for macroeconomic factors and spatial proximities to amenities and facilities.[36]

Another feature of the subjective measures is that they are based on people's understanding of their condition. However, prices form based on people's expectation of the situation of the house that they are paying for. People may not be aware of all aspects of their purchase before settling down in the house that they are buying. In the case of a relief of the urban limit, for example, many people may prefer to pay relatively high prices in order to settle into newly developed areas. However, the buyers are very likely to be unaware of the frustration involved in living far away from central areas. If the policy makers use price signals in this specific case, they may not achieve a sustainable development, which aims to increase residents' satisfaction with their living environment. Therefore, the price may not really account for all of the important factors.

Our focus on residential satisfaction does not mean that prices are not informative. The main advantage of subjective measures is the information that they provide about consumer's unrevealed preferences. Prices can still be used for various purposes. A change in the trend of prices may be a good indicator of a change in macroprudential conditions. For example, a general increase in prices may be caused by an increase in the prices that residents can afford to pay as a result of a decrease in interest rates[37] (Torshizian, 2016a). A change in the relative prices amongst areas may be an indicator of the impacts of a change in local government's regulations (e.g., Torshizian, 2016b) or a change in local government's services, such as a change in transportation services (Grimes & Young, 2013).

Summary

Using prices as the outcome of interest is common amongst urban economic studies. One of the shortcomings of the price measure is that it is not informative about people's desired choice that has not been provided in the market; in other words, prices do not present unrevealed preferences.

The use of subjective measures in the economics literature became more common after Easterlin's (1974) study on the economics of happiness. Many economists have discussed that subjective measures (as opposed to objective ones) are measures of the consumer's utility

[36] Many timeseries studies using objective measures have discussed the importance of accounting for exogeneity of the output of interest with respect to the past values.[36] Due to the usual data limitations, accounting for future values is almost impossible. Cross-section studies of objective measures use comprehensive datasets and advanced econometric methods, such as two-stage least squares, to account for the expected levels of the objective factors. However, the wide range of data is usually not available for Hedonic studies.

[37] Here we are only referring to using prices in order to inform policy. However, using a specific modelling method may limit prices' contribution to policy. For example, while hedonic models use prices as an outcome of interest, they do not account for macroprudentials; this is one of the criticisms of the use of hedonic models.

because they account for psychology as well as economics, and also account for the desired choices of consumers that have not been provided in the market. A brief comparison between subjective and objective measures is presented in the following table:

Table 4. Features of subjective and objective measures.

Subjective measures	**Objective measures**
Account for both revealed and unrevealed preferences	Only account for revealed preferences; that is, do not hold information about unrevealed preferences
People's evaluation of their living condition	Researcher's evaluation of people's living condition
Reflect people's understanding of their current situation	Hold information about future expectations
Subject to relative positions of individuals	Mainly based on people's absolute position
Adjust for people's expectations based on their cognitive dissonance	The role of expectations is not clear (mostly due to data limitations)

5 Common Methodology

Chapter outline

This chapter explains the basics of the methodologies of the following chapters. This chapter begins with a description of the general form of the utility function used in the current study. I then describe the econometric approach, followed by a discussion of the nature of the dependent variable (that is, the outcome variable of interest).

5.1 Utility function

Based on demand theory, households maximise their utility subject to their income constraint. We can separate utility into two components: residential satisfaction and satisfaction with other aspects of life. The budget constraint is equal to the expenditure on the residential good plus the expenditure on non-residential goods. The outcome of interest throughout of this thesis is residential satisfaction (RS), which is a function of a vector of house-specific factors (H) and a vector of neighbourhood characteristics (N):

$$\mathrm{RS} \equiv \mathrm{f(H, N)}, \qquad (4)$$

$$\mathrm{H} \equiv (\mathrm{H_1}, \mathrm{H_2}, \dots), \mathrm{N} \equiv (\mathrm{N_1}, \mathrm{N_2}, \dots)$$

Household utility (U) depends on the utility derived from some composite residential good (R) and the utility of consuming non-residential goods (C):

$$\mathrm{U} = \mathrm{u(R)} + \gamma.\mathrm{v(C)} \qquad (5)$$

Where u(R) is actually equal to RS. Based on demand theory, households increase their utility (U) subject to their income constraint (Y). Hence, I maximise U subject to the budget constraint:

$$\mathrm{s.t.\,Y} = \mathrm{C.P_C} + \mathrm{X.P_X} \qquad (6)$$

Where $\mathrm{P_C}$ is the price of C and $\mathrm{P_X}$ is a vector of prices related to each aspect of R. Then the Lagrangian will be as follows:

$$\mathrm{L} = \mathrm{u(R)} + \gamma.\mathrm{v(C)} + \lambda(\mathrm{Y} - \mathrm{C.P_C} - \mathrm{X.P_X})$$

The derivatives are as follows:

$$\mathrm{dL}/_{\mathrm{dR}} = \mathrm{u_R} - \lambda.\mathrm{P_R} = 0 \qquad (7)$$

$$\mathrm{dL}/_{\mathrm{dC}} = \mathrm{v_C} - \lambda.\mathrm{P_C} = 0 \qquad (8)$$

Dividing (8) by (9) derives:

$$\frac{u_R}{v_C} = \frac{P_R}{P_C} \tag{9}$$

Using equations (7) and (10) give each of R and C and, hence, u(R) and v(C), in terms of Y and $\frac{P_R}{P_C}$.

P_R is a composite price of H and N, as explained in equation (5). In other words, P_R is the sum of hedonic prices of each component of H and N. As the price of each components rises, the quantity consumed of that component will fall and, hence, RS will fall. Thus, I can model RS directly as a function of the components of H and N after controlling for the economic factors that affect the overall allocation of resources between R and C.

As will be discussed in the econometric section, I will control for the impact of Y in the subsequent chapters. In addition to Y, to solve the maximisation problem I must identify the relationship between P_C and P_R[38]. I control for a variable that indicates respondents' valuation of housing prices (expensive or not). I also control for the time fixed effects in order to account for changes in cost of housing across surveys.

In equation (7), g(C) differs across people with different backgrounds. The econometric methodology presented in Chapter 7 is used for solving the maximisation problem. This method indicates the importance of the factors of space, including crowding and density, to people with different backgrounds. If space is expensive, that will be more negative for someone who likes space than for someone who is indifferent to it.

5.2 The econometric basis for welfare maximisation

A typical empirical study consists of a relationship between a dependent variable and an independent variable[39] that may have an impact on the dependent variable. In the present study, I have named the dependent variable y and the variables of interest x. To account for identification problems, including the issues that were discussed in Section 2.4.4, a number of variables are controlled for. Consequently, the typical model will be as follows:

$$\mathbf{E(y_i|x_i, c_i) = \sum_{i=1}^{n} w_i \beta_i} \tag{10}$$

where c_i represents the control variables that enable us to abstract from the effect of all other variables when examining the impact of the variable of interest. x_i is a column vector of the variable of interest for n observations. w_i consists of a unity column that takes into account the intercept for n observations, and a

[38] To find the optimal choice, I need to know the proportion of income that has been used for residential purposes. This point happens at the tangent point of the utility curve and the budget constraint line, the slope of which is equal to the ratio of the residential good's price to the non-residential good's price. This price ratio is what we ordinally know as how expensive housing is (compared to other goods).

[39] There may be more independent variables.

matrix of the values of c_i plus an additional column of x_i. By assuming k-1 control variables, w_i is a n*k+1 matrix. The coefficients to be estimated are included in β_i , which is a column vector of length k+1.

Our dependent variable throughout this dissertation follows a Bernoulli distribution; in other words, the dependent variable is a dummy variable. Thus, the estimated values derived from the right-hand side of equation (11) should be modified such that the outcome is between 0 and 1. To do this, I adopted a categorical dependent variable transform function. Two common functions are logit and probit. Consequently, equation (11) is as follows:

$$\mathbf{E(y_i|x_i,\ c_i) = G(Z) = G(\sum_{i=1}^{n} w_i\boldsymbol{\beta}_i)}$$

In this study, a logit transform function is used, the reason for which will be discussed in the next chapter. Consequently, the log-odds of the probability of being satisfied (dependent variable = 1) are:

$$\mathbf{log(\frac{\boldsymbol{\pi}_i}{1-\boldsymbol{\pi}_i}) = \sum_{i=1}^{n} w_i\boldsymbol{\beta}_i} \tag{11}$$

5.2.1 The appropriate measure of goodness of fit

There are a number of tests for finding how well a statistical model fits the set of actual observations. A list of measures of goodness of fit includes the coefficient of determination (also known as R^2), likelihood ratio test, chi-square test, Akaike information criterion (AIC), Bayesian information criterion (BIC), Hosmer–Lemeshow test, Kolmogorov–Smirnov test, Cramér–von Mises criterion, Anderson–Darling test and Shapiro–Wilk test. The use of these tests is case-specific; that is, it depends on the modelling approach.

The most common measure of goodness of fit (GOF) is the coefficient of determination, which measures the variability in the outcome that has been explained by the model. However, the logistic method is based on maximum likelihood estimation, which is an iterative process. Since this process does not minimise the variance, it is not possible to measure the coefficient of determination.

All of the GOF measures have pros and cons. For example, the Hosmer–Lemeshow test proposes grouping observations based on their predicted values from the logistic regression model. This test has been criticised by many econometricians because of the arbitrary choice of the number of groups. In this thesis, AIC and BIC[40] are used as measures of GOF. Both of these measures estimate the quality of each model relative to each of the other models.

Suppose L is the maximum value of the likelihood function and k is the number of parameters in the model. The value of AIC for this model will then be equal to 2k – 2ln(L). Therefore, AIC rewards GOF and

[40] AIC will be discussed in the next paragraph. For BIC, see Kenny (2012).

penalises an increasing number of estimators, which discourages overfitting. The model with the minimum AIC is the preferred model.

Summary

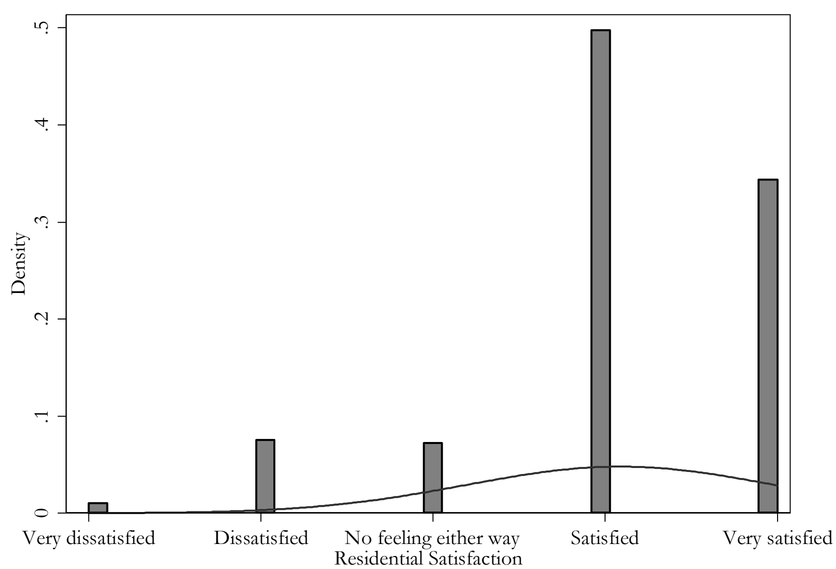
This chapter introduces the basics of the methodology used in this thesis. First, the utility function and the corresponding maximisation problem are presented. Then, in the 'econometric basis' section, I discuss the solution for the maximisation problem. The chosen transformation function throughout this thesis is logit. Amongst the measures of the GOF, AIC and BIC are introduced as appropriate measures whose values will be reported for the subsequent chapters' models.

5.3 Changing the ordinal dependent variable to a binary variable

As discussed in Section 2.2, while the outcome of interest is an ordinal variable, it is converted to a binary variable in this study. Two main questions need to be answered: Do results change if we use a binary variable instead of the original five-category variable? And, if not, which categories of the five-category variable should be aggregated to one group? Both of these questions are answered in the current section. According to Appendix ii of Chapter 6, the differences in the slope of the coefficients of the variables of interest is not significant. These reasons are discussed more thoroughly in the current section.

Section 2.2 introduces the asymmetric distribution of the residential satisfaction variable as a reason for the chosen categorisation of the final binary variable. Figure 11 depicts the distribution of the ordinal residential satisfaction variable. The proportion of the 'very dissatisfied' group is only 1 per cent. As depicted, the distribution is very negatively skewed such that 84 per cent of observations fit into the two major groups, namely 'very satisfied' and 'satisfied'. Considering the high proportion of the 'satisfied' group, the only sizeable recategorisation such that the objective will be a minority, that the majority of people are seeking, needs to distinguish between 'satisfied' and 'very satisfied' groups.

Figure 11. Distribution of the five-category residential satisfaction variable.



Source: Torshizian (2017).

As discussed in Section 3.3, 'the criteria for grouping some categories are first the diversity of the category, e.g. MELAA ethnicity, consisting of Middle Eastern, Latin American and African ethnicities, is a highly diverse category, thus it is grouped with 'Other', and second the results derived from different categories, i.e. two categories are grouped if they show very similar patterns in regression models'.

As an example, I compare the impact of one variable of interest, PPBR,[41] on different categories of the ordinal measure of residential satisfaction (ORS). The first estimation is simply based on a regression of the ordinal variable on PPBR and its squared form:

$$ORS_i = \alpha_0 + \alpha_1 . PPBR_i \; + \alpha_2 . PPBR_i^2 \; + \varepsilon_i \tag{12}$$

Where α_1 and α_2 are the parameters of interest; and ε_i is the error term. This estimation is done by using ordinal logit method. The results of the estimation are reported in the 'ordered logit' column of table 5.[42] I need to test the parallel assumption of the regression by using the Brant test:

Table 5. Ordered logit and Brant test.

	Ordered logit		Brant test: different categories			
	Estimates	p>chi2	ORS>1	ORS>2	ORS>3	ORS>4
PPBR	-0.89	0.78	-1.12	-0.74	-1.03	-0.75
PBBR^2	0.014	0.58	0.3	-0.03	0.11	-0.08
cut1	-5.53		5.4	3.18	2.56	0.11
cut2	-3.28					
cut3	-2.54					
cut4	-0.14					
N	4706					
brant_chi2	5.001					
brant_df	6					
brant_p	0.544					

The null hypothesis of the parallel-lines model is not violated here (P_value is equal to 0.544). The detailed results of the test are depicted in the last four columns of the table, which are titled 'Brant test: different categories'. This is a series of binary logistic regressions. First, 'ORS>1' is based on category 1 ('very dissatisfied') versus the other four categories, namely 'dissatisfied', 'no feeling either way', 'satisfied' and 'very satisfied'. The 'ORS>2' column is based on the categories 'very dissatisfied' and 'dissatisfied' versus the other three categories and so on. Considering that the estimated coefficients PPBR and PBBR^2 are not statistically different amongst these four binary logistic estimation, the null hypothesis of the parallel-line model is not violated.[43] The Brant test's result suggests that, for the variable of interest, I can rely on the parallel assumption of the ordered logit method. This implies that I do not need to use other methods such as generalised ordered logit and multinomial logit.

In the next step, I recategorise the five-category dependent variable to a binary one, which makes it easier to present and interpret the results. However, I need to ensure that no information will be lost as a result of the re-grouping. I compare the estimation results derived from an ordered logit with two binary logistic regressions with different definitions for the binary outcome of interest.

[41] As will be explained in the next chapter, PPBR is the main measure of crowding used in this thesis.

[42] The impacts of PPBR and PPBR^2 are jointly significant at better than the 1 per cent significance level.

[43] This is also shown by the Brant test's p_value.

Based on the distribution of the variable presented in Figure 11, the two most appropriate definitions for the binary outcome of interest are as follows:

$$\text{RS1} = \begin{cases} 1 & \text{if RS} = \text{"Very Satisfied"} \\ 0 & \text{if RS} = \text{Other groups} \end{cases} \tag{13}$$

$$\text{RS2} = \begin{cases} 1 & \text{if RS} = \text{"Very Satisfied" or "Satisfied"} \\ 0 & \text{if RS} = \text{Other groups} \end{cases} \tag{14}$$

As presented in equation (14), one definition is to consider the very satisfied category as one group and all the other categories as the other group. Equation (15) suggests considering two upper categories – very satisfied and satisfied – as one group and the other categories as another. The distribution of the variables constructed based on equations (14) and (15) is depicted in Table 6:

Table 6. Distribution of the binary dependent variable for different definitions.

	RS1		RS2	
	Freq.	Per cent	Freq.	Per cent
Dissatisfied	3080	65.45	728	15.47
Satisfied	1626	34.55	3978	84.53
Total	4607	100	4607	100.00

Note: As discussed in Section 7.4, after correcting the dataset and merging it with other dataset, 4607 observations are used for the analyses presented in the current thesis.

Then equation (13) needs to be estimated using the two different measures. For a reliable result, a wide range of control variables[44] is added to equation (13). Results are depicted in Table 7. The marginal effect of the variable of interest on RS1 is equal to -0.33, while the marginal effect for the ordered logit model is equal to -0.28. The marginal effect of PPBR in RS2 model is equal to -0.06. Therefore, assuming that ordered logit has derived correct results, I do not find any material differences between results of the binary logit using RS1 and ordered logit in terms of the marginal effect of the variable of interest. However, the other categorisation of RS results in a very different marginal effect for the variables of interest; this supports the use of RS1 instead of RS2. Also, the results suggest that the impact of PPBR on RS is not significantly different when I use a binary RS (RS2) instead of the ordered RS (ORS).

Table 7. Ordered logit compared to binary logit.

	Ordered logit	Binary logit	
	ORS	RS1	RS2

[44] A list of control variables is presented in section 6.6.1.

PPBR	0.452*	0.762	0.174**
	(0.193)	(0.386)	(0.125)
Squared PPBR	1.168	0.913	1.676*
	(0.226)	(0.222)	(0.512)
Control variables	Yes	Yes	Yes
cut1			
Constant	0.000631***		
	(0.001)		
cut2			
Constant	0.00989***		
	(0.010)		
cut3			
Constant	0.0259***		
	(0.026)		
cut4			
Constant	0.542		
	(0.535)		
Observations	4706	4706	4706
N_pop	925815.0	925815.0	925815.0
Marginal Effect	-0.33	-0.28	-0.06

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

A number of studies have argued that the impact of explanatory variables on being satisfied is not the same as the impact of them on being dis-satisfied (for example, Huppert and Whittington (2003)). However, the results of the Brant test confirm that the ordered logit model is reliable as the proportional odds assumption is not violated. This confirms that I can view satisfaction and dissatisfaction as essentially being the opposite of each other; that is, there is no reason to model factors explaining satisfaction and dissatisfaction separately.

Summary

To use the ordered logit method, I need to ensure that its parallel lines assumption is not violated. The Brant test's result suggests that, for the variable of interest, I can rely on the parallel lines assumption of the ordered logit method.

It will be easier to present and interpret the results if the five-category residential satisfaction variable is aggregated to a binary variable. However, I need to ensure that no information will be lost as a result of the re-grouping. Hence, I compare the estimation results derived from three models: a model using ordered logit method, and two binary logistic regressions with different definitions for the binary outcome of interest.

The results reveal no material differences between the ordered logit's results and the results derived from the binary variable. Hence, residential satisfaction in the current study is aggregated into two categories such that the 'very satisfied' category is considered as one group (0) and the other four categories are grouped as the other category (1).

5.4 Addressing the potential issues of subjective measures

In addition to the issues described in Section 2.4, our outcome of interest is subject to a range of measurement issues. As discussed in Section 2.2, similar to many subjective measures, our residential satisfaction is measured using a five-point Likert-type scale. An advantage of the Likert scale measurement is its cognitively light method of self-reporting (Knowles & Nathan, 1997). However, this scale is likely to be biased towards its innermost categories, which is called central tendency bias. Respondents may avoid choosing the extreme responses, which leads to a lower sensitivity of this scale. The other possible bias is over-optimism, or so-called acquiescence bias, which happens when respondents have a tendency to answer positively. For a debate about measurement methods of subjective measures, see Diener (2009). This section explains a number of potential measurement issues and the strategy adopted in this strategy to address them.

Survey responses may be susceptible to over-optimism, meaning that some respondents may answer questions with an overly positive (negative) attitude. This may be related to ethnic features. For example, some ethnicities are happier with different aspects of their life because of their traditional wisdoms, education level, length of stay in their host country and other individual characteristics. Hence, the RS variable may be over-estimated or under-estimated for some respondents. For example, some may say something like, 'as long as I am healthy I am the happiest person', which may lead to an exaggeration of other subjective statuses, such as satisfaction with residential environment (Larwood & Whittaker, 1977). To control for this effect, I included two other satisfaction variables that may reflect the same propensity to respond overly positively or negatively in my analyses; namely, satisfaction with health status and satisfaction with skills and abilities. This is a similar approach to the method that Helliwell (2003) took to address the high association issue (Halpern, 2005).

Econometrics considers the presence of measurement error in both the regressand and in the regressor. The potential over-optimism error considered in the previous paragraph is an example of a measurement error in the regressand. While the presence of measurement error in regressors may be solved by using certain techniques, the measurement error in the regressand needs to be addressed very carefully as the error will lead to biased standard errors, which may lead to a null hypothesis being mistakenly rejected. For the purpose of illustration, I assume that the following simple model needs to be estimated:

$$y_i = \alpha_0 + \alpha_1 x_{1i} + \alpha_2 x_{2i} + \epsilon_i \tag{15}$$

Where y_i is the true values of the outcome of interest, x_{1i} is the variable of interest, x_{2i} is a control variable and ϵ_i is the random error term. When the outcome of interest is systematically biased, the true value of it (y_i) is actually equal to the observed value ($\tilde{y}_i$) subtracted by the systematic error (ξ_i):

$$\tilde{y}_i = y_i + \xi_i \tag{16}$$

In other words, instead of estimating equation (17), the following model is being estimated:

$$\tilde{y}_i = \alpha_0 + \alpha_1 x_{1i} + \alpha_2 x_{2i} + \epsilon_i + \xi_i \quad (17)$$

Therefore, the presence of error in the outcome of interest leads to an additional term in the error, which reduces the power of the statistical test. However, the measurement error of the dependent variable does not lead to biased estimates of the coefficients (e.g., Wooldridge, 2015).

My solution is to consider a control variable with similar construction to the (potentially) biased output of interest. Hence, instead of the true value of x_{2i} I have $\tilde{x}_{2i}$, which is equal to:

$$\tilde{x}_{2i} = x_{2i} + \xi_i \quad (18)$$

Therefore, assuming the independency between the regressors, the estimated α_1 is the impact of x_{1i} on $\tilde{y}_i$ when x_{2i} is kept constant; that is:

$$\widehat{\alpha_1} = {}^{d\tilde{y}_i}/_{dx_{1i}}\Big|_{\tilde{x}_{2i}=\overline{x_{2i}+\xi_i}}$$

This shows that, in my estimation of the impact of the variable of interest (x_{1i}), I have also controlled for the individual specific error (ξ_i). Hence, ξ_i will no longer be a part of the error term, which means that the estimated standard error will be precise.

As Hamermesh (2004) discussed, including subjective variables on both sides of the equation does not provide us with a precise estimation of the impact of the explanatory subjective variable. This is because I am actually estimating the impact of an explanatory variable without excluding its measurement error. Therefore, the estimated impact of the variables used for controlling for over-optimism will be biased. However, the impact of the variable of interest will be estimated precisely. In the current study, except for subjective understanding of crowding (perceived crowding), which is the only subjective variable of interest in this thesis, all variables of interest are constructed based on objective values, such as the number of bedrooms and the household size. As I will discuss in the next chapter, perceived crowding will not be used as a variable of interest in the following chapters.

In the questionnaire of the survey used in the current study, both variables used for measuring the over-optimism issue are addressed in the modules prior to the housing module. Therefore, individuals are likely to think about their residential satisfaction isolated from any problem that they may have with their health status and abilities. This may imply that I interpret any significant relationship between residential satisfaction and the other measures of satisfaction as a correlation between the measurement errors.

Two other common issues are the source of concern in studies that use qualitative data; namely, misunderstanding of respondents and inaccuracy of responses. The misunderstanding of respondents may lead to measurement error in RS. However, since respondents were asked explicitly about their satisfaction with a number of specific aspects of life, the chance of misunderstanding is considered small.[45] The accuracy of qualitative data is discussed in a study of firms by Fabling, Grimes, and Stevens (2012). They showed that subjective responses by firms matched well with objective data for the same firms. They also emphasised that the accuracy of this data was enhanced by the presence of a 'don't know' category, allowing respondents the option of choosing the 'don't know' category when they are not sure. The inclusion of this category in my survey data mitigates the possibility of incorrect responses.

[45] The modules included in the questionnaire of GSS are presented in Section 2.2.

6 Is Residential Satisfaction Affected by Crowding? A Study of Raw, Adjusted and Perceived Crowding Measures

Abstract

The goal of this chapter is to compare the information content of subjective (perceived) versus objective measures of household crowding when estimating people's evaluation of satisfaction with their residential environment. Crowding can be measured using a number of indexes, some of which – such as the Canadian National Occupation Standard (CNOS) index – account for cultural issues and household composition, while others, such as the people per bedroom measure (PPBR), do not. In the present study, these measures are compared with one another and also with people's perceptions about the crowding of their dwellings in terms of their explanatory power for residential satisfaction. After controlling for a wide range of other influences, my results indicate that higher residential satisfaction is associated more with a perception of less crowding than with a lower value of an objective crowding measure. Amongst objective measures, CNOS is a better predictor of residential satisfaction than PPBR. Nevertheless, PPBR remains a useful predictor of residential satisfaction in the absence of the other measures.

6.1 Introduction

Residential satisfaction (RS) is related to many factors derived from the social and physical aspects of the living environment, such as safety and noise level. These two factors may be affected by the crowding and area/residential density of the living environment, as well as by other subjective and objective factors (Amérigo & Aragonés, 1997). The causal relationship is discussed by Rodgers (1982), who highlighted the causal impact of crowding and density on both dwelling and neighbourhood satisfaction. While crowding can be measured both subjectively and objectively, density remains an objective measure of intensification. The crowding scales have different shades of subjectivity; that is, some are more quantitative and some a more qualitative.

As discussed in Section 2.2, residential satisfaction is a combination of objective and subjective characteristics of the living environment. Individuals respond to a question about their satisfaction with their living environment based on their understanding of the characteristics of a desirable living environment. Therefore, a good measure of crowding is the one that best explains the variations of a nation's residential satisfaction.

In the next section I review the crowding concept and its three main measures. Section 6.3 discusses the hypothesis of this chapter. Section 6.4 presents the data and some information about the three measures of crowding: perceived crowding (PC), CNOS and PPBR. The research design section discusses the methods used in this study. The final section includes my results and suggestions for future research.

6.2 Literature review

In one New Zealand study, Maani, Vaithianathan and Wolfe (2006) considered the health impact of crowding. Their results show that the positive correlation between income inequality and health outcomes, such as the number of infectious diseases, can be partially explained by household crowding, which in they measured using the CNOS.[46] The definition of crowding differs amongst different studies, depending on their focus. For example, in a study on educational achievements, crowding may be defined as the number of persons per bedroom, while in a health economics study, it may be defined based on the composition of people living in a house, which can be measured by a crowding index.

Bonnes et al. (1991) took a contextual approach towards the negative relationship between crowding and residential satisfaction. The sample population of this study is 461 residents of a neighbourhood in Rome. The study used a factor analysis together with a regression analysis. However, due to lack of data, the study did not control for a wide range of variables. Their results indicate a significant spatial effect in studying the impact of crowding on residential satisfaction. They highlighted the need for further study on the differences between crowding perceptions and the objective measures of density and crowding. They also called for a careful definition of neighbourhood boundaries and concluded that residential satisfaction is affected to a significant degree by the perception of crowding.

6.2.1 Crowding measures

Floor area per person is regularly used by the United Nations and World Health Organisation as a quality of life indicator, more particularly as a measure of sustainable human settlement development. The people per room (PPR) index, also known as the American Crowding Index (ACI), is another commonly used crowding index. In the US, a house in which a room is shared with more than one person is considered to be crowded and a room with more than 1.5 people is considered to be very crowded. For instance, in a housing quality study, Cook and Bruin (1994) used the PPR measure, which is defined as the ratio of the number of persons in the house to the number of rooms in the house, as a crowding measure. The PPR measure is a raw crowding measure that does not take into account the type of the rooms nor the characteristics of the residents. For more details, see Table 8.

Table 8. Criteria of crowding indexes, inspired by Goodyear, Fabian and Hay (2011).

Index	Based on	Couple status	Age that boys and girls can share	Age that same sex children can share
PPR	Rooms	Not used	-	-
PPBR	Bedrooms	Not used	-	-
APPBR	Bedrooms	Used	-	-
CNOS	Bedrooms	Used	Under 5	Under 18
HIR	Bedrooms	Used	Under 10	Under 10

[46] As discussed in Section 2.4, however, the causal inference needs further investigation.

The dependent variable of this study is residential satisfaction, which is a subjective variable that may be affected by both objective and subjective factors. One measure of objective household crowding is the PPR index, which is a pure objective index; that is, it does not account for cultural issues like other indexes do. Another objective measure is people per bedroom (PPBR),[47] which is the number of people living in the house divided by the number of bedrooms in the house. As shown in equations (20) and (21), the main difference between PPR and PPBR is in the denominator; for PPR the denominator is the number of rooms and for PPBR it is the number of bedrooms.

$$\text{PPR} = \frac{\text{Number of people in the dwelling}}{\text{Number of rooms}} \tag{19}$$

$$\text{PPBR} = \frac{\text{Number of people in the dwelling}}{\text{Number of bedrooms}} \tag{20}$$

I derive another measure, called adjusted people per bedroom (APPBR), after controlling for the number of couples in the dwelling; that is, in the denominator calculations, couples are considered as one person rather than two. In the definition of PPBR and APPBR, the number of bedrooms does not take into account the number of balconies, bathrooms, foyers, halls and half rooms; it does take into account a room furnished as a bedroom and also includes caravans, sleep-outs and other rooms in cases where they are the only bedroom facilities available in the dwelling. In this definition, a bedroom should include the necessities of a bedroom.

Perceived crowding[48] (PC) reflects people's subjective evaluations of the number of people living in a small area, generally their dwelling. Since perceptions are subjective, these may be affected by their socioeconomic status. For example, people with a higher income level may not appreciate living in (an objectively) crowded space as much as those with a lower income level. Stokols (1972) argues that one person's understanding of his or her crowding level depends on his/her relative intensity of spatial, social and personal factors and the possibility of changing them.

Depending on the overall interpretation of crowding amongst a nation's citizens, measurement of it may vary from one society to another. Thus, a crowding index may be constructed based on specific aspects of the living environment in a region but measured on other aspects in another region. The relevance of the people's evaluations and their cultural backgrounds is well-known amongst philosophers as 'cultural relativism'. For example, sleeping in the living area is prevalent in Japanese culture but not common in some other nations. Thus, households with the same characteristics living in similar living environments

[47] This measure is also called room density. To avoid confusion, I refer to it in this study as people per bedroom.

[48] This measure is also called perceived density. To avoid confusion, I refer to it in this study as perceived crowding.

may be considered crowded in one country but not in another. The variation in the definition of indexes may even occur amongst small geographic scales, such as amongst municipalities of a region.

PC reflects objective crowding and a range of socioeconomic factors. Some of the previously mentioned objective measures do not account for the composition of households or only account for couple status. The CNOS does consider household composition; in this measure, children under 18 years old may share a bedroom if they are of the same sex.[49] In New Zealand, crowding standards are defined based on the Housing Improvement Regulations 1947 (HIR) (Yoshikawa & Ohtaka, 1989). Based on the HIR, the household is considered to be crowded if it violates some well-defined conditions: children between one and 10 years old are counted as half a person and the number of people per bedroom should not exceed 2.5. Children more than 10 years of age may share a room only if they are of the same sex. As each person requires a specific amount of floor area, the size of bedrooms is important in this definition of household crowding.

In a study of Pacific Island households in Auckland, Schluter, Carter and Kokaua (2007) compared PC, PPR and CNOS to report the best measure of crowding. That study did not take into account the PPBR and APPBR measures. They used data derived from interviews with mothers of a cohort of infants born during 2000 in Auckland, which yielded a total of 1224 observations. The dependent variable in their study is the overall satisfaction with the home meeting the family's needs, which is measured on a Likert scale base. Using a logistic method and controlling for demographics, smoking status and house type (residential house, townhouse, unit and caravan), they found PC to be the best measure for crowding. However, the findings of that study should be investigated further by using a population-representative sample with a wider range of explanatory variables to control for different measurement; I do this in the remainder of this chapter.

Rodgers (1982) studied the relationship between satisfaction with residential environment and the measures of perceived crowding and environmental density. He used 1970 observations of census data at the tract level collected in 1970 in the Detroit metropolitan area. He also used interview data from 1194 respondents during 1974 and 1975, with a response rate of 70 per cent. However, the definition of crowding in his study is different from the current study. While I consider household crowding as an objective as well as a subjective concept, Rodgers considered household crowding as a subjective matter. Rodgers (1982) also considered both objective measures of crowding and the measure of density as the objective measure of

[49] As described, the definition of crowding differs amongst countries depending on the culture, the policies of policy makers and the structure of the housing market. Consequently, besides PPBR, as an objective measure of crowding, I account for PC, which is measured by the subjective evaluation of the interviewee about the size of his/her dwelling. The house is supposed to be more crowded if an interviewee says his or her dwelling is too small. As a person states the size of a dwelling by considering many aspects, including cultural ones, this measure is considered as a subjective crowding measure. Besides PC and the objective measures, including PPR, PPBR and APPBR, I use CNOS as an adjusted crowding measure that accounts for some characteristics of individuals, such as partnership status and the number of children.

crowding.[50] Furthermore, the satisfaction measure in Rodgers' study consists of three measures: satisfaction with community, satisfaction with neighbourhood and satisfaction with dwelling unit. Therefore, caution should be taken regarding difference in definitions when comparing the results of Rodgers' study with the current study. Rodgers concluded that the relationship between satisfaction measures and PC is stronger than the correlation between objective density measures (including objective crowding measures). Rodgers (1982) also concluded that most of the predictive power of objective measures is not explained by perceived crowding. He introduced socioeconomic factors, such as income levels, as the main reason for the difference between the explanatory power of objective measures of crowding and the subjective one. In the current study, I control for the impact of socioeconomic factors in my analysis of the predictive power of crowding measures for residential satisfaction.

6.3 Hypothesis development

Satisfaction as a subjective matter depends on both objective and subjective measures. However, in most urban economics studies, this point is not clearly recognised. In the present study, to the extent that the available dataset makes it possible, I made a distinction between objective crowding and perceived crowding after controlling for identification problems. On the basis of previous studies, I hypothesise that:

H1: Higher residential satisfaction is affected more by a perception of less crowding than by a lower value of the raw (objective) crowding measure.

Summary

Household crowding can be measured using a number of objective and subjective measures. An objective measure of crowding is the people per bedroom measure (PPBR), which is the number of people living in each bedroom. The Canadian National Occupation Standard (CNOS) is a measure of crowding that accounts for cultural aspects and the characteristics of the household. Therefore, CNOS is an adjusted measure of crowding. The perceived crowding measure (PC) is a subjective measure of crowding that is based on people's perception of the crowding of their dwelling. In the current study, I hypothesise that higher residential satisfaction is affected more by a perception of less crowding than by a lower value of the raw (objective) crowding measure.

6.4 Data and sample

As discussed in Section 3.1, this study was conducted using three waves of the NZGSS – 2008, 2010 and 2012 – each of which has approximately 8500 observations. The total number of observations in the Auckland region is 5832, with 19,901 observations from the rest of New Zealand (the latter are not included in the main part of this study's analysis). In the survey, respondents answer questions about their usual

[50] However, I distinguish between objective measures of crowding, such as PPBR, and the measure of density. For a definition of density measures see Section 2.3.1.

residence. Therefore, they are considering their primary residence in their evaluation of their living environment.

6.5 Descriptive statistics

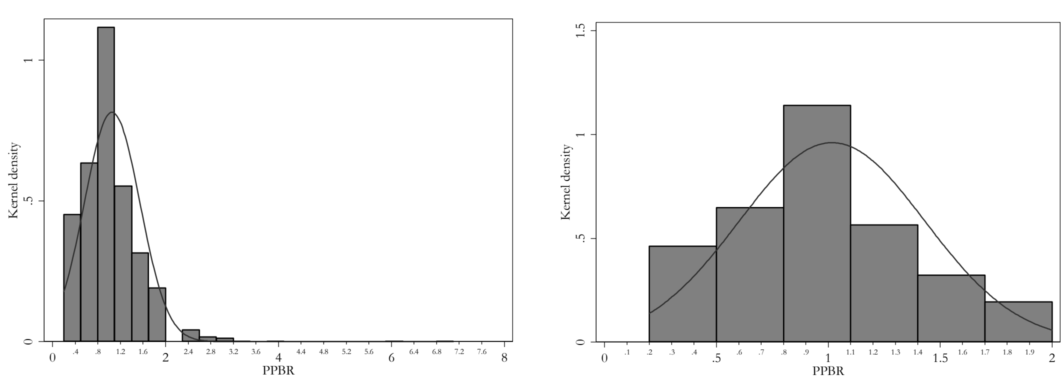
This section describes the variables that are specifically used for answering the research question, explained in Section 6.3. The other variables used in this study are described in Section 3.3.

PPBR, which is one of the variables of interest, is derived by dividing household size by the number of bedrooms in the dwelling. As illustrated in Table 9, PPBR takes values between 0.2 and 8. The mean difference amongst years is not significant and the overall mean is equal to 1.05. As depicted on the left-hand side of Figure 12, 54 per cent of observations have a PPBR less than 0.8, 25 per cent are equal to 1 and 25.12 per cent are greater than 1 and less than 2.1. To derive a more reliable measure, I dropped the outliers that had a reported PPBR between 2.1 and 8 (this only represented 101 observations, which is less than 1.8 per cent of the sample population). The kernel density plot by imposing this restriction on the PPBR variable is illustrated on the right-hand side of Figure 12, where the mean is equal to 1.02.

Table 9. PPBR summary statistics.

	Observations	Mean			P_value[51]	Min	Max	Std error
		2008	2010	2012				
Unrestricted PPBR	5816	1.041	1.081	1.041	0.158	0.2	8.0	0.004
Restricted PPBR	5715	1.006	1.038	1.006	0.171	0.2	2.0	0.003

Figure 12. The distribution of PPBR.



*** On the left-hand side, the kernel density is illustrated before omitting the outliers. The figure on the right-hand side depicts the kernel density after restricting the variable to 98.2 per cent of the sampling population. The horizontal axis is the number of people per bedroom.**

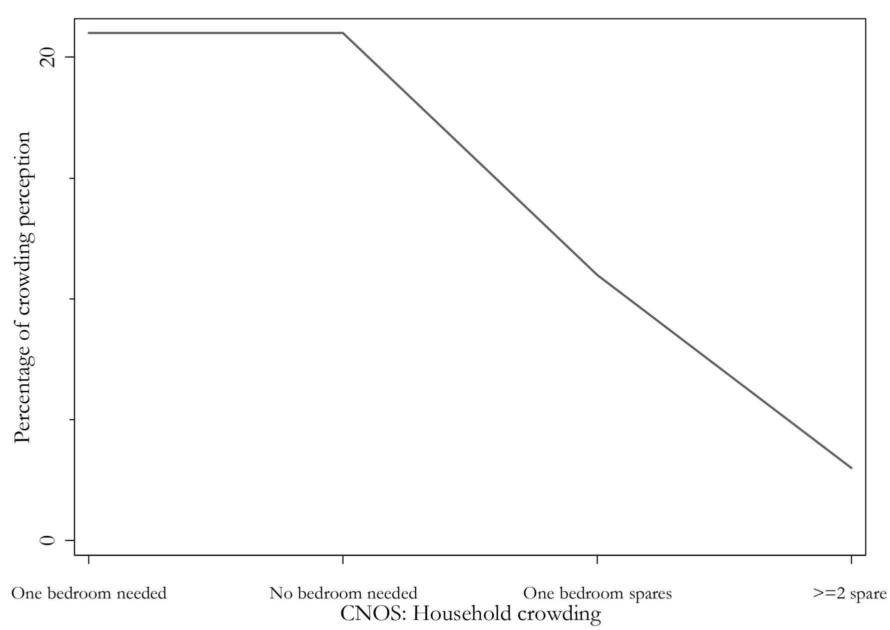
[51] This P_value corresponds to the test of the null hypothesis that of the means of this variable across three waves of data, namely 2008, 2010 and 2012, are equal.

In Figure 13, PC is tabulated against CNOS. As expected, in dwellings with more spare rooms, people perceive less of a small-house problem (PC). It is interesting that where the index reports no need for an extra bedroom, 21 per cent of respondents still perceive their house to be small. Also, where the CNOS measure indicates the need for one extra bedroom, 79 per cent of respondents are satisfied with the size of their dwellings.

Another point to be considered is the association between CNOS and household size. As illustrated in Table 10, as the household composition is counted as a component of CNOS, there is a negative correlation between household composition and CNOS. In other words, as the household size increases, CNOS indicates the need for extra bedrooms.

As shown in equation (21), PPBR is measured as a ratio of household size. In Table 11, PPBR is tabulated versus household size. Since PPBR is a continuous measure, it is categorised into six groups based on its distribution (Abas, Vanderpyl, Robinson, Le Prou, & Crampton, 2006). As depicted, amongst one-person families, PPBR is mainly between 0.3 and 0.6. The positive correlation between PPBR and household size suggests that the members of a large household are not likely to be provided with a commensurate number of bedrooms. In Table 12, PPBR is tabulated versus PC. As shown, the percentage of people who perceive their house to be small increases as PPBR increases.

Figure 13. Canadian National Occupation Standard (CNOS) versus perceived crowding (PC).



Source: Torshizian (2017).

Table 10. Household size versus CNOS.

	Household crowding (CNOS)

Household size	One bedroom needed	No bedrooms needed	One bedroom spare	Two or more bedrooms spare
One person	0.00	0.19	0.34	0.48
Two people	0.01	0.08	0.26	0.65
Three people	0.02	0.20	0.48	0.30
Four people	0.05	0.24	0.38	0.33
Five people	0.11	0.43	0.31	0.15
Six people	0.25	0.36	0.27	0.12
Seven people	0.26	0.34	0.30	0.10
Eight people	0.40	0.31	0.20	0.10
All households	0.14	0.27	0.32	0.28
N	5715			

Note: Rows sum to 1.0; decimals are rounded.

Table 11. Household size versus PPBR.

Household size	Household crowding (PPBR)					
	0< PPBR ≤ 0.3	0.3< PPBR ≤ 0.6	0.6< PPBR ≤ 0.9	0.9< PPBR ≤ 1.3	1.3<PPBR≤ 2	2<PPBR
One person	0.12	0.69	0.00	0.19	0.00	0.00
Two people	0.00	0.24	0.48	0.24	0.00	0.04
Three people	0.00	0.00	0.33	0.53	0.13	0.00
Four people	0.00	0.00	0.14	0.38	0.41	0.07
Five people	0.00	0.00	0.00	0.63	0.37	0.00
Six people	0.00	0.00	0.00	0.28	0.34	0.38
Seven people	0.00	0.00	0.00	0.00	1.00	0.00
Eight people	0.00	0.00	0.00	0.00	0.73	0.27
N	5715					

Note: Rows sum to 1.0; decimals are rounded.

Table 12. PPBR versus PC (small-house problem).

People per bedroom	PC: Small-house problem	
	No	Yes
0 < PPBR ≤ 0.3	0.98	0.02
0.3 < PPBR ≤ 0.6	0.97	0.03
0.6 < PPBR ≤ 0.9	0.97	0.03

0.9 < PPBR ≤1.3	0.90	0.10
1.3 < PPBR ≤ 2	0.81	0.19
2 < PPBR	0.68	0.32
N	5715	

Note: Rows sum to 1.0; decimals are rounded.

The APPBR calculations control for couple status, so I expect lower levels of crowding at each household size level compared to the PPBR measure. This is shown in Table 13, where APPBR is tabulated versus the household size. Similar to the results derived from the tabulation of PPBR versus APPBR, the tabulation of APPBR versus PC shows that the percentage of people who perceive a small house problem is increasing as the objective crowding measure (APPBR) increases (see Table 14).

Table 13. Household size versus adjusted people per bedroom (APPBR).

Household size	Household crowding (APPBR)					
	0< APPBR ≤ 0.3	0.3< APPBR ≤ 0.6	0.6< APPBR ≤ 0.9	0.9< APPBR ≤ 1.3	1.3< APPBR ≤ 2	2< APPBR
One person	0.12	0.69	0.00	0.19	0.00	0.00
Two people	0.21	0.58	0.10	0.10	0.00	0.01
Three people	0.00	0.21	0.45	0.30	0.04	0.00
Four people	0.00	0.00	0.36	0.43	0.18	0.02
Five people	0.00	0.00	0.09	0.54	0.37	0.00
Six people	0.00	0.00	0.00	0.52	0.36	0.13
Seven people	0.00	0.00	0.00	0.26	0.74	0.00
Eight people	0.00	0.00	0.00	0.00	0.83	0.17
N	5715					

Note: Rows sum to 1.0; decimals are rounded.

Table 14. Adjusted people per bedroom (APPBR) versus PC.

Adjusted people per bedroom (APPBR)	PC: Small-house problem	
	No	Yes
0 < APPBR ≤ 0.3	0.99	0.01
0.3 < APPBR ≤ 0.6	0.94	0.06
0.6 < APPBR ≤ 0.9	0.92	0.08

0.9 < APPBR ≤1.3	0.87	0.13
1.3 < APPBR ≤ 2	0.80	0.20
2 < APPBR	0.68	0.32
N	5715	

Given the multiplicity of available crowding measures, which tend to show similar information, a principal component approach can be taken to summarise these measures of crowding (APPBR is excluded from calculation of this measure owing to its close relationship with PPBR). Based on the covariance matrix, the corresponding eigenvector for the first principal component is as follows:

Crowding Component (CC) = 0.649PPBR + 0.402PC – 0.646CNOS

6.6 Research design

Our research aims to test which crowding measure(s) has the greatest explanatory power for determining residential satisfaction (RS) after controlling for other influences on RS. The dataset provides us with three cross-sections. The equation to be estimated is as follows:

$$y_i = \alpha_0 + \alpha_1 . x_i \ + \alpha_2 . x_i^2 \ + \alpha_3 . w_i \ + \beta . \text{Controls}_i + \varepsilon_i \qquad (21)$$

in which the dependent variable (y_i) is individual i's residential satisfaction (measured as a binary variable), and x_i is a vector of our variables of interest, including PPBR, PC and CC. Given that I control for marital status, I do not include APPBR as a variable of interest separate from PPBR. x_i^2 is the squared form of x_i, which is used in relation to the raw crowding measure (PPBR). The assumption of a quadratic relationship between the measures of crowding and residential satisfaction is consistent with Rodgers' (1982) findings that these relationships are better described in a curvilinear form than a linear one. Given the categorical structure of our other variable of interest (CNOS), it is included in the equation as a vector of dummy variables, depicted by w_i. Thus, α_1 and α_2 are the parameters of interest; and ε_i is the error term. Control variables are explained in the next section. The responses to residential satisfaction are an ordinal limited dependent variable. The appropriate models to be used are logit and probit, in the case of a binary dependent variable, and ordered logit and ordered probit, in the case of a non-dummy categorical dependent variable. As explained in Appendix ii, the logistic regression[52] is used for ease of interpretation in this study's estimates.

6.6.1 Control variables

In addition to the variables of interest, the effects of other factors on RS need to be controlled for. In equation (22), the control variables are indicated by Controls_i. The control variables are chosen based on a hypothesis that they may have an independent effect on the dependent variable. Since two single people

[52] The logistic function is the inverse of the logit.

usually need two separate rooms in a house, but couples do not, marital status is one of the control variables. A higher PC may derive from dissatisfaction with the living environment (Chan, 1999); for example, a person may consider his or her dwelling as small if he or she has some problems with his or her living environment. Thus, the impact of problems both with the dwelling and the neighbourhood should be controlled for. The housing problems include difficult access from the street, poor condition, damp, cold, having pests and being expensive. Neighbourhood problems include being located far from work, located far from other destinations, being unsafe, having noise or vibration and having air pollution. The other control variables included in the estimations are named and described in Section 3.3.

Summary

Out of 5832 observations derived from three cross-sections of the General Social Survey for the Auckland region, 5715 observations are used in this study. The primary analysis depicts a difference between people's understanding of their crowding levels and the figures derived from the indices of crowding, such as CNOS. The members of a large household are not likely to be provided with a commensurate number of bedrooms. For a better understanding of the best measure of crowding, I constructed a crowding component (CC) by taking a principal component approach. This study uses a logit model for its analyses of the impacts that the variables of interest – which are PPBR, CNOS, PC and CC – have on RS.

6.7 Models and results

In order to compare the results derived from including PPBR, CNOS, PC and CC, eight models are included in Table 15. For the sake of brevity, the table does not report the impact of the control variables (the complete table and the impact of control variables are presented in Appendix iii). The coefficients are reported in the exponential form; that is, the impact is negative if the coefficient is less than 1 and it is positive if the coefficient is greater than one.[53]

The same control variables are included in models 1 to 8. In the first three models, each of the variables of interest is included separately. Of these models, based on the GOF criteria (AIC and pseudo R-squared), PC has the highest predictive power for RS. CNOS has slightly higher predictive power than PPBR. Since I have assumed a quadratic form for the models that include PPBR, the statistical significance of the coefficients of the variable and its squared form ($PPBR^2$) are reported jointly at the bottom of the table in the '†PPBR Wald test' row.

The next three columns include three combinations of two variables at a time; namely, PPBR and PC, PPBR and CNOS, and PC and CNOS. Model (7) incorporates all three crowding variables. Of these, model

[53] Appendix II presents the difference between exponentiated and non-exponentiated forms.

6 is the best model in terms of the GOF. The simultaneous presence of CNOS in the equations is associated with a lower statistical significance of PPBR (see models 5 and 7).

Model 8 includes the crowding principal component (CC). As expected, according to t-statistic, R-squared and AIC, this component provides a more precise prediction than PPBR and CNOS. However, this model has less predictive power than any model that includes PC, such as models 2, 4, 6 and 7.

Based on the AIC and R-squared, the best fit amongst all of the eight models is obtained in model 6 of Table 15, which includes PC and CNOS. PC is included in all of the best performing models, which confirms the important association between subjective evaluations of crowding and RS. A model including all control variables but no crowding variable has a predictive power of 0.1688.

CNOS has a higher predictive power than the raw measure (PPBR) but the difference between them is very slight. Therefore, while PPBR does not have quite as much explanatory power as PC or CNOS, the results across the three variables provide a very similar picture. PPBR is the simplest of the three measures to construct and is the variable used henceforth as the prime measure of crowding.

Table 15. Models of people per bedroom, subjective density (small-house perception) and CNOS.

	Model(1) PPBR	Model(2) PC	Model(3) CNOS	Model(4) PPBR &PC	Model(5) PPBR &CNOS	Model(6) PC &CNOS	Model(7) PPBR & PC & CNOS	Model(8) CC	Models 1,2,3,8 M.E. (dy/dx)
RS (dummy)									
PPBR	0.762†			0.802	1.591		1.361		-0.099
	(0.386)			(0.408)	(0.909)		(0.789)		
PPBR ^{2}	0.913†			0.973	0.835		0.929		
	(0.222)			(0.242)	(0.217)		(0.250)		
PC		0.199***		0.212***		0.218***	0.217***		-0.274***
		(0.041)		(0.044)		(0.046)	(0.045)		(-7.57)
CNOS (Two or more spare bedrooms)									
= One bedroom needed			0.277***		0.270***	0.344***	0.300***		-0.228**
			(0.093)		(0.107)	(0.116)	(0.121)		(-3.14)
= No bedrooms needed			0.674***		0.639**	0.821	0.740		-0.0421
			(0.095)		(0.131)	(0.115)	(0.152)		(-1.40)
= One bedroom spare			0.676***		0.644***	0.736***	0.696***		-0.0653**
			(0.064)		(0.082)	(0.069)	(0.088)		(-3.25)
CC								0.784***	-0.053***
								(0.031)	
Observations	4706	4706	4706	4706	4706	4706	4706	4706	
Population size	925815	925815	925815	925815	925815	925815	925815	925815	
†PPBR Wald	0.003	.	.	0.103	0.711	.	0.725	.	

test								
Pseudo R^2	0.1723	0.1860	0.1742	0.1872	0.1743	0.1888	0.1888	0.1778
AIC	5184	5099	5175	5096	5178	5088	5092	5148

Reported coefficients are exponentiated ones. Standard errors are reported in parentheses. Marginal effects are calculated based on the model (7). For marginal effects t-statistics are reported in the second row.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

† The significance level for PPBR and its squared (Squared PPBR) variables is reported jointly on the PPBR Wald test row. The reported AICs are for the models that do not take into account jackknife replications. The base group for each variable is written in parentheses. For example, for the Age variable, the 'more than 75 years old' category is supposed as the base group.

Based on model 2, and as illustrated in the last column of the table in the marginal effects column, people who perceive their house to be too small are almost 27 per cent less likely to be satisfied with their residential environment. Using our objective measure, as depicted in Figure 14, a higher PPBR leads to lower residential satisfaction. A unit increase in PPBR results in roughly a 9 per cent decrease in RS;[54] that is, more crowded households are more likely to have lower residential satisfaction. The curvature of the probability curve is derived from the nonlinearity assumption (equation (22)). The partial effect of PPBR on the probability of being satisfied, P(RS=1|x), depends on its coefficient (α_1) and the coefficient of its squared form (α_2). The turning point in the quadratic function is at $(-\frac{\alpha_1}{2\alpha_2})$, which is approximately equal to -1.45 (derived from $\frac{\log(0.762)}{2\log(0.913)}$); which means that, as PPBR is always greater than 0, the correlation between RS and PPBR is always negative.

As depicted in Figure 15, based on estimations in model 3, the probability of being dissatisfied for those who need one or more bedrooms is 15.3 per cent higher than for those with two or more spare bedrooms. As both of these figures show, the decreasing trend of the likelihood of residential satisfaction is ascending. This implies that an increase in household crowding decreases the likelihood of satisfaction of those who already live in a crowded household more than those who live in a less crowded house.

[54] The reported marginal effect of PPBR accounts for the non-linear relationship between RS and PPBR.

Figure 14. The probability of being satisfied amongst different PPBR bins (derived from Model 1).

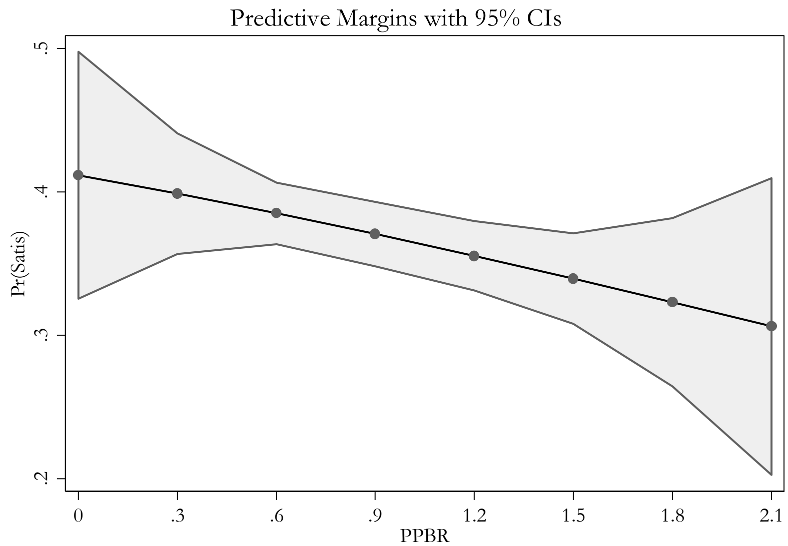
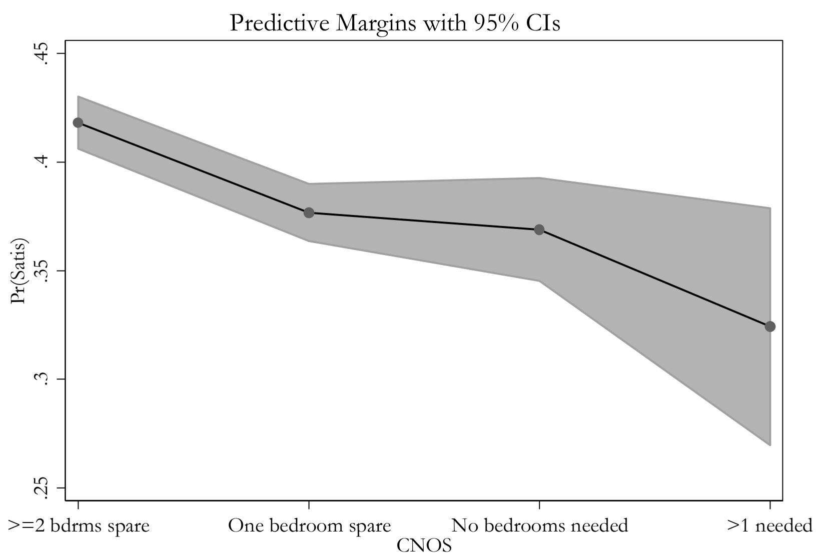


Figure 15. The probability of being satisfied amongst different categories of CNOS.



At the upper-left hand side of Figure 16, the probability of being satisfied is compared between people who have stated that their dwelling is too small and those who have not stated this. This suggests a significantly higher residential satisfaction for those with a positive perception towards their living environments than those with who perceive their house to be small. The lower part of the figure illustrates the difference between these two. An upward slope of this difference curve suggests that the probability of being residentially satisfied for people who perceive their house to have adequate space compared to those who do not increases for higher levels of household crowding, as gauged by PPBR.

The probability of being satisfied for people with different perception towards their dwelling is depicted versus the adjusted crowding measure (CNOS) at the upper-right hand side of Figure 16. The difference between the impact of positive and negative subjective evaluation of the size of the house on the probability of residential satisfaction remains almost constant amongst different objective crowding levels; that is, people with a positive perception towards the size of their dwelling are almost 25 per cent more likely to be satisfied with their residential environment at any household crowding level, gauged by CNOS.

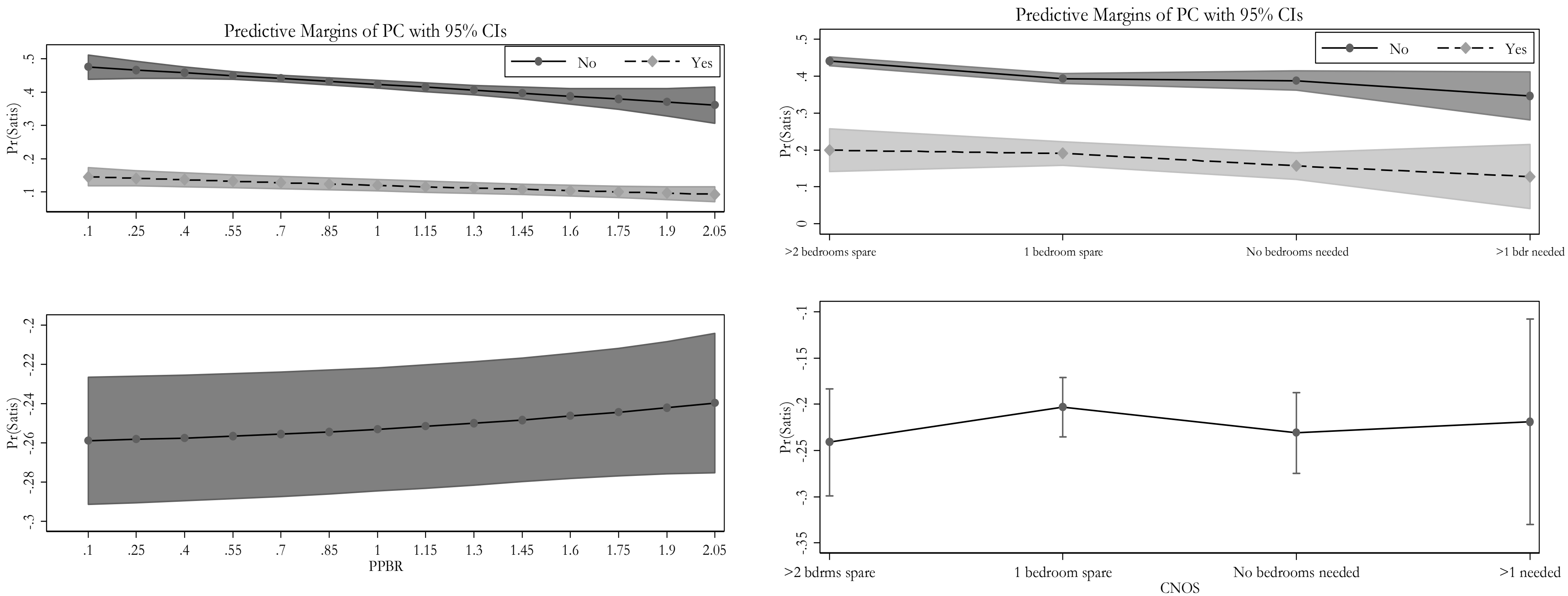


Figure 16. Perceived crowding versus the adjusted crowding measure (CNOS).

The last column of Table 15 includes the marginal effects of each included variable based on models 1, 2, 3 and 8. For example, the marginal effect of PPBR on RS is calculated based on model 1 and the marginal effect of CC is reported based on model 8. Since PPBR is included as a squared form, the reported marginal effect is measured jointly. A one-unit increase in PPBR is associated with a 9.9 per cent decrease in the likelihood of being satisfied.

In order to assess how closely the crowding variables are related, the results of regressing one on another are illustrated in Table 16. In model 1, PC is regressed on PPBR and its squared form using a logit method, while model 2 regressed the variables using an OLS equation. The results indicate that a one-unit increase in the explanatory variable (PPBR) results in a 10.6 per cent increase in the likelihood of reporting a higher PC.[55] In the third model, CNOS is regressed on PPBR and its squared form by using a generalised ordered logit estimation. Based on the marginal effects, a one-unit increase in PPBR is associated with a 14.5 per cent higher CNOS. Based on the fourth model, 60.4 per cent of the variation of PPBR is explained by CNOS. The results of this table suggest that the crowding variables significantly explain one another and therefore all indicate the same thing: crowding.

Table 16. Relationships between the variables of interest; namely, PC, PPBR and CNOS.

Dependent variable	(1) PC Logit	(2) PPBR OLS	(3) CNOS Generalised ordered logit			(4) PPBR OLS	(5) PC Logit
			Need bdrs	No need	1 spare		
PPBR	3.386***		-26.7***	-4.980***	2.709		
	(0.635)		(3.501)	(0.692)	(2.054)		
Squared PPBR	-0.630**		6.854***	-0.613*	-5.48***		
	(0.261)		(1.053)	(0.313)	(1.402)		
Marginal effect (%)	0.106***			0.145***			
PC		0.321***					
		(0.025)					
Household Crowding - CNOS (Two or more spare bedrooms)							
= One bedroom needed						-0.317***	-0.159
						(0.028)	(0.276)
Marginal effect (%)							-0.0277
= No bedrooms needed						-0.705***	-0.828***
						(0.024)	(0.287)
Marginal effect (%)							-0.118**
= One spare bedroom						-1.048***	-2.532***
						(0.024)	(0.309)
Marginal effect (%)							-0.214***
Constant	-5.037***	0.968***	26.30***	7.550***	1.880***	1.733***	-1.163***
	(0.361)	(0.007)	(2.894)	(0.365)	(0.688)	(0.023)	(0.260)
Observations	4706	4706		5715		5715	4706
N_pop	925815.0	925815.0		1119973.0		1119973.0	925815.0
F	94.06	171.2		374.2		1307.4	55.40
R^2		0.0578				0.604	

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
In all models, based on the Wald test, the joint probability of PPBR and its squared is statistically significant at the 0.001 level.

[55] The joint marginal effect of PPBR and $PPBR^2$ is reported on a row below 'squared PPBR'.

All R-squared coefficients are reported based on OLS estimations. The ME rows include the marginal effects of the variable in the upper row; for example, in model 5, the ME row includes the marginal effect of the CNOS variable.

I also tested whether there is a difference in the impact of perceived crowding (PC) on the residential satisfaction of people with different demographics. After including interaction terms, I found no significant difference between the impact of PC on RS of women versus men, rich versus poor, or elderly versus youth. In addition to the impact of the variables of interest, presented in the current section, I have discussed the impact of the controlling variables in Appendix iii.

6.8 Conclusion

This chapter hypothesises that residential satisfaction is affected more by a perception of crowding than by objective crowding level. To test this hypothesis, this chapter attempts to find the best measure of household crowding. Residential satisfaction is a measure of both objective aspects and subjective characteristics of the living environment, where the latter includes cultural perspectives on crowding. Hence, the PC measure is used in this chapter for a subjective measurement of household crowding. Also, the objective aspects of residential satisfaction are determined by using objective measures. The CNOS is used as an adjusted crowding measure that accounts for some characteristics of individuals, such as partnership status and the number of children. I also combined these measures by using a principal component approach. The short list of the crowding measures and the factors they account for are presented in Table 8. PPBR is calculated as follows:

$$\text{PPBR} = \frac{\text{Number of people in the dwelling}}{\text{Number of bedrooms}}$$

The impact of the variables of crowding on residential satisfaction is estimated based on a quadratic function by fitting a logistic model for our binary response outcome, residential satisfaction. Controls for personal characteristics are included in all regressions. The results indicate that an increase in household crowding, measured based on any measure of crowding, lowers residential satisfaction significantly. Those who already live in a crowded household are affected more negatively by an increase in their crowding level than those who live in a less crowded house.

The lower predictive power of the raw and adjusted raw measures of crowding compared to the perceived measure suggests that perceptions are an important manifestation of people's evaluations of their household crowding. Also, those people who have a positive perception towards their house size have a significantly higher residential satisfaction than those who perceive their house to be small.

All three measures of household crowding – PPBR, CNOS and PC – are significantly related to each other; in other words, each gives a similar qualitative impact on residential satisfaction with all measures of greater crowding impacting negatively on residential satisfaction. This suggests that PPBR remains a useful predictor of residential satisfaction in the absence of the other measures and, given its wide availability, is used as the measure for household crowding in the following chapters.

Another reason for using an objective measure of crowding in the next chapters is the potential issue with including subjective measures on both sides of an equation. As presented in Section 5.4, Hamermesh (2004) discussed that the estimated impact of a subjective measure (PC) on another subjective measure (RS) may be biased as both of these measures are subject to a measurement error. As discussed in 5.4, in measurement of the impact of the objective measure of crowding on RS, I have used two control variables – satisfaction with health status and satisfaction with abilities and skills – to overcome the potential measurement issues of RS. Respondents were asked about both of the control variables prior to answering the residential satisfaction question. Therefore, I expected the control variables to only capture the impact of any potential measurement error.

The results of this chapter confirm Bonnes et al.'s (1991) results that the relationship between crowding and residential satisfaction is negative. The results also confirm the importance of perceived crowding compared to objective measures of crowding, as concluded by Rodgers (1982).[56] Rodgers suggested that, for a better understanding of the predictive power of objective crowding measures, one must account for a wide range of socioeconomic factors. While I have done this, the measures of crowding used in the current study are different from those in previous studies.[57]

The minor results of this chapter are similar to the findings of the previous hedonic pricing studies. For example, as discussed in Appendix iii, consistent with Eichholtz and Lindenthal's (2014) results, a higher level of income, better health status and becoming older are all positively correlated with residential satisfaction. While the results of that study indicate a positive relationship between education levels and the demand for housing, I found no significant relationship between education levels and residential satisfaction.

Due to computational limits, I used Auckland's sample population in all chapters of this thesis. However, the estimations based on Auckland's sample population derive the same results as the estimations using New Zealand's sample population, as presented in Appendix i.

A future study could address measurement issues with the components of the constructed measures. For example, the PPBR measure is constructed using the number of people and the number of bedrooms, both of which may be subject to measurement error.

An extension to this study could disaggregate the current findings for different groups. For example, an interesting question is if a specific ethnic group exhibits a different relationship between PC and PPBR compared to the other ethnic groups (see for example Schluter, Carter, and Kokaua (2007)). This is an assessment of the cultural relativism hypothesis that I discussed in section 6.2.1.

---

[56] As discussed in 6.2, however, Rodgers' measure of objective crowding is very different from the current study's measure of objective crowding.

[57] For details see 6.2.

As described in section 6.5, the current study omits the outliers that had a reported PPBR between 2.1 and 8 (this only represented 101 observations, which is less than 1.8 per cent of the sample population). I have described the changes in the sample population but the characteristics of the outliers may be of interest to a future study.

Appendix I. Results derived from the New Zealand sample

In this appendix, the same equations that were described in Section 6.6 are estimated based on New Zealand's sample population. The equations presented in Table 15 are identical to the equations presented in Table 15 and Appendix iii.

Table 17. Models of PPBR, PC and CNOS using New Zealand's sample population.

	Model(1) PPBR	Model(2) PC	Model(3) CNOS	Model(4) PPBR & PC	Model(5) PPBR & CNOS	Model(6) PC &CNOS	Model(7) PPBR & PC & CNOS	Model(8) CC	Marginal Effect (dy/dx)
Residential Satisfaction (dummy)									
PPBR[1]	0.754†			0.766†	1.353†		1.171†		-0.065
	(0.174)			(0.179)	(0.372)		(0.323)		(-4.28)
Squared PPBR	0.915†			0.991†	0.823†		0.919†		
	(0.102)			(0.114)	(0.100)		(0.116)		
Perceived Crowding		0.254***		0.272***		0.276***	0.278***		-0.242
		(0.021)		(0.023)		(0.024)	(0.024)		(-21.9)
Household Crowding (Two or more bedrooms spare)									
= One bedroom needed			0.457***		0.541***	0.566***	0.591**		
			(0.077)		(0.112)	(0.098)	(0.127)		
= No bedrooms needed			0.642***		0.698***	0.770***	0.782**		
			(0.047)		(0.077)	(0.058)	(0.086)		
= One bedroom spare			0.743***		0.751***	0.806***	0.802***		
			(0.033)		(0.049)	(0.036)	(0.053)		
Crowding Component								0.809***	
								(0.015)	
Gender[3]	0.834***	0.830***	0.834***	0.827***	0.833***	0.827***	0.827***	0.829***	-0.043
	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)	(-4.57)
Partner	1.131**	1.061	0.976	1.129**	0.994	1.026	1.025	1.079	0.028
	(0.064)	(0.056)	(0.051)	(0.065)	(0.064)	(0.055)	(0.067)	(0.057)	(2.15)
Age (>74)									
= 15–19[5]	0.989	0.843	0.883	0.961	0.888	0.893	0.888	1.024	-0.010
	(0.170)	(0.140)	(0.146)	(0.168)	(0.153)	(0.151)	(0.156)	(0.172)	(-0.23)
= 20–24	0.757**	0.666***	0.699***	0.741***	0.702***	0.704***	0.700***	0.787**	-0.073
	(0.085)	(0.075)	(0.080)	(0.084)	(0.080)	(0.080)	(0.080)	(0.090)	(-2.69)
= 25–34	0.540***	0.523***	0.519***	0.561***	0.521***	0.547***	0.545***	0.566***	-0.136
	(0.055)	(0.053)	(0.053)	(0.057)	(0.053)	(0.056)	(0.055)	(0.058)	(-5.65)
= 35–54	0.535***	0.519***	0.520***	0.550***	0.521***	0.539***	0.538***	0.557***	-0.140
	(0.043)	(0.041)	(0.041)	(0.044)	(0.041)	(0.043)	(0.043)	(0.044)	(-7.15)
= 55–59	0.737***	0.758***	0.731***	0.757***	0.728***	0.750***	0.749***	0.741***	-0.068
	(0.064)	(0.068)	(0.064)	(0.067)	(0.064)	(0.067)	(0.066)	(0.065)	(-3.14)
= 60–64	0.827**	0.843*	0.816**	0.835**	0.814**	0.826**	0.825**	0.819**	-0.044

	(0.073)	(0.075)	(0.072)	(0.074)	(0.072)	(0.074)	(0.073)	(0.073)	(-2.02)
= 65–69	0.854*	0.876	0.841**	0.866*	0.839**	0.855*	0.854*	0.844**	-0.035
	(0.072)	(0.074)	(0.071)	(0.073)	(0.071)	(0.073)	(0.073)	(0.072)	(-1.69)
= 70–74	0.864	0.872	0.851*	0.866	0.850*	0.855*	0.855*	0.852*	-0.036
	(0.079)	(0.081)	(0.079)	(0.080)	(0.079)	(0.079)	(0.079)	(0.079)	(-1.57)
Length of living in NZ (<4 years)									
= 4–10 years	0.972	0.994	1.014	0.987	1.001	1.015	1.011	0.997	-0.003
	(0.130)	(0.135)	(0.137)	(0.133)	(0.133)	(0.138)	(0.135)	(0.134)	(-0.10)
= 10–25 years	0.850	0.870	0.906	0.862	0.890	0.898	0.894	0.886	-0.033
	(0.117)	(0.122)	(0.125)	(0.121)	(0.125)	(0.127)	(0.127)	(0.123)	(-1.06)
>25 years[15]	0.972	1.023	1.011	1.003	0.998	1.028	1.023	0.992	0.001
	(0.099)	(0.108)	(0.105)	(0.105)	(0.103)	(0.110)	(0.108)	(0.103)	(0.03)
Ethnicity (European)									
= Maori	0.916	0.873**	0.914	0.901*	0.920	0.903	0.905	0.937	-0.024
	(0.057)	(0.055)	(0.057)	(0.056)	(0.057)	(0.056)	(0.056)	(0.058)	(-1.71)
= Pacific	0.779*	0.690***	0.761*	0.747**	0.774*	0.742**	0.746**	0.815	-0.065
	(0.108)	(0.095)	(0.106)	(0.104)	(0.109)	(0.104)	(0.106)	(0.112)	(-2.19)
= Asian	0.661***	0.619***	0.653***	0.639***	0.658***	0.637***	0.639***	0.666***	-0.097
	(0.086)	(0.079)	(0.084)	(0.083)	(0.085)	(0.082)	(0.083)	(0.087)	(-3.72)
= Other	1.009	0.977	1.012	0.994	1.013	0.997	0.998	1.024	-0.001
	(0.104)	(0.102)	(0.104)	(0.104)	(0.104)	(0.104)	(0.104)	(0.105)	(-0.05)
Education (No qualification)									
= Diploma[20]	0.990	1.011	0.990	1.000	0.990	0.999	1.000	0.985	0.000
	(0.060)	(0.061)	(0.060)	(0.060)	(0.060)	(0.061)	(0.060)	(0.061)	(0.00)
= Degree	0.903	0.931	0.909	0.916	0.907	0.919	0.918	0.898*	-0.020
	(0.057)	(0.060)	(0.057)	(0.059)	(0.057)	(0.059)	(0.059)	(0.058)	(-1.35)
Household income ($150,001 or more)									
= Zero	0.854	0.980	0.871	0.920	0.875	0.931	0.936	0.835	-0.021
	(0.257)	(0.283)	(0.254)	(0.272)	(0.255)	(0.271)	(0.271)	(0.250)	(-0.29)
=$1–$5,000	0.877	0.918	0.907	0.881	0.912	0.900	0.905	0.872	-0.031
	(0.345)	(0.357)	(0.365)	(0.343)	(0.363)	(0.354)	(0.354)	(0.355)	(-0.33)
=$5000–$50,001	0.577***	0.611***	0.594***	0.597***	0.596***	0.609***	0.611***	0.583***	-0.120
	(0.046)	(0.048)	(0.047)	(0.048)	(0.048)	(0.048)	(0.049)	(0.047)	(-6.41)
=$50,001–$70,000[25]	0.681***	0.714***	0.690***	0.706***	0.692***	0.712***	0.713***	0.688***	-0.083
	(0.060)	(0.064)	(0.060)	(0.064)	(0.061)	(0.063)	(0.064)	(0.061)	(-3.89)
=$70,001–$150,000	0.720***	0.735***	0.730***	0.729***	0.730***	0.736***	0.736***	0.725***	-0.076
	(0.048)	(0.048)	(0.048)	(0.048)	(0.049)	(0.049)	(0.049)	(0.049)	(-4.69)
Health (Very dissatisfied)									
=Dissatisfied	1.185	1.187	1.193	1.185	1.191	1.190	1.190	1.188	0.035
	(0.185)	(0.189)	(0.183)	(0.190)	(0.184)	(0.187)	(0.188)	(0.186)	(1.08)
= No feeling either way	1.294*	1.304*	1.308**	1.295*	1.305*	1.304*	1.304*	1.296*	0.054
	(0.177)	(0.178)	(0.176)	(0.179)	(0.177)	(0.178)	(0.178)	(0.178)	(1.96)

= Satisfied	1.485***	1.502***	1.498***	1.492***	1.496***	1.498***	1.499***	1.489***	0.086
	(0.199)	(0.202)	(0.197)	(0.203)	(0.198)	(0.201)	(0.201)	(0.200)	(3.16)
= Very satisfied[30]	1.869***	1.865***	1.883***	1.857***	1.884***	1.868***	1.870***	1.870***	0.138
	(0.269)	(0.270)	(0.267)	(0.271)	(0.268)	(0.269)	(0.270)	(0.269)	(4.63)
Abilities (Very dissatisfied)									
= Dissatisfied	0.719	0.667	0.735	0.660	0.729	0.679	0.675	0.705	-0.095
	(0.340)	(0.294)	(0.340)	(0.296)	(0.342)	(0.305)	(0.305)	(0.329)	(-0.93)
= No feeling either way	0.562	0.513	0.576	0.502	0.568	0.516	0.513	0.541	-0.158
	(0.264)	(0.227)	(0.265)	(0.226)	(0.265)	(0.233)	(0.232)	(0.252)	(-1.53)
= Satisfied	0.740	0.673	0.758	0.657	0.749	0.675	0.671	0.710	-0.097
	(0.345)	(0.294)	(0.345)	(0.292)	(0.346)	(0.301)	(0.301)	(0.327)	(-0.94)
= Very satisfied	1.233	1.126	1.270	1.096	1.253	1.132	1.125	1.186	0.021
	(0.569)	(0.484)	(0.570)	(0.481)	(0.571)	(0.496)	(0.496)	(0.539)	(0.21)
Region (Auckland)									
= Wellington[35]	1.000	1.010	1.012	1.006	1.012	1.015	1.015	1.006	0.001
	(0.060)	(0.062)	(0.061)	(0.062)	(0.061)	(0.063)	(0.062)	(0.061)	(0.09)
= Northland group	1.274***	1.265***	1.271***	1.262***	1.274***	1.262***	1.263***	1.263***	0.054
	(0.090)	(0.092)	(0.090)	(0.092)	(0.090)	(0.092)	(0.092)	(0.090)	(3.15)
= Rest of North Island	1.321***	1.336***	1.328***	1.322***	1.327***	1.328***	1.328***	1.313***	0.065
	(0.091)	(0.091)	(0.092)	(0.090)	(0.091)	(0.090)	(0.090)	(0.091)	(4.08)
= Canterbury	0.980	0.981	0.985	0.981	0.985	0.986	0.986	0.985	-0.004
	(0.068)	(0.067)	(0.068)	(0.067)	(0.068)	(0.068)	(0.068)	(0.069)	(-0.28)
= Rest of South Island	1.220***	1.238***	1.230***	1.234***	1.230***	1.242***	1.242***	1.227***	0.048
	(0.085)	(0.088)	(0.086)	(0.087)	(0.085)	(0.087)	(0.087)	(0.086)	(2.97)
Bad street access[40]	0.772	0.821	0.769	0.822	0.774	0.821	0.823	0.783	-0.045
	(0.132)	(0.141)	(0.131)	(0.142)	(0.131)	(0.141)	(0.141)	(0.134)	(-1.13)
Poor condition	0.351***	0.372***	0.349***	0.374***	0.350***	0.371***	0.372***	0.357***	-0.226
	(0.043)	(0.046)	(0.043)	(0.047)	(0.043)	(0.047)	(0.047)	(0.044)	(-7.86)
Damp dwelling	0.568***	0.608***	0.573***	0.608***	0.573***	0.612***	0.611***	0.584***	-0.114
	(0.055)	(0.058)	(0.055)	(0.059)	(0.055)	(0.059)	(0.059)	(0.056)	(-5.18)
Difficult to heat	0.543***	0.543***	0.541***	0.544***	0.541***	0.543***	0.543***	0.542***	-0.140
	(0.041)	(0.041)	(0.041)	(0.041)	(0.041)	(0.041)	(0.041)	(0.041)	(-8.03)
Having pests	1.252**	1.239**	1.250**	1.246**	1.250**	1.246**	1.246**	1.255**	0.051
	(0.116)	(0.114)	(0.114)	(0.116)	(0.114)	(0.114)	(0.115)	(0.116)	(2.38)
Expensive house	0.749**	0.777**	0.746**	0.774**	0.745**	0.771**	0.770**	0.752**	-0.059
	(0.089)	(0.092)	(0.089)	(0.091)	(0.089)	(0.092)	(0.091)	(0.089)	(-2.17)
Far from work[48]	0.813*	0.887	0.813*	0.876	0.812*	0.874	0.874	0.820	-0.030
	(0.098)	(0.108)	(0.098)	(0.107)	(0.098)	(0.106)	(0.106)	(0.100)	(-1.08)
Far from things	0.686**	0.715**	0.673**	0.710**	0.673**	0.700**	0.700**	0.680**	-0.079
	(0.105)	(0.107)	(0.102)	(0.107)	(0.103)	(0.105)	(0.105)	(0.103)	(-2.29)
Unsafe neighbourhood	0.644***	0.687**	0.638***	0.689**	0.639***	0.685**	0.684**	0.654***	-0.086
	(0.106)	(0.109)	(0.102)	(0.110)	(0.103)	(0.107)	(0.108)	(0.104)	(-2.34)

Noise and vibration	0.543***	0.566***	0.540***	0.567***	0.541***	0.565***	0.565***	0.550***	-0.130
	(0.042)	(0.045)	(0.042)	(0.045)	(0.042)	(0.045)	(0.045)	(0.042)	(-7.22)
Air pollution	0.971	1.022	0.978	1.023	0.976	1.029	1.028	0.990	0.005
	(0.149)	(0.159)	(0.150)	(0.161)	(0.150)	(0.162)	(0.162)	(0.154)	(0.14)
Tenure status (Owner)									
= Renter[53]	0.667***	0.641***	0.690***	0.659***	0.687***	0.674***	0.673***	0.698***	-0.096
	(0.036)	(0.036)	(0.037)	(0.036)	(0.038)	(0.037)	(0.037)	(0.038)	(-7.71)
= Family trust	1.259***	1.249***	1.240***	1.245***	1.242***	1.231***	1.232***	1.240***	0.050
	(0.074)	(0.073)	(0.074)	(0.073)	(0.074)	(0.073)	(0.073)	(0.074)	(3.71)
Free Time (Not enough free time)									
= The right amount of free time[55]	1.156***	1.118**	1.164***	1.116**	1.162***	1.121***	1.121***	1.148***	0.025
	(0.050)	(0.048)	(0.050)	(0.048)	(0.050)	(0.048)	(0.048)	(0.050)	(2.56)
= Too much free time	1.082	1.052	1.083	1.046	1.081	1.045	1.045	1.068	0.010
	(0.085)	(0.083)	(0.086)	(0.084)	(0.086)	(0.084)	(0.084)	(0.085)	(0.56)
Socialising (Every day)									
= At least once in the last four weeks	0.842*	0.834*	0.843*	0.842*	0.843*	0.842*	0.841*	0.853	-0.039
	(0.083)	(0.084)	(0.084)	(0.085)	(0.084)	(0.086)	(0.086)	(0.085)	(-1.69)
= Around once a fortnight	0.876*	0.877*	0.873*	0.878*	0.875*	0.875*	0.876*	0.875*	-0.030
	(0.066)	(0.066)	(0.067)	(0.066)	(0.067)	(0.066)	(0.066)	(0.067)	(-1.73)
= Around 1–6 times a week	0.902*	0.896*	0.901*	0.898*	0.903*	0.897*	0.898*	0.902*	-0.025
	(0.054)	(0.054)	(0.054)	(0.054)	(0.054)	(0.054)	(0.054)	(0.054)	(-1.79)
Local political involvement[60]	0.911**	0.920*	0.915*	0.926	0.915*	0.928	0.928	0.923*	-0.018
	(0.043)	(0.043)	(0.043)	(0.043)	(0.043)	(0.043)	(0.043)	(0.043)	(-1.65)
Recycling									
= None/A little	0.681***	0.685***	0.678***	0.682***	0.680***	0.681***	0.682***	0.677***	-0.088
	(0.046)	(0.049)	(0.047)	(0.048)	(0.047)	(0.048)	(0.048)	(0.047)	(-5.36)
= Some[60]	0.711***	0.706***	0.711***	0.704***	0.710***	0.704***	0.704***	0.708***	-0.080
	(0.054)	(0.054)	(0.054)	(0.054)	(0.054)	(0.054)	(0.054)	(0.054)	(-4.54)
= Most	0.841***	0.830***	0.836***	0.832***	0.837***	0.829***	0.829***	0.838***	-0.042
	(0.038)	(0.038)	(0.037)	(0.038)	(0.037)	(0.037)	(0.037)	(0.038)	(-4.04)
Water use (Never)									
= A little of the time[64]	1.097	1.106	1.101	1.112	1.101	1.116	1.116	1.107	0.024
	(0.086)	(0.089)	(0.086)	(0.089)	(0.086)	(0.089)	(0.089)	(0.086)	(1.33)
= Some of the time	1.194**	1.184**	1.190**	1.189**	1.192**	1.188**	1.189**	1.193**	0.040
	(0.095)	(0.093)	(0.094)	(0.093)	(0.094)	(0.093)	(0.093)	(0.095)	(2.21)
= Most of the time	1.200***	1.195***	1.199***	1.204***	1.204***	1.206***	1.207***	1.208***	0.043
	(0.074)	(0.076)	(0.074)	(0.076)	(0.074)	(0.076)	(0.076)	(0.076)	(2.92)
Energy use (Never)									
= A little of the time	0.842*	0.847	0.832*	0.848	0.832*	0.841*	0.840*	0.838*	-0.038
	(0.087)	(0.087)	(0.086)	(0.087)	(0.086)	(0.087)	(0.087)	(0.088)	(-1.60)
= Most of the time	0.876	0.881	0.871	0.875	0.869	0.872	0.871	0.868	-0.031
	(0.095)	(0.096)	(0.094)	(0.096)	(0.095)	(0.095)	(0.095)	(0.095)	(-1.22)
Council services (Very satisfied)									
= Very dissatisfied	0.756*	0.776*	0.760*	0.759*	0.759*	0.761*	0.761*	0.744*	-0.063
	(0.114)	(0.118)	(0.113)	(0.116)	(0.114)	(0.115)	(0.116)	(0.112)	(-1.80)
= Dissatisfied[70]	0.683***	0.695***	0.684***	0.685***	0.681***	0.685***	0.684***	0.676***	-0.087
	(0.062)	(0.064)	(0.062)	(0.063)	(0.062)	(0.063)	(0.063)	(0.062)	(-4.11)
= No feeling	0.531***	0.532***	0.531***	0.525***	0.529***	0.524***	0.523***	0.524***	-0.148

either way	(0.034)	(0.035)	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)	(-9.95)
= Satisfied	0.590***	0.593***	0.593***	0.587***	0.590***	0.589***	0.588***	0.586***	-0.122
	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)	(-9.35)
Facilities access (All of them)									
= Never been to	0.191*	0.202*	0.190*	0.198*	0.189*	0.197*	0.197*	0.190*	-0.372
	(0.172)	(0.191)	(0.175)	(0.186)	(0.172)	(0.187)	(0.186)	(0.173)	(-1.73)
= A few of them	0.986	0.969	0.986	0.976	0.988	0.976	0.977	0.989	-0.006
	(0.149)	(0.145)	(0.149)	(0.147)	(0.150)	(0.148)	(0.148)	(0.150)	(-0.16)
= None of them	0.797***	0.799***	0.795***	0.800***	0.795***	0.799***	0.799***	0.797***	-0.051
	(0.044)	(0.043)	(0.044)	(0.043)	(0.044)	(0.043)	(0.044)	(0.044)	(-4.10)
Green space access (Few of them)									
= Most of them[76]	0.951	0.955	0.943	0.946	0.944	0.940	0.941	0.936	-0.013
	(0.076)	(0.076)	(0.075)	(0.076)	(0.076)	(0.076)	(0.076)	(0.076)	(-0.68)
= All of them	1.060	1.059	1.054	1.049	1.054	1.044	1.044	1.044	0.011
	(0.095)	(0.093)	(0.094)	(0.093)	(0.094)	(0.092)	(0.092)	(0.094)	(0.54)
Green space state (Very satisfied)									
= Not been to	0.850	0.834	0.852	0.834	0.852	0.832	0.833	0.847	-0.042
	(0.193)	(0.191)	(0.191)	(0.193)	(0.191)	(0.190)	(0.190)	(0.194)	(-0.79)
= Very dissatisfied	1.123	1.147	1.152	1.159	1.164	1.190	1.196	1.158	0.034
	(0.357)	(0.382)	(0.362)	(0.387)	(0.367)	(0.391)	(0.394)	(0.368)	(0.44)
= Dissatisfied	1.007	0.994	1.010	0.993	1.011	0.996	0.997	1.002	-0.002
	(0.099)	(0.098)	(0.101)	(0.098)	(0.100)	(0.099)	(0.099)	(0.100)	(-0.07)
= No feeling either way	0.898	0.888	0.892	0.892	0.894	0.889	0.889	0.894	-0.026
	(0.081)	(0.079)	(0.079)	(0.080)	(0.079)	(0.079)	(0.079)	(0.080)	(-1.28)
Coastline access (Few of them)									
= Most of them[82]	0.890	0.901	0.892	0.895	0.891	0.896	0.896	0.887	-0.025
	(0.073)	(0.072)	(0.072)	(0.071)	(0.072)	(0.071)	(0.071)	(0.072)	(-1.39)
= All of them	0.985	1.002	0.979	0.997	0.980	0.993	0.994	0.978	-0.001
	(0.095)	(0.096)	(0.093)	(0.096)	(0.094)	(0.094)	(0.095)	(0.093)	(-0.03)
Coastline state (Very satisfied)									
= Not been to	1.052	1.003	1.033	1.014	1.034	1.003	1.003	1.039	0.003
	(0.280)	(0.268)	(0.275)	(0.270)	(0.275)	(0.266)	(0.267)	(0.276)	(0.05)
= Very dissatisfied	0.620**	0.615**	0.619**	0.622**	0.617**	0.621**	0.620**	0.628**	-0.109
	(0.119)	(0.120)	(0.117)	(0.122)	(0.117)	(0.120)	(0.121)	(0.119)	(-2.41)
= Dissatisfied	0.770***	0.777***	0.768***	0.774***	0.771***	0.774***	0.775***	0.768***	-0.059
	(0.067)	(0.069)	(0.067)	(0.069)	(0.067)	(0.069)	(0.069)	(0.067)	(-2.87)
= Satisfied[87]	0.701***	0.696***	0.701***	0.696***	0.703***	0.698***	0.698***	0.700***	-0.083
	(0.042)	(0.044)	(0.043)	(0.043)	(0.042)	(0.044)	(0.043)	(0.043)	(-5.78)
Constant	6.904***	5.794***	6.201***	7.344***	5.645***	6.613***	6.275***	6.904***	
	(3.490)	(2.841)	(3.110)	(3.670)	(2.869)	(3.304)	(3.178)	(3.490)	
Observations	21388	21388	21388	21388	21388	21388	21388	21388	
Population size	2770262	2770262	2770262	2770262	2770262	2770262	2770262	2770262	
†PPBR Wald test	<001***			<001***	0.2		0.8		
Pseudo R2	0.1454	0.1578	0.1466	0.1589	0.1469	0.1598	0.1598	0.1510	
AIC	3178969	3133067	3174488	3128709	3173621	3125489	3125375	3158438	

Exponentiated coefficients; Standard errors are reported in parentheses. Marginal effects are calculated based on Model (7). For marginal effects t-statistics are reported in the second row.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

† The significance level for PPBR and its squared (Squared PPBR) variables is reported jointly on the PPBR Wald test row.

The reported AICs are for the models that do not take into account jackknife replications.
Base group for each variable is written in the brackets. For example, for Age the more-than-75-years-old category is supposed as the base group.

The regression models presented in the following table correspond to those in Table 16:

Table 18. Relationships between the variables of interest using New Zealand's sample population.

	(1)	(2)		(3)		(4)	(5)
Dependent variable	PC	PPBR		CNOS		PPBR	PC
	Logit	OLS		MLogit		OLS	logit
			Need bdrs	No need	1 spare		
PPBR	3.272***		15.88***	-0.244	-0.783		
	(0.325)		(2.550)	(1.708)	(1.455)		
Squared PPBR	-0.549***		1.855	5.861***	4.100***		
	(0.135)		(1.142)	(1.104)	(0.996)		
ME	0.174***			0.109***			
Perceived Crowding		0.336***					
		(0.011)					
= One bedroom needed						-0.312***	-0.0892
						(0.016)	(0.128)
ME							-0.0193
= No bedrooms needed						-0.702***	-0.873***
						(0.015)	(0.126)
ME							-0.123
= One bedroom, spare						-1.053***	-2.306***
						(0.015)	(0.135)
ME							-0.209
Constant	-5.096***	0.910***	-23.14***	-6.388***	-2.54***	1.710***	-1.157***
	(0.194)	(0.003)	(1.913)	(0.631)	(0.487)	(0.014)	(0.120)
Observations	25466	25466		25466		25466	25466
N_pop	3344395	3344395		3344395		3344395	3344395
F	549.5	979.3		1129.1		4191.1	372.6
R2		0.065				0.588	0.073

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
All R-squared coefficients are reported based on OLS estimations. The ME rows include the marginal effects.

Appendix II. A comparison between categorical dependent variable models

A comparison between Table 17 and Table 19 reveals the reason for choosing the logit model with exponentiated coefficient in the results reported in 6.7. In the first column of table 19 (column 4a), the same model as model (4) of Table 17 is used, but the reported coefficients are not in exponentiated form. For example, gender log(0.827) is roughly equal to -0.19, which is the coefficient that is reported in model (4a) of Table 19. The coefficients of logit (model 1) are odds ratios, which state the chance of gaining a higher level of the dependent variable for one group versus another one. For example, an odds-ratio equal to -0.19 for gender represents a 19 per cent lower chance of being satisfied for males than for females. In the second model, if the exponentiated coefficients, also called log-odds, are greater than one, this means the effect of the explanatory variable is positive.

As shown in Table 19, the results derived from logit and probit are very similar. The differences in coefficients stem from the differences in the link functions, which usually leads to flatter tails in the logit estimations. While there is no specific rule for deciding between these two estimation methods, the GOF may still be the criterion. Since the logit estimation derives a lower AIC, the estimations are based on the logit function hereafter.

In models 3 and 4 the dependent variable is changed to an ordered categorical one with three values: satisfied, no feeling either way, and dissatisfied. The common method for regressing the ordered categorical variable on a number of explanatory variables seems to be the ordered logit, which is reported in model 3. However, this method assumes the same relationship between each pair of outcome groups; that is, the coefficients derived from a regression of the highest category (satisfied) versus the lowest category (dissatisfied) are the same as the coefficients derived from regressing the highest category (satisfied) versus the middle category (no feeling either way). This assumption, which is referred to as the parallel regression assumption, leads to equal coefficients between pairs of groups of the dependent variable and, consequently, only one model. For a comprehensive discussion on the use of the logit regression versus the ordered logit regression, see Section 5.3.

A likelihood ratio test is used to check on the validity of the parallel regression assumption. The null hypothesis assumes the equality of coefficients between the pairs of groups. The null hypothesis is rejected at 99.99 confidence level. Consequently, a method should be used that relaxes the parallel regression assumption. This is done in model 4 by serving a generalised ordered logit method. As illustrated, some of the coefficients change between the two categories, but the estimated coefficients for the variables of interest remain the same. This issue has been discussed more thoroughly in section 5.3.

Table 19. Results based on logit, probit, ordered logit and generalised ordered logit.

	(4a) Logit dummy	(4b) Probit dummy	(4c) Ologit categories	(4d) GOL Dissatisfied	(4d) GOL Satisfied
PPBR	-0.267	-0.190	-0.385*	-0.350*	-0.350*
	(0.23)	(0.14)	(0.20)	(0.20)	(0.20)
Squared PPBR	-0.009	0.010	0.074	0.050	0.050
	(0.11)	(0.07)	(0.09)	(0.09)	(0.09)
Perceived Crowding	-1.301***	-0.731***	-1.239***	-1.251***	-1.251***
	(0.09)	(0.05)	(0.08)	(0.06)	(0.06)
Gender	-0.190***	-0.113***	-0.149***	-0.049	-0.189***
	(0.04)	(0.02)	(0.04)	(0.06)	(0.04)
Partner	0.122**	0.068*	0.117**	0.120**	0.120**
	(0.06)	(0.03)	(0.05)	(0.05)	(0.05)
Age (>74)					
= 15–19	-0.040	-0.034	-0.125	-0.093	-0.093
	(0.18)	(0.10)	(0.16)	(0.14)	(0.14)
= 20–24	-0.300***	-0.190***	-0.343***	-0.311***	-0.311***
	(0.11)	(0.07)	(0.11)	(0.11)	(0.11)
= 25–34	-0.578***	-0.350***	-0.534***	-0.381***	-0.581***
	(0.10)	(0.06)	(0.09)	(0.11)	(0.09)

= 35–54	-0.598***	-0.367***	-0.549***	-0.382***	-0.592***
	(0.08)	(0.05)	(0.07)	(0.09)	(0.08)
= 55–59	-0.279***	-0.175***	-0.259***	-0.248***	-0.248***
	(0.09)	(0.05)	(0.09)	(0.09)	(0.09)
= 60–64	-0.180**	-0.110**	-0.200**	-0.196**	-0.196**
	(0.09)	(0.05)	(0.08)	(0.09)	(0.09)
= 65–69	-0.143*	-0.088*	-0.134*	-0.137*	-0.137*
	(0.08)	(0.05)	(0.08)	(0.08)	(0.08)
= 70–74	-0.144	-0.090	-0.105	0.215	-0.149
	(0.09)	(0.06)	(0.09)	(0.18)	(0.09)
Length of living in NZ (<4 years)					
= 4–10 years	-0.014	-0.006	-0.053	-0.050	-0.050
	(0.13)	(0.08)	(0.11)	(0.11)	(0.11)
= 10–25 years	-0.148	-0.090	-0.190	-0.186	-0.186
	(0.14)	(0.08)	(0.12)	(0.12)	(0.12)
>25 years	0.003	0.004	-0.007	0.009	0.009
	(0.11)	(0.06)	(0.09)	(0.10)	(0.10)
Ethnicity (European)					
= Maori	-0.104*	-0.060	0.007	0.256***	-0.093
	(0.06)	(0.04)	(0.06)	(0.10)	(0.08)
= Pacific	-0.292**	-0.157*	-0.118	0.164	-0.292*
	(0.14)	(0.08)	(0.12)	(0.15)	(0.15)
= Asian	-0.447***	-0.268***	-0.208**	0.249**	-0.442***
	(0.13)	(0.08)	(0.09)	(0.13)	(0.11)
= Other	-0.006	0.001	0.107	0.384***	0.000
	(0.10)	(0.06)	(0.09)	(0.15)	(0.11)
Education (No qualification)					
= Diploma	0.000	-0.001	-0.030	-0.024	-0.024
	(0.06)	(0.04)	(0.06)	(0.05)	(0.05)
= Degree	-0.088	-0.054	-0.156**	-0.283***	-0.106*
	(0.06)	(0.04)	(0.06)	(0.08)	(0.06)
Household income ($150,001 or more)					
= Zero	-0.084	-0.068	-0.128	-0.173	-0.173
	(0.30)	(0.18)	(0.32)	(0.32)	(0.32)
= $1–$5,000	-0.126	-0.077	-0.424	-0.390	-0.390
	(0.39)	(0.23)	(0.43)	(0.39)	(0.39)
= $5000–$50,001	-0.515***	-0.314***	-0.440***	-0.448***	-0.448***
	(0.08)	(0.05)	(0.07)	(0.08)	(0.08)
= $50,001–$70,000	-0.348***	-0.210***	-0.266***	-0.271***	-0.271***
	(0.09)	(0.05)	(0.08)	(0.08)	(0.08)
= $70,001–$150,000	-0.316***	-0.191***	-0.286***	-0.289***	-0.289***
	(0.07)	(0.04)	(0.06)	(0.07)	(0.07)
Health (Very dissatisfied)					
=Dissatisfied	0.170	0.096	0.175	0.150	0.150
	(0.16)	(0.09)	(0.14)	(0.12)	(0.12)
= No feeling either way	0.259*	0.144*	0.319***	0.297**	0.297**

	(0.14)	(0.08)	(0.12)	(0.12)	(0.12)
= Satisfied	0.400***	0.226***	0.447***	0.435***	0.435***
	(0.14)	(0.08)	(0.12)	(0.12)	(0.12)
= Very satisfied	0.619***	0.360***	0.614***	0.433***	0.654***
	(0.15)	(0.08)	(0.13)	(0.14)	(0.12)
Abilities (Very dissatisfied)					
= Dissatisfied	-0.415	-0.213	-0.399	-0.385	-0.385
	(0.45)	(0.26)	(0.39)	(0.38)	(0.38)
= No feeling either way	-0.690	-0.380	-0.680*	-0.670*	-0.670*
	(0.45)	(0.26)	(0.39)	(0.38)	(0.38)
= Satisfied	-0.421	-0.217	-0.299	-0.098	-0.390
	(0.45)	(0.26)	(0.39)	(0.38)	(0.38)
= Very satisfied	0.092	0.093	0.158	0.132	0.132
	(0.44)	(0.26)	(0.39)	(0.38)	(0.38)
Region (Auckland)					
= Wellington	0.006	0.003	0.019	0.016	0.016
	(0.06)	(0.04)	(0.05)	(0.06)	(0.06)
= Northland group	0.233***	0.148***	0.242***	0.245***	0.245***
	(0.07)	(0.04)	(0.06)	(0.06)	(0.06)
= Rest of North Island	0.279***	0.175***	0.215***	0.067	0.271***
	(0.07)	(0.04)	(0.06)	(0.08)	(0.06)
= Canterbury	-0.019	-0.007	-0.049	-0.038	-0.038
	(0.07)	(0.04)	(0.06)	(0.06)	(0.06)
= Rest of South Island	0.210***	0.130***	0.135**	-0.032	0.205***
	(0.07)	(0.04)	(0.06)	(0.09)	(0.07)
Bad street access	-0.196	-0.119	-0.139	-0.148	-0.148
	(0.17)	(0.10)	(0.14)	(0.14)	(0.14)
Poor condition	-0.984***	-0.545***	-1.394***	-1.522***	-0.909***
	(0.13)	(0.07)	(0.10)	(0.09)	(0.12)
Damp dwelling	-0.497***	-0.270***	-0.608***	-0.600***	-0.600***
	(0.10)	(0.05)	(0.07)	(0.07)	(0.07)
Difficult to heat	-0.609***	-0.356***	-0.610***	-0.610***	-0.610***
	(0.08)	(0.04)	(0.06)	(0.06)	(0.06)
Having pests	0.220**	0.138**	0.106	-0.033	0.224**
	(0.09)	(0.06)	(0.08)	(0.10)	(0.10)
Expensive house	-0.256**	-0.152**	-0.448***	-0.604***	-0.238**
	(0.12)	(0.07)	(0.09)	(0.11)	(0.11)
Far from work	-0.132	-0.081	-0.117	-0.125	-0.125
	(0.12)	(0.07)	(0.11)	(0.10)	(0.10)
Far from things	-0.343**	-0.187**	-0.554***	-0.761***	-0.321**
	(0.15)	(0.09)	(0.14)	(0.15)	(0.14)
Unsafe neighbourhood	-0.373**	-0.213**	-0.556***	-0.698***	-0.349**
	(0.16)	(0.09)	(0.13)	(0.13)	(0.16)

Noise and vibration	-0.567***	-0.328***	-0.510***	-0.506***	-0.506***
	(0.08)	(0.05)	(0.06)	(0.06)	(0.06)
Air pollution	0.023	0.017	0.029	0.030	0.030
	(0.16)	(0.09)	(0.13)	(0.11)	(0.11)
Tenure status (Owner)					
= Renter	-0.418***	-0.252***	-0.475***	-0.604***	-0.427***
	(0.05)	(0.03)	(0.05)	(0.07)	(0.06)
= Family trust	0.219***	0.137***	0.201***	0.201***	0.201***
	(0.06)	(0.04)	(0.05)	(0.05)	(0.05)
Free Time (Not enough free time)					
= The right amount of free time	0.109**	0.068***	0.117***	0.119***	0.119***
	(0.04)	(0.03)	(0.04)	(0.04)	(0.04)
= Too much free time	0.045	0.026	-0.009	-0.008	-0.008
	(0.08)	(0.05)	(0.07)	(0.08)	(0.08)
Socialising (Every day)					
= At least once in the last four weeks	-0.171*	-0.099	-0.247***	-0.421***	-0.172*
	(0.10)	(0.06)	(0.09)	(0.11)	(0.09)
= Around once a fortnight	-0.131*	-0.079*	-0.194***	-0.316***	-0.145**
	(0.08)	(0.05)	(0.07)	(0.09)	(0.07)
= Around 1–6 times a week	-0.108*	-0.065*	-0.108**	-0.111**	-0.111**
	(0.06)	(0.04)	(0.05)	(0.05)	(0.05)
Local political involvement[60]	-0.077	-0.048*	-0.049	-0.050	-0.050
	(0.05)	(0.03)	(0.04)	(0.04)	(0.04)
Recycling (All)					
= None/A little	-0.382***	-0.232***	-0.286***	-0.000	-0.397***
	(0.07)	(0.04)	(0.07)	(0.11)	(0.08)
= Some[60]	-0.350***	-0.214***	-0.254***	-0.053	-0.340***
	(0.08)	(0.05)	(0.07)	(0.11)	(0.07)
= Most	-0.184***	-0.112***	-0.159***	-0.044	-0.198***
	(0.05)	(0.03)	(0.04)	(0.07)	(0.05)
Water use (Never)					
= A little of the time[64]	0.106	0.063	0.068	0.069	0.069
	(0.08)	(0.05)	(0.07)	(0.07)	(0.07)
= Some of the time	0.173**	0.103**	0.157**	0.153**	0.153**
	(0.08)	(0.05)	(0.07)	(0.06)	(0.06)
= Most of the time	0.186***	0.111***	0.137**	0.138**	0.138**
	(0.06)	(0.04)	(0.06)	(0.06)	(0.06)
Energy use (Never)					
= A little of the time	-0.165	-0.096	-0.095	-0.085	-0.085
	(0.10)	(0.06)	(0.10)	(0.11)	(0.11)
= Most of the time	-0.133	-0.077	-0.058	-0.046	-0.046
	(0.11)	(0.07)	(0.11)	(0.11)	(0.11)
Council services (Very satisfied)					
= Very dissatisfied	-0.275*	-0.154*	-0.420***	-0.731***	-0.240*
	(0.15)	(0.09)	(0.15)	(0.18)	(0.14)
= Dissatisfied[70]	-0.378***	-0.223***	-0.411***	-0.389***	-0.389***
	(0.09)	(0.06)	(0.09)	(0.08)	(0.08)
= No feeling either way	-0.644***	-0.381***	-0.634***	-0.613***	-0.613***
	(0.06)	(0.04)	(0.06)	(0.08)	(0.08)

= Satisfied	-0.532***	-0.317***	-0.468***	-0.259***	-0.524***
	(0.06)	(0.03)	(0.05)	(0.08)	(0.06)
Facilities access (All of them)					
= Never been to	-1.618*	-0.869*	-0.519	2.254***	-1.601*
	(0.94)	(0.50)	(0.44)	(0.65)	(0.85)
= A few of them	-0.025	-0.024	0.028	0.025	0.025
	(0.15)	(0.09)	(0.13)	(0.13)	(0.13)
= None of them	-0.223***	-0.132***	-0.178***	-0.187***	-0.187***
	(0.05)	(0.03)	(0.04)	(0.04)	(0.04)
Green space access (Few of them)					
= Most of them[76]	-0.055	-0.034	0.042	0.254***	-0.054
	(0.08)	(0.05)	(0.07)	(0.08)	(0.07)
= All of them	0.048	0.030	0.095	0.071	0.071
	(0.09)	(0.05)	(0.08)	(0.07)	(0.07)
Green space state (Very satisfied)					
= Not been to	-0.182	-0.113	-0.355	-0.724**	-0.191
	(0.23)	(0.14)	(0.24)	(0.30)	(0.21)
= Very dissatisfied	0.148	0.110	0.223	0.214	0.214
	(0.33)	(0.19)	(0.34)	(0.29)	(0.29)
= Dissatisfied	-0.007	-0.012	-0.052	-0.044	-0.044
	(0.10)	(0.06)	(0.09)	(0.09)	(0.09)
= No feeling either way	-0.114	-0.071	-0.186***	-0.187***	-0.187***
	(0.09)	(0.05)	(0.07)	(0.07)	(0.07)
Coastline access (Few of them)					
= Most of them[82]	-0.111	-0.068	-0.099	-0.099	-0.099
	(0.08)	(0.05)	(0.06)	(0.07)	(0.07)
= All of them	-0.003	-0.006	-0.020	-0.018	-0.018
	(0.10)	(0.06)	(0.08)	(0.08)	(0.08)
Coastline state (Very satisfied)					
= Not been to	0.014	-0.005	0.020	0.051	0.051
	(0.27)	(0.16)	(0.25)	(0.25)	(0.25)
= Very dissatisfied	-0.474**	-0.282**	-0.629***	-0.612***	-0.612***
	(0.20)	(0.12)	(0.18)	(0.17)	(0.17)
= Dissatisfied	-0.256***	-0.155***	-0.259***	-0.265***	-0.265***
	(0.09)	(0.05)	(0.08)	(0.08)	(0.08)
= Satisfied[87]	-0.362***	-0.220***	-0.358***	-0.362***	-0.362***
	(0.06)	(0.04)	(0.05)	(0.06)	(0.06)
Constant	1.994***	1.185***		4.274***	1.859***
	(0.50)	(0.29)		(0.45)	(0.44)
cut1					
Constant			-4.698***		
			(0.43)		
cut2					
Constant			-1.729***		
			(0.43)		
Observations	21388	21388	21388	21388	
AIC	3128709	3131788			

Reported coefficients are exponentiated ones. Standard errors are reported in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Appendix III. Expanded table of results

Table 20 depicts the estimates results for the variables of interest, as illustrated in Table 15, and also for the control variables. Since CNOS accounts for a number of individuals' characteristics, such as the partnership status, including it in the model leads to a number of control variables having less statistical significance. Residential problems have an almost permanent negative correlation with residential satisfaction. Amongst housing problems, living in poor conditions is not one of the most common problems; however, compared to other housing problems, with a marginal effect equal to negative 22.6 per cent, it does seem to affect RS the most. The next challenging housing issue is dampness. A one-unit increase in the dampness of a dwelling is associated with a 11.4 per cent decrease in the likelihood of being satisfied. Noise and vibration in the environment is the most challenging neighbourhood problem.

In addition to the impact of the variables of interest, discussed in Section 6.7, the impacts of the controlling variables are of potential interest. The impact of control variables are very similar to the results of the previous studies.[58] In general, men are less likely to be satisfied with their residential environment. Marriage does not affect residential satisfaction significantly. Elderly people are more likely to be satisfied with their living environment. Longer length of staying in New Zealand has no significant impact on RS. Europeans are generally more likely to report higher residential satisfaction, while Asians are significantly less likely than Europeans to be satisfied with their living environment. While a few previous studies have expected different residential satisfaction levels for different education levels, our results do not suggest any significant relationship between education level and RS. The results indicate a positive correlation between satisfaction with health and residential satisfaction, which confirms the existence of over-optimism bias in respondents' evaluation of their satisfaction level. Compared to Auckland, the residents of areas farther from the central urban area are more likely to be residentially satisfied.

As explained earlier, respondents have also reported problems with their dwellings and with their neighbourhoods. Results indicate that poor access to streets does not affect RS significantly. However, living in a poorly conditioned dwelling lowers residential satisfaction significantly. People who live in damp houses are significantly less satisfied with their dwellings. Having an expensive house is associated with lower chance of satisfaction. All other problems, including having pests, being far from facilities, living in an unsafe neighbourhood, and having noise and vibration in the neighbourhood, are associated with a lower chance of residential satisfaction. Having air pollution in the neighbourhood does not affect RS significantly.

Compared to homeowners, renters are significantly less likely to be satisfied with their living environment. However, those people who live in family trusts are more likely to be residentially satisfied than

[58] For a review of previous studies, see section 2.3.

homeowners). Having enough free time, in conjunction with socialising more and being more involved in political affairs, measure an individual's attention to his or her neighbourhood issues. The results indicate a higher probability of residential satisfaction for those people who have enough free time (against those who claim they do not have enough time), and those who socialise more. However, local political involvement does not have any significant impact on residential satisfaction, which may be justified by the wide range of similar variables that I have controlled for, such as homeownership.

The other set of variables are measures of individuals caring about their environment and neighbourhood. People who care more about their water use and recycle more often are more likely to be residentially satisfied. Holding all other variables constant, being more satisfied with council services is associated with a higher level of residential satisfaction.

The control variables also include a number of accessibility variables. After controlling for the wide range of control variables, access to green spaces and the state of green spaces do not affect RS significantly. Having better accessibility to facilities, including shops, schools, post shops, libraries and medical services, is associated with a higher RS. Coastline accessibility does not affect RS significantly, but low-quality coastlines do lower residential satisfaction.

Table 20. Estimates results, including the impact of control variables.

	Model(1) PPBR	Model(2) PC	Model(3) CNOS	Model(4) PPBR&PC	Model(5) PPBR&CNOS	Model(6) PC&CNOS	Model(7) PPBR & PC & CNOS	Model(8) CC	Marginal Effect (dy/dx)
Residential Satisfaction (dummy)									
PPBR	0.762† (0.386)			0.802 (0.408)	1.591 (0.909)		1.361 (0.789)		-0.099
Squared PPBR	0.913† (0.222)			0.973 (0.242)	0.835 (0.217)		0.929 (0.250)		
Perceived Crowding		0.199*** (0.041)		0.212*** (0.044)		0.218*** (0.046)	0.217*** (0.045)		-0.325*** (-7.63)
Household Crowding (Two or more spare bedrooms)									
= One bedroom needed			0.277*** (0.093)		0.270*** (0.107)	0.344*** (0.116)	0.300*** (0.121)		-0.228** (-3.14)
= No bedrooms needed			0.674*** (0.095)		0.639** (0.131)	0.821 (0.115)	0.740 (0.152)		-0.0421 (-1.40)
= One bedroom, spare			0.676*** (0.064)		0.644*** (0.082)	0.736*** (0.069)	0.696*** (0.088)		-0.0653** (-3.25)
Crowding component								0.784*** (0.031)	-0.053***
Gender	0.957 (0.085)	0.971 (0.088)	0.949 (0.084)	0.967 (0.087)	0.951 (0.084)	0.959 (0.087)	0.959 (0.086)	0.951 (0.085)	-0.00905 (-0.47)
Partner	1.161 (0.130)	1.105 (0.113)	1.000 (0.101)	1.165 (0.133)	0.971 (0.122)	1.058 (0.109)	1.012 (0.130)	1.113 (0.114)	0.0121 (0.55)

Age (>74)									
= 15–19	1.208	1.026	1.088	1.152	1.040	1.085	1.028	1.266	0.0174
	(0.444)	(0.370)	(0.399)	(0.426)	(0.385)	(0.400)	(0.384)	(0.466)	(0.22)
= 20–24	0.747	0.655*	0.684	0.724	0.657	0.686	0.654	0.778	-0.0804
	(0.193)	(0.165)	(0.174)	(0.187)	(0.168)	(0.174)	(0.169)	(0.201)	(-1.49)
= 25–34	0.415***	0.402***	0.399***	0.426***	0.392***	0.415***	0.406***	0.436***	-0.188***
	(0.095)	(0.093)	(0.090)	(0.100)	(0.089)	(0.096)	(0.095)	(0.101)	(-3.78)
= 35–54	0.431***	0.428***	0.420***	0.447***	0.413***	0.441***	0.433***	0.451***	-0.175***
	(0.081)	(0.081)	(0.078)	(0.085)	(0.077)	(0.084)	(0.082)	(0.085)	(-4.28)
= 55–59	0.654**	0.699*	0.647**	0.695*	0.645**	0.687**	0.686**	0.661**	-0.0801*
	(0.125)	(0.132)	(0.124)	(0.131)	(0.123)	(0.130)	(0.129)	(0.125)	(-1.99)
= 60–64	0.833	0.849	0.819	0.840	0.817	0.830	0.830	0.818	-0.0397
	(0.190)	(0.198)	(0.185)	(0.194)	(0.185)	(0.191)	(0.191)	(0.186)	(-0.81)
= 65–69	0.824	0.838	0.799	0.826	0.806	0.806	0.811	0.802	-0.0460
	(0.156)	(0.156)	(0.147)	(0.157)	(0.151)	(0.148)	(0.152)	(0.150)	(-1.17)
= 70–74	0.746	0.744	0.737	0.738	0.738	0.732	0.732	0.731	-0.0666
	(0.165)	(0.167)	(0.162)	(0.165)	(0.162)	(0.163)	(0.163)	(0.161)	(-1.41)
Length of living in NZ (<4 years)									
= 4–10 years	0.835	0.854	0.900	0.851	0.901	0.908	0.917	0.866	-0.0205
	(0.184)	(0.192)	(0.201)	(0.190)	(0.202)	(0.205)	(0.207)	(0.192)	(-0.43)
= 10–25 years	0.811	0.821	0.886	0.813	0.887	0.871	0.883	0.847	-0.0296
	(0.179)	(0.180)	(0.200)	(0.179)	(0.207)	(0.197)	(0.207)	(0.186)	(-0.61)
>25 years	1.046	1.126	1.112	1.103	1.119	1.157	1.173	1.079	0.0310
	(0.187)	(0.207)	(0.202)	(0.203)	(0.206)	(0.216)	(0.223)	(0.196)	(0.78)
Ethnicity (European)									
= Maori	0.925	0.894	0.929	0.920	0.922	0.929	0.922	0.952	-0.0161
	(0.173)	(0.177)	(0.180)	(0.180)	(0.177)	(0.185)	(0.183)	(0.180)	(-0.37)
= Pacific	0.768	0.672**	0.780	0.726	0.777	0.748	0.738	0.811	-0.0607
	(0.154)	(0.132)	(0.164)	(0.144)	(0.163)	(0.154)	(0.152)	(0.163)	(-1.47)
= Asian	0.661**	0.623**	0.667**	0.642**	0.665**	0.652**	0.648**	0.671*	-0.0870*
	(0.134)	(0.126)	(0.135)	(0.131)	(0.136)	(0.133)	(0.133)	(0.136)	(-2.24)
= Other	1.075	0.998	1.124	1.029	1.120	1.068	1.065	1.113	0.0148
	(0.253)	(0.244)	(0.262)	(0.249)	(0.261)	(0.257)	(0.256)	(0.258)	(0.27)
Education (No qualification)									
= Diploma	1.118	1.148	1.116	1.132	1.119	1.132	1.135	1.097	0.0264
	(0.196)	(0.200)	(0.194)	(0.199)	(0.195)	(0.198)	(0.199)	(0.193)	(0.71)
= Degree	0.997	1.043	1.008	1.025	1.012	1.034	1.039	0.987	0.00705
	(0.173)	(0.179)	(0.171)	(0.178)	(0.172)	(0.178)	(0.179)	(0.171)	(0.19)
Household income ($150,001 or more)									
= Zero	0.914	0.989	0.945	0.945	0.994	0.999	1.038	0.873	-0.000252
	(0.460)	(0.528)	(0.490)	(0.487)	(0.500)	(0.521)	(0.525)	(0.454)	(-0.00)
=$1–$5,000	2.556	2.687	2.533	2.589	2.627	2.494	2.596	2.665	0.195
	(2.167)	(2.225)	(2.087)	(2.166)	(2.151)	(2.021)	(2.095)	(2.304)	(1.13)

=$5000–$50,001	0.627***	0.673***	0.638***	0.653***	0.652***	0.655***	0.669***	0.630***	-0.0903**
	(0.086)	(0.094)	(0.090)	(0.089)	(0.088)	(0.093)	(0.092)	(0.088)	(-2.96)
=$50,001– $70,000	0.780	0.832	0.793	0.820	0.803	0.826	0.836	0.792	-0.0409
	(0.130)	(0.143)	(0.131)	(0.141)	(0.135)	(0.142)	(0.145)	(0.133)	(-1.12)
=$70,001–$150,000	0.751**	0.770**	0.755**	0.762**	0.761**	0.761**	0.767**	0.755**	-0.0584*
	(0.091)	(0.094)	(0.093)	(0.093)	(0.093)	(0.095)	(0.095)	(0.091)	(-2.19)
Health (Very dissatisfied)									
= Dissatisfied	2.415**	2.469**	2.340**	2.449**	2.352**	2.348**	2.360**	2.410**	0.182*
	(0.906)	(0.923)	(0.846)	(0.922)	(0.852)	(0.853)	(0.856)	(0.901)	(2.34)
= No feeling either way	2.853***	2.964***	2.802***	2.919***	2.821***	2.825***	2.849***	2.858***	0.222***
	(0.872)	(0.896)	(0.828)	(0.891)	(0.835)	(0.835)	(0.842)	(0.871)	(3.50)
= Satisfied	3.414***	3.534***	3.326***	3.485***	3.351***	3.344***	3.369***	3.414***	0.258***
	(1.056)	(1.083)	(0.997)	(1.080)	(1.006)	(1.007)	(1.013)	(1.056)	(4.00)
= Very satisfied	4.401***	4.557***	4.321***	4.501***	4.366***	4.357***	4.398***	4.419***	0.314***
	(1.440)	(1.494)	(1.381)	(1.488)	(1.397)	(1.411)	(1.422)	(1.452)	(4.53)
Abilities (Very dissatisfied)									
= Dissatisfied	1.325	0.908	1.319	0.935	1.331	0.980	0.984	1.233	-0.00428
	(1.663)	(1.009)	(1.586)	(1.033)	(1.598)	(1.069)	(1.077)	(1.442)	(-0.02)
= No feeling either way	1.446	0.949	1.420	0.960	1.429	1.004	1.007	1.288	0.000792
	(1.775)	(1.024)	(1.667)	(1.029)	(1.675)	(1.063)	(1.070)	(1.467)	(0.00)
= Satisfied	1.768	1.153	1.754	1.157	1.763	1.215	1.222	1.568	0.0416
	(2.144)	(1.225)	(2.031)	(1.220)	(2.039)	(1.266)	(1.277)	(1.760)	(0.19)
= Very satisfied	2.716	1.814	2.754	1.800	2.773	1.924	1.943	2.415	0.140
	(3.306)	(1.928)	(3.199)	(1.899)	(3.217)	(2.004)	(2.030)	(2.714)	(0.63)
Bad street access	0.683	0.758	0.670	0.746	0.675	0.739	0.740	0.680	-0.0647
	(0.319)	(0.348)	(0.306)	(0.347)	(0.307)	(0.337)	(0.336)	(0.314)	(-0.67)
Poor condition	0.326***	0.345***	0.318***	0.348***	0.318***	0.344***	0.342***	0.327***	-0.228***
	(0.100)	(0.106)	(0.099)	(0.108)	(0.100)	(0.108)	(0.109)	(0.101)	(-3.41)
Damp dwelling	0.556***	0.599**	0.569***	0.596**	0.569***	0.598**	0.598**	0.570***	-0.110*
	(0.114)	(0.125)	(0.116)	(0.124)	(0.116)	(0.124)	(0.125)	(0.117)	(-2.48)
Difficult to heat	0.462***	0.461***	0.466***	0.464***	0.465***	0.468***	0.468***	0.469***	-0.162***
	(0.083)	(0.082)	(0.085)	(0.084)	(0.086)	(0.086)	(0.086)	(0.087)	(-4.18)
Having pests	1.361	1.323	1.356	1.332	1.351	1.336	1.333	1.363	0.0619
	(0.369)	(0.363)	(0.364)	(0.368)	(0.360)	(0.368)	(0.364)	(0.372)	(1.05)
Expensive house	0.564***	0.589**	0.560***	0.589**	0.559***	0.582**	0.581**	0.567***	-0.116*
	(0.122)	(0.125)	(0.121)	(0.125)	(0.121)	(0.124)	(0.125)	(0.122)	(-2.52)
Far from work	0.954	1.040	0.951	1.036	0.950	1.033	1.032	0.968	0.00688
	(0.216)	(0.243)	(0.218)	(0.243)	(0.219)	(0.244)	(0.245)	(0.224)	(0.14)
Far from things	0.620	0.631	0.613	0.621	0.612	0.615	0.614	0.606	-0.104
	(0.270)	(0.278)	(0.266)	(0.275)	(0.265)	(0.271)	(0.270)	(0.266)	(-1.11)
Unsafe neighbourhood	0.656	0.710	0.663	0.712	0.663	0.730	0.729	0.667	-0.0671
	(0.198)	(0.215)	(0.196)	(0.215)	(0.196)	(0.217)	(0.216)	(0.196)	(-1.06)
Noise and	0.563***	0.582***	0.557***	0.591***	0.556***	0.588***	0.586***	0.577***	-0.113**

vibration	(0.091)	(0.095)	(0.091)	(0.096)	(0.091)	(0.097)	(0.097)	(0.093)	(-3.25)
Air pollution	0.958	0.985	0.969	0.979	0.969	0.991	0.991	0.956	-0.00184
	(0.238)	(0.260)	(0.243)	(0.259)	(0.243)	(0.263)	(0.263)	(0.242)	(-0.03)
Tenure status (Owner)									
= Renter	0.598***	0.582***	0.626***	0.598***	0.628***	0.619***	0.622***	0.634***	-0.102***
	(0.072)	(0.071)	(0.076)	(0.073)	(0.077)	(0.075)	(0.076)	(0.078)	(-3.94)
= Family trust	1.277**	1.278**	1.254**	1.279**	1.252**	1.258**	1.255**	1.270**	0.0489*
	(0.132)	(0.132)	(0.131)	(0.133)	(0.132)	(0.132)	(0.133)	(0.134)	(2.18)
Free Time (Not enough free time)									
= The right amount of free time	1.280***	1.210**	1.301***	1.212**	1.303***	1.232**	1.235**	1.272**	0.0446*
	(0.120)	(0.114)	(0.121)	(0.114)	(0.121)	(0.115)	(0.115)	(0.119)	(2.24)
= Too much free time	1.328	1.255	1.325	1.260	1.323	1.258	1.259	1.325	0.0490
	(0.248)	(0.236)	(0.250)	(0.238)	(0.250)	(0.240)	(0.240)	(0.251)	(1.21)
Socialising (Every day)									
= At least once in the last four weeks	0.876	0.843	0.876	0.860	0.874	0.859	0.856	0.898	-0.0325
	(0.151)	(0.142)	(0.153)	(0.149)	(0.153)	(0.149)	(0.150)	(0.159)	(-0.88)
= Around once a fortnight	0.856	0.834	0.856	0.837	0.858	0.840	0.840	0.852	-0.0373
	(0.118)	(0.118)	(0.116)	(0.119)	(0.116)	(0.117)	(0.117)	(0.118)	(-1.25)
= Around 1–6 times a week	0.851	0.831*	0.854	0.834*	0.857	0.837*	0.838*	0.849*	-0.0379
	(0.083)	(0.082)	(0.083)	(0.082)	(0.083)	(0.082)	(0.082)	(0.083)	(-1.81)
Local political involvement	0.966	0.981	0.982	0.986	0.982	0.997	0.997	0.986	-0.000675
	(0.098)	(0.101)	(0.102)	(0.102)	(0.102)	(0.105)	(0.105)	(0.101)	(-0.03)
Recycling									
= None/A little	1.906	1.922	1.952	1.999	1.927	2.018	2.003	2.058	0.150
	(0.956)	(0.971)	(0.957)	(1.022)	(0.948)	(0.999)	(0.993)	(1.047)	(1.42)
= Some	0.838	0.825	0.824	0.825	0.825	0.809	0.810	0.837	-0.0453
	(0.293)	(0.289)	(0.295)	(0.292)	(0.296)	(0.289)	(0.291)	(0.305)	(-0.59)
= Most	0.599***	0.597***	0.606***	0.595***	0.603***	0.597***	0.597***	0.601***	-0.110**
	(0.103)	(0.100)	(0.103)	(0.100)	(0.103)	(0.100)	(0.100)	(0.102)	(-3.09)
Water use (Never)									
= A little of the time	1.068	1.071	1.103	1.068	1.105	1.089	1.094	1.085	0.0183
	(0.154)	(0.155)	(0.159)	(0.154)	(0.159)	(0.158)	(0.159)	(0.157)	(0.59)
= Some of the time	1.034	1.025	1.041	1.023	1.037	1.032	1.029	1.025	0.00665
	(0.134)	(0.133)	(0.134)	(0.135)	(0.133)	(0.135)	(0.134)	(0.136)	(0.24)
= Most of the time	0.947	0.954	0.955	0.951	0.957	0.958	0.961	0.947	-0.00906
	(0.100)	(0.099)	(0.098)	(0.102)	(0.099)	(0.100)	(0.100)	(0.101)	(-0.41)
Energy use (Never)									
= A little of the time	0.754	0.734	0.744	0.744	0.743	0.743	0.739	0.750	-0.0634
	(0.136)	(0.138)	(0.137)	(0.138)	(0.136)	(0.141)	(0.139)	(0.135)	(-1.58)
= Most of the time	0.844	0.860	0.831*	0.865	0.829*	0.856	0.852	0.847	-0.0332
	(0.088)	(0.091)	(0.088)	(0.091)	(0.088)	(0.092)	(0.091)	(0.089)	(-1.46)
Council services (Very satisfied)									
= Very dissatisfied	1.081	1.081	1.096	1.041	1.097	1.046	1.053	1.038	0.00959
	(0.387)	(0.386)	(0.383)	(0.374)	(0.386)	(0.367)	(0.370)	(0.368)	(0.13)
= Dissatisfied	0.619***	0.619***	0.616***	0.603***	0.614***	0.596***	0.598***	0.601***	-0.111**
	(0.107)	(0.107)	(0.108)	(0.103)	(0.107)	(0.103)	(0.103)	(0.105)	(-3.00)
= No feeling either way	0.448***	0.440***	0.437***	0.427***	0.434***	0.417***	0.417***	0.426***	-0.187***
	(0.072)	(0.072)	(0.071)	(0.069)	(0.070)	(0.068)	(0.067)	(0.069)	(-5.28)

= Satisfied	0.477***	0.477***	0.478***	0.468***	0.476***	0.467***	0.468***	0.468***	-0.163***
	(0.060)	(0.059)	(0.061)	(0.057)	(0.060)	(0.058)	(0.058)	(0.059)	(-6.06)
Facilities access (All of them)									
= Never been to	0.0236	0.0273	0.0247	0.0257	0.0245	0.0262	0.0263	0.0235	-0.778
	(0.093)	(0.106)	(0.096)	(0.100)	(0.095)	(0.100)	(0.101)	(0.092)	(-0.95)
= A few of them	0.384	0.369	0.374	0.361	0.370	0.354	0.353	0.369	-0.222
	(0.602)	(0.586)	(0.590)	(0.571)	(0.584)	(0.559)	(0.560)	(0.581)	(-0.66)
= None of them	0.840	0.778	0.842	0.802	0.846	0.802	0.801	0.854	-0.0471
	(0.422)	(0.387)	(0.418)	(0.402)	(0.421)	(0.398)	(0.398)	(0.429)	(-0.44)
Green space access (Few of them)									
= Most of them	0.419	0.460	0.426	0.487	0.429	0.502	0.498	0.457	-0.147
	(0.305)	(0.350)	(0.305)	(0.373)	(0.309)	(0.370)	(0.369)	(0.337)	(-0.94)
= All of them	0.959	0.951	0.977	0.958	0.979	0.979	0.982	0.975	-0.00445
	(0.281)	(0.271)	(0.278)	(0.278)	(0.279)	(0.277)	(0.277)	(0.288)	(-0.07)
Green space state (Very satisfied)									
= Not been to	0.365*	0.356*	0.355*	0.345*	0.357*	0.334*	0.337*	0.344*	-0.234
	(0.202)	(0.207)	(0.200)	(0.201)	(0.203)	(0.196)	(0.199)	(0.200)	(-1.88)
= Very dissatisfied	1.833	1.658	1.888	1.781	1.924	1.860	1.866	1.926	0.132
	(2.692)	(2.249)	(2.872)	(2.513)	(2.919)	(2.733)	(2.722)	(2.964)	(0.42)
= Dissatisfied	0.868	0.869	0.886	0.870	0.887	0.888	0.889	0.871	-0.0254
	(0.232)	(0.237)	(0.238)	(0.238)	(0.238)	(0.243)	(0.243)	(0.235)	(-0.43)
= No feeling either way	1.021	1.030	1.026	1.043	1.027	1.051	1.047	1.033	0.0106
	(0.213)	(0.215)	(0.214)	(0.217)	(0.213)	(0.219)	(0.217)	(0.214)	(0.24)
Coastline access (Few of them)									
= Most of them	1.678	1.818	1.654	1.822	1.646	1.749	1.755	1.775	0.119
	(0.842)	(1.000)	(0.833)	(0.996)	(0.830)	(0.942)	(0.948)	(0.905)	(1.04)
= All of them	1.067	1.064	1.057	1.083	1.058	1.066	1.064	1.095	0.0137
	(0.279)	(0.285)	(0.273)	(0.291)	(0.274)	(0.285)	(0.284)	(0.288)	(0.24)
Coastline state (Very satisfied)									
= Not been to	1.896	1.726	1.830	1.749	1.820	1.712	1.693	1.845	0.115
	(1.285)	(1.196)	(1.271)	(1.203)	(1.272)	(1.179)	(1.180)	(1.288)	(0.78)
= Very dissatisfied	1.078	1.081	1.064	1.087	1.052	1.079	1.067	1.072	0.0162
	(0.747)	(0.736)	(0.761)	(0.744)	(0.756)	(0.758)	(0.753)	(0.754)	(0.11)
= Dissatisfied	0.999	1.034	0.995	1.019	0.999	1.016	1.021	0.989	0.00339
	(0.192)	(0.203)	(0.190)	(0.199)	(0.191)	(0.197)	(0.197)	(0.189)	(0.08)
= Satisfied	0.835	0.826	0.841	0.818	0.843	0.826	0.828	0.818	-0.0409
	(0.136)	(0.133)	(0.138)	(0.132)	(0.139)	(0.134)	(0.135)	(0.133)	(-1.18)
Constant	0.855	0.951	0.808	1.172	0.649	1.086	0.918	0.644	
	(1.173)	(1.194)	(1.071)	(1.481)	(0.871)	(1.338)	(1.164)	(0.842)	
Observations	4706	4706	4706	4706	4706	4706	4706	4706	
Population size	925815	925815	925815	925815	925815	925815	925815	925815	
†PPBR Wald test	0.003	.	.	0.103	0.711	.	0.725	.	
Pseudo R2	0.1723	0.1860	0.1742	0.1872	0.1743	0.1888	0.1888	0.1778	
AIC	5184	5099	5175	5096	5178	5088	5092	5148	

Reported coefficients are exponentiated ones. Standard errors are reported in parentheses. Marginal effects are calculated based on the model (7). For marginal effects t-statistics are reported in the second row.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

† The significance level for PPBR and its squared (Squared PPBR) variables is reported jointly on the PPBR Wald test row. The reported AICs are for the models that do not take jackknife replications into account. The base group for each variable is written in parentheses. For example, for the Age variable, the 'more than 75 years old' category is supposed as the base group.

7 Residential Satisfaction, Household Crowding and Density: Evidence over Different Geographic Scales in Auckland

Abstract

This study analyses the role of households' absolute residential positions, particularly the level of household crowding and the level of area density, in determining the satisfaction of residents. In their residential environment evaluations, people may think of their neighbourhood boundaries at a very local level, such as an area just 'over the road', at the area known as their neighbourhood or suburb or an even larger area. This study introduces the best definition for neighbourhood boundaries in people's evaluations of their living environment. I then estimate the role of absolute crowding and density levels on residents' satisfaction with their living environment. The study highlights the importance of accounting for people's expected levels crowding and density in our estimation of the impact of housing features on residential satisfaction. The results indicate that if the absolute crowding level is higher than an individual's expected level of crowding, his or her residential satisfaction will be lower. A sizeable proportion of the impact of household crowding can be addressed by the subjective perception of their crowding level. Evidence indicates that people's evaluations of the density levels of their living environment mainly consider the level of density at the area that is conventionally known as their suburb. After taking people's location choice into account, a higher level of density, including both area density and residential density measures, lowers residential satisfaction significantly.

7.1 Introduction

In order to understand the effectiveness of urban development schemes, it is necessary to have a better understanding of the impact of the characteristics of the residential environment on residents' satisfaction. There is little evidence about the effectiveness of such schemes in delivering increased satisfaction. Residential satisfaction (RS) consists of a range of objective and subjective factors, so the study must include the subjective evaluations' complexities.

There has been a great deal of discussion between planners and economists about the impact and effectiveness of intensification. The two main goals are affordability and residential satisfaction (RS)[59]. It is generally believed that an increase in the supply of housing triggered by an intensification policy leads to an increase in affordability (though this remains to be verified). However, our

[59] This can be seen in Auckland Council's main vision, which is 'to create the world's most liveable city and deliver great value for money'.

knowledge of the impact of intensification on people's satisfaction with their residential environment remains limited.[60]

One of the complexities regards the definition of the living environment, especially the definition of neighbourhood boundaries. In a Chinese survey used by Kingdon and Knight (2007), over two-thirds of respondents (68 per cent) stated that they compare themselves with others in their 'own' village, which is a sign of the importance of geographic proximities. However, there is not a clear understanding of a person's definition of his or her 'own' neighbourhood. This is important when considering the effects of density since there may be a huge gap between the density levels of areas. As a result, while an area may be considered highly dense, an individual may not consider it dense because his or her definition of the neighbourhood boundaries differs from the conventional one. Therefore, the validity of the neighbourhood's boundaries at different geographic scales needs to be checked.

Another complexity of satisfaction analysis regards individuals' preferences. While a factor may be undesirable for one person, it may be desirable for the other. For example, while it seems that household crowding negatively affects the satisfaction of individuals with their living environment, some people may like to live in a crowded environment because of their cultural backgrounds. Stokols (1972) argues that:

> 'In situations where either normative or physical constraints inhibit overt behavioural adjustments of spatial variables, perceptual and cognitive modes of reducing the salience of restricted space will be more likely to occur. In such cases, the person may modify his standards of spatial adequacy, enhance the attractiveness of his task, or attempt to achieve a greater degree of coordination with others in the group, as a means of alleviating the sensation of crowding.'

The preferences issue will be addressed by adopting an appropriate estimative method, discussed in Section 7.5.

As concluded in the previous chapter, amongst three measures of crowding (perceived crowding (PC), Canadian National Occupation Standard (CNOS) and people per bedroom (PPBR)), the one comprising individuals' perception (PC) is the best predictor of their residential satisfaction. The objective measure (PPBR) has the lowest predictive power amongst the three measures.

[60] Planners have differing opinions about intensification. Some favour intensification by arguing that a higher intensification leads to an increase in walkability, which provides a better environment for businesses and is also associated with better health outcomes. On the other hand, some planners claim that a higher intensification is associated with a lower residential satisfaction as it leads to a decrease in the amenity value of a neighbourhood.

Nevertheless, PPBR remains a good measure for RS in the absence of the other measures of household crowding and is highly correlated with PC.

While a higher level of household crowding (PPBR) may be undesirable for one person, it may well be desirable to the other. The undesirability derives from the space shortage, which may lead to the lack of privacy and in some cases to health issues, such as the expansion of infectious diseases (Maani et al., 2006).[61] On the other hand, the desirability of crowding may stem from a number of individual characteristics. People's priorities are widely affected by their ethnicity. Some ethnic groups define crowding as a positive because they believe it creates a warm family environment.

Other than ethnicity, the impact of crowding on residential satisfaction is affected by individuals' age group, household size, gender, income group and neighbourhood location, all of which are known as 'own group' factors. These variables are also valid for considering density and other characteristics of the households' residential position. Therefore, as will be discussed in Section 7.2.1, this study accounts for the endogeneity of its variables of interest, namely crowding and density.

In studying the impact of crowding, it is important to consider the role of area density. The opportunities to access green spaces are not equal amongst different neighbourhoods. People located in business centres often benefit less from public green spaces. The results of hedonic analyses suggest that the impact of green spaces on prices depends on neighbourhoods' density levels and socio-demographics, such as income and the rate of crime (Troy & Grove, 2008). Barbosa et al. (2007) reported on the distance of different social groups from public green spaces. They concluded that the elderly and most deprived social groups enjoy the highest access to public green spaces. Even though people may not live in a crowded dwelling, they may think of a greener neighbourhood in their evaluations of the condition of their living environment.

The use of these two specific measures of density (area density and residential density) is controversial (for a review of residential density and area density concepts see Section 2.3.1 and for further details see Boyko and Cooper (2011)). While amongst urban planning studies there is an emphasis on the residential density measure, this needs to be confirmed by taking people's perceptions into account. The present study examines which of these measures best determines residential satisfaction.

The remainder of this chapter is organised as follows. Section 7.2 reviews the literature, including an introduction to residential satisfaction, the definitions and differences of crowding measures, the neighbourhood effects and the definitions of neighbourhood boundaries, and an explanation of the average approach taken in this study. Section 7.3 describes the datasets and the following section presents descriptive statistics. As presented in this section, the area density measures are calculated based on defining neighbourhood boundaries at different geographic scales. The methodology is

[61] The causal relationship between crowding and the expansion of diseases needs further investigation.

presented in Section 7.6. The last section includes the estimated results and concluding remarks. This chapter finishes with some suggestions for future research.

7.2 Literature review

As discussed in Section 2.2, RS is considered as a binary variable. Within this dissertation, Section 2.4 has already discussed issues of measurement error, Section 2.3.1 has covered the differences between crowding and density measures, and Section 2.4.4 has examined the complexity of a causal inference in our context. In the current section, I discuss the importance of a clear definition of neighbourhood boundaries in identifying the effects of neighbourhoods on the residents. I then present a review of the studies on the impact of individual and household socio-demographic characteristics on residential satisfaction.

Cheshire and Sheppard (1998) estimated the demand characteristics for individual housing by using a Box-Cox form hedonic price function. They compared the estimated demand elasticities for different features of housing between two medium-sized British cities. Their results indicate that an increase in income leads to a lower price elasticity of demand for all housing attributes, except for internal space and land area. They observed that the demand for land is much more income-elastic than any other attribute of housing, especially for higher income groups. In a hedonic analysis, Anderson and West (2006) studied the impact of proximity to open space on sale prices. Their results indicate that proximity to open space is valued more highly in denser areas located nearby the CBD.

Song and Knaap (2003) attempted to characterise various neighbourhoods in Washington County (USA) by developing a number of measures of urban form. Their results indicate a lower willingness to pay for houses in denser areas, and the areas with more commercial, multifamily and public uses.

As noted in the previous chapter, Bonnes et al. (1991) studied the relationship between crowding and residential satisfaction by using factor analysis and regression analysis. Their sample population was 461 residents of Rome. However, their dataset did not provide them with enough control variables. Their results indicate a significant spatial effect in studying the impact of crowding on residential satisfaction. They highlighted the need for a further study on the differences between crowding perceptions and the objective measures of density and crowding. They also highlighted the need for a careful definition of neighbourhood boundaries. Despite the limitations of their study, they concluded that residential satisfaction is affected to a remarkable degree by the perception of crowding. The second-most important factor is perceived density of the spatial features of the environment, such as building density and building height. Bonnes et al.'s results suggest that the socioeconomic aspects, measured by age, estimated income, profession and educational background, and the temporal experienced residency and the spatial location affect the satisfaction of residents with their living environment significantly.

Williams, Jenks and Burton (2000) study the relationship between intensification and the overall satisfaction with the area of residence. The result of their statistical analysis suggests that the residents of denser areas are not dissatisfied with higher levels of intensification. Also, people generally have high satisfaction with dense areas. The study notes the importance of accounting for the role of location choice and the features of suburbs in our understanding of the impact of intensification on satisfaction of residence with their area.

7.2.1 Location choice

The purpose of this section is to address the location choice issue. Morrison (2015b) referred to the neglected role of geography in the growth of poverty. He discussed that one reason for not accounting for residential sorting is the lack of evidence of the impact that neighbourhoods have on residents. He highlighted the importance of location choice by shedding light on the contradiction between the results of the studies that account for residential sorting and those that do not. For example, in a longitudinal study, Bolster et al. (2007) found no negative relationship between neighbourhood disadvantage and subsequent income growth. On the other hand, as noted by Cheshire (2007), a number of studies have concluded that social homogeneity, derived from residential sorting, facilitates job-matching and communication.

Morrison (2015b) introduced the disappointing experience with mixed neighbourhoods as another reason for the ignorance of the implications of residential sorting. He argued that because of the freedom of choice in democratic societies, people, especially New Zealanders, prefer not to live next to areas with a different affluence level (Morrison, 2015a). I distinguish between people who have randomly located in their current living area and those who have aimed to locate where they are living.

Morrison (2014a), in his review of Marans and Stimson's (2011) book on 'Investigating Quality of Urban Life', referred to the authors' result on the independence of life satisfaction and the region of residence. In this relationship, Morrison noted that people self-select their region of residence such that their lifestyle and their residence be brought into harmony. He then highlighted the need to account for the impact that personal attributes have on an individual's choice of location in econometric modellings.

It is very likely that the segregation/desegregation in neighbourhoods is affected by ethnicities and occupations. Ethnicity distribution is important because of the networking within ethnicities; that is, individuals with the same ethnicity are likely to settle into a specific neighbourhood in order to form social and financial activities. Also, people working in a similar field are likely to locate in specific areas. For instance, farm workers may settle into the fringe of the cities. In order to account for neighbourhood effects, the distribution of ethnicities and occupations needs to be addressed.

Therefore, for each mesh-block (MB), a ratio of different ethnicities is defined as the ratio of population of each ethnicity to the MB's total population (the MB is the smallest geographic segment, which will be introduced more thoroughly in Section 7.4.2); this is known as the MB's ethnicity ratios. The same approach is taken for different occupations to calculate the MB's occupation ratios. In addition to these ratios, which are about the ethnic and occupational compositions of neighbourhoods, the own occupation proportion and the own ethnicity proportion are included in the estimations. These proportions are the person's own occupation's and own ethnicity's ratios. Therefore, the main difference between the proportions and the ratios is that the ratios are about the compositions of neighbourhoods – that is, the neighbourhoods' characteristics –, while the proportions are about the individual's ethnicity as a ratio of the total population; that is, the individuals' characteristics.

A reasonable location decision rule is that individual 1 chooses neighbourhood 'a' if the utility of locating in neighbourhood 'a' is higher than the utility of locating in any other neighbourhood 'b' (Glaeser & Gottlieb, 2009). This utility function represents a relationship between error terms (ε_{1n},ε_{2n}), which is observable by individual 1 and 2 but not by econometricians, and individuals' characteristics (x_{1n},x_{2n}). This leads to the inconsistent estimation of parameters. A solution to this problem is the use of instrumental variables. However, as Van Ham and Feijten (2008) mentioned, the complex relationship between the behaviour of the housing market and neighbourhood characteristics makes it difficult to identify a valid instrument. A conventional solution is to ignore the locational adjustment when the speed of that process is slower than the speed of social interactions formation. In the present study, I used a proxy for the locational decision. In a study on the factors that affect Aucklanders' location choice, Mare and Coleman (2011) emphasised the effect of 'own group' attraction, such as ethnicity, income and the country of birth. Consequently, by including these variables as control variables, the group membership is controlled for.

As shown in the descriptive statistics section (Appendix i), 64 per cent of our sample are of European ethnicity, 8 per cent are Maori, 8 per cent Pacific Island, 15 per cent Asian and 5 per cent of other ethnicities. In practice, ethnicities are not distributed evenly across neighbourhoods. In order to know more about the population patterns, I am interested in the percentage of neighbourhoods that experience the concentration of specific ethnic groups at different levels over the mean of the ethnic groups' mean in the sample population. Therefore, I defined four categories for the level of excess concentration over the mean of each category; this is called 'excess concentration relative to mean'.

As illustrated in Table 21, for example, 21 per cent of neighbourhoods have at least 25 per cent extra concentration of Europeans compared to the mean of the European category in the sample; that is, 21 per cent of neighbourhoods have at least 89 per cent European residents. In total, 45 per cent of neighbourhoods have an excess concentration of a specific ethnicity of at least 25 per cent greater than their means. This is evidence of the presence of ethnic segregation in the sample.

Similar to ethnicity, the other variables included in Table 21 also show patterns of segregation amongst the neighbourhoods. These variables are very individual-specific; that is, they have a lot to do with individuals' backgrounds. This reflects the individual's character and his or her preferences. The segregation amongst neighbourhoods seems very apparent. For example, for the variable measuring the number of generations in New Zealand, a concentration of at least 25 per cent for people with different numbers of year staying in New Zealand happens in a total of 19 per cent of neighbourhoods. Therefore, as will be explained in 7.6, these variables will be used for capturing individuals' location choice.

Table 21. The distribution of people with different characteristics amongst neighbourhoods.

			Excess concentration relative to mean			
Variables	Categories	Mean	10%	15%	20%	25%
Ethnicity	European	0.64	0.4	0.33	0.26	0.21
	Maori	0.08	0.13	0.11	0.07	0.05
	Pacific	0.08	0.16	0.14	0.12	0.1
	Asian	0.15	0.26	0.19	0.13	0.09
	Other	0.05	0.07	0.04	0.02	0.01
	Total	1	1.03	0.81	0.59	0.45
The number of people who raised the individual who were born in New Zealand	None	0.44	0.34	0.23	0.17	0.1
	One	0.14	0.12	0.05	0	0
	Two	0.37	0.25	0.14	0.09	0.06
	More than two	0.04	0.07	0.03	0.02	0
	Total	1	0.81	0.46	0.29	0.17
The number of generations in New Zealand	First generation	0.37	0.3	0.17	0.15	0.08
	Second generation	0.11	0.14	0.08	0.05	0.03
	Third generation or more	0.52	0.3	0.2	0.12	0.07
	Total	1	0.77	0.48	0.33	0.19
Identity Expression	Very easy	0.40	0.27	0.15	0.11	0.06
	Easy	0.44	0.25	0.18	0.1	0.07
	Sometimes easy, sometimes difficult	0.12	0.13	0.04	0.03	0.02
	Difficult	0.02	0.02	0.01	0	0
	Very difficult	0.01	0.01	0	0	0
	Total	1	0.71	0.4	0.25	0.16
Length of stay in New Zealand	< 4 years	0.04	0.05	0	0	0
	= 4–10 years	0.11	0.12	0.06	0.05	0.03
	= 10–25 years	0.12	0.16	0.11	0.07	0.05
	>25 years	0.73	0.32	0.21	0.14	0.06
	Total	1	0.65	0.38	0.26	0.14

* The non-stated categories of the variables are not included in the table.

Summary

As previous studies have noted, people choose the location of their residence in a way that will maximise their utility. For example, a person may be happy to live in a crowded house because his or her priority is to live nearby his or her work place and may not express dissatisfaction with his or her household crowding. This leads to an under-estimation of the negative impact of crowding on residential satisfaction. Therefore, it is important to control for the level of crowding that an individual is likely to have. I then distinguish between the effect of the likely level of crowding on RS and the level of crowding imposed by the limitations of the available choices in the market. I capture the pattern of segregation amongst suburbs by using a number of factors related to residents' backgrounds. These background variables are used to estimate the levels of crowding and density that an individual may be indifferent to.

7.3 Hypothesis development

I hypothesise that a higher household crowding level lowers residential satisfaction. However, people may endogenise their location choice, and hence their crowding level, in which case their chosen household crowding level may not affect RS. However, if people fail to maximise their residential satisfaction subject to the exogenous constraints they face, the impact of household crowding will not vanish after taking the factors of location choice into account. For example, a family of eight, with parents who believe that each child should have his or her own room, may not be able to fit into a standard apartment size near a business centre.

If an individual's crowding level can be explained by his or her previous life circumstances, he or she may consider it as a normal or even desirable level of crowding, even though that level of crowding would be considered by the public as a negative factor. However, if the person is not able to settle into a place with a desired level of crowding, defined based on the norms stemming from her backgrounds, he or she may feel unhappy with his or her crowding level. Therefore, I hypothesise that:

H1: A higher level of expected crowding has a non-negative effect on residential satisfaction.

H2: A higher level of unexpected crowding lowers residential satisfaction.

There are a wide range of definitions for density. Empirically, I have chosen the one that predicts residential satisfaction the best. A higher level of density may reduce residential satisfaction because of greater pressures on available local amenities.

On the other hand, a higher density level may be associated with agglomeration advantages for residents and may therefore increase residential satisfaction. Again, because of locational choices, the density level that will be undesirable for an individual is different from that person's expected living area's density level. Therefore, I hypothesise that:

H3: A density level different from an individual's expected density level lowers residential satisfaction.

7.4 Data and sample

As explained in Section 3.1, two main datasets are used for the purpose of this study; namely, three cross-sections of NZGSS and Census 2006 dataset. This section explains the merging of these datasets and also the use of resampling weights.

In order to derive the variables of interest, the NZGSS (Statistics New Zealand, 2009, 2011) and the New Zealand Census of Population and Dwellings (Statistics New Zealand, 2006) are matched based on mesh-block IDs, for which there is data in both datasets. The individuals' variables, such as demographics, derived from NZGSS can then be used as control variables, and the area-level variables, such as the density of mesh-blocks, can be derived from the Census.[62]

NZGSS provides us with a total of 5816 observations located in the Auckland region. In order to have comparable results between the following chapters, the number of observations is henceforth fixed at 4607 observations. In addition to the reasons for the dropping of observations explained in the last chapter (see section 6.5), mismatches between the datasets necessitates that extra observations will be dropped. As will be explained in the following chapters, one of the reasons for the loss of observations is to avoid including areas with fewer than 12 residents, which drops 3.5 per cent of the observations. Outliers of the constructed measures, 2 per cent of the sample size, are also dropped while 6 per cent of observations are excluded because of the lack of information about their neighbouring areas in our geo-spatial analysis, which will be explained in Appendix i.1 in Chapter 6. The rest of the observations are dropped in order to avoid cases of perfect multi-collinearity in the data.

One potential concern regards the use of different cross-sections of data, as there might be some changes not captured in between. However, since I am not studying a longitudinal change, this is not a major source of concern in this study. Also as Meen, Nygaard, and Meen (2013) argued, neighbourhoods' social structures are persistent over long periods such that their relative spatial patterns may not change for decades.

In the neighbourhood-level studies, availability of a dataset that provides data at small geographic scale plays a prominent role. To the best of my knowledge, no previous study has attempted to use the NZGSS (Statistics New Zealand, 2009, 2011, 2013) matched with New Zealand Census of Population and Dwellings (Statistics New Zealand, 2006) to have data on both individuals and

[62] In order to add a density measure based on the developable land area, a dataset provided by Auckland Council Research and Evaluation Unit (RIMU) was matched with the outcome of the merging of Census with NZGSS.

neighbourhoods at a very local level. This point has been noted by Morrison (2015b). However, Clark and Kim (2012) used three cross-sections of Census at the neighbourhood level to study the effect of diversity in neighbourhoods on volunteering, although they did not integrate the Census data with the GSS unit record data. Another example is Sengupta et al. (2013), which links New Zealand Census areal data to the Quality of Life Survey.

7.4.1 Resampling by using replicate weights

Different methods of resampling may serve to increase the precision of estimates using survey designs by deriving more robust standard errors, proportions, odds ratios and regression coefficients. To do this, a random set of observations is left out at each time of estimation. Replication of this leads to the estimation of the bias of a statistic. This method is called jackknifing. NZGSS provides us with replicate weights produced by the delete-a-group jackknife method (Kott, 2001). In the dataset, 100 groups are derived by using primary sampling units (PSUs) that are randomly sorted into each stratum. This strategy results in 100 replicate samples, in each of which one of the groups is omitted and weights are adjusted accordingly. Using these weights leads to estimates that tend asymptotically to true values.

7.4.2 The geographic scales

As discussed in Section 2.4, in order to understand the effect of neighbourhoods I should first have a clear definition of neighbourhood, which has been defined differently amongst studies. To have a reliable definition of the neighbourhood, I would ideally rely on the resident's perception of their neighbourhood boundaries that can be derived from face-to-face interviews (Witten et al., 2003). People's understanding of their neighbourhoods' boundaries depends on their socioeconomics, such as income and education levels and the geographic distribution of their relationships. In this study, I have controlled for a wide range of socioeconomic factors to account for individuals' characteristics in their definitions for their neighbourhoods' boundaries. A typical definition of neighbourhood consists of a certain number of neighbours. This definition is very similar to the definition of a mesh-block as an area consisting of dwellings that are located close to each other. This is defined by SNZ as a very small neighbourhood, which consists of a few houses located next to each other, with an average population of 110 people in the Auckland region. As I assume that the reference group consists of people who live shoulder-to-shoulder with an individual, the MB design suits our purpose well.

In the context of New Zealand, the larger neighbourhoods are defined as area units. Most people consider themselves as a member of their suburbs, to which area units are an approximation. Hence, to take perceptions into account, area units may also be considered as a good definition of neighbourhoods. Neither mesh-blocks nor area units are based on administrative boundaries. To take a step forward into the perceptional definition of neighbourhoods, I consider the boundaries of territorial authorities (TLAs), which emphasise the importance of an administravity based community of interest. These three different scales of 'neighbourhood' each take into account proximities (albeit to differing degrees) and allow for common bonds, such as feeling of attachment, or entities, such as a church group that may have differing reaches in terms of scale.

Summary

This chapter has three major research questions about the impact that expected crowding, unexpected crowding and unexpected density levels have on residential satisfaction. Due to a number of reasons, the number of observations for the rest of the thesis is decreased to 4607. Using resampling weights, the precision of the estimates are increased. For a good understanding of the effects of neighbourhoods on residents' satisfaction, it is important to have a good definition of 'neighbourhood'. To find the best scale at which the consequences of density impacts individuals the most, this study relies on residents' perception of their neighbourhood boundaries.

7.5 Descriptive statistics

Chapter 6's analysis was done by using Auckland's sample population. The variables of interest across this chapter and the following chapters are very similar. Therefore, to remain consistent and to derive comparable results across this chapter and the following chapters, the sample population needs to be identical. However, due to spatial analysis in the following chapters, some observations will be omitted from the sample population. Therefore, the sample population has been limited to the smallest sample population amongst the next three chapters. A description of the variables in common between this chapter and Chapters 8 and 9 is presented in Appendix i. A glossary of all the constructed measures in this thesis is presented in the Glossary chapter.

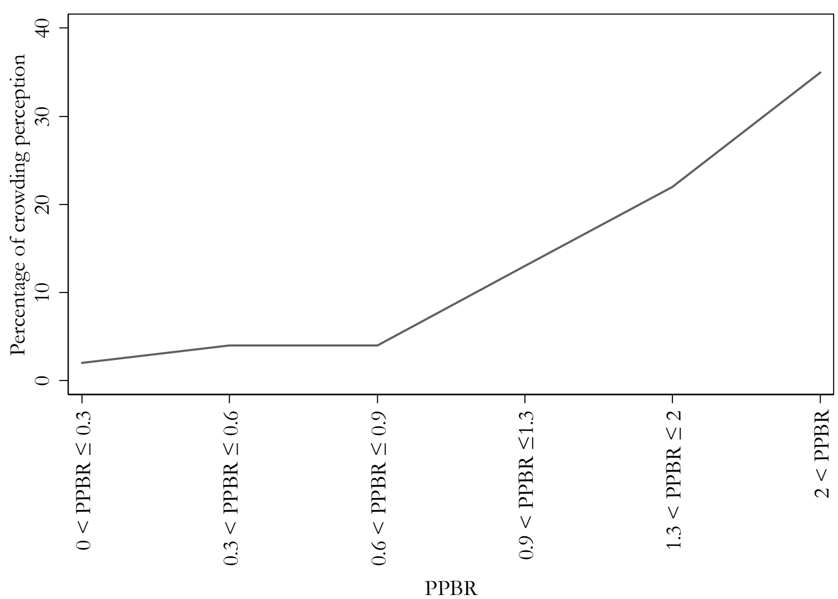
As shown in equation (21), the objective crowding measure, people per bedroom (PPBR), is measured as a ratio of household size to the number of bedrooms. An updated figure of the relationship between PPBR and household size, based on the sample population used in the current chapter, is depicted in Table 22. Also, as depicted in Figure 17, the percentage of people who perceive a small-house problem increases as the objective crowding measure (PPBR) increases.

Table 22. Household size versus PPBR.

Household size	Household crowding (PPBR)					
	0<PPBR≤ 0.3	0.3<PPBR≤ 0.6	0.6<PPBR≤ 0.9	0.9< PPBR ≤ 1.3	1.3< PPBR ≤ 2	2< PPBR
One person	12	73	0	15	0	0
Two people	0	23	49	24	0	3
Three people	0	0	29	55	15	0
Four people	0	1	11	34	45	9
Five people	0	0	6	50	44	0
Six people	0	0	1	23	34	42
Seven people	0	0	0	9	91	0
Eight people	0	0	0	0	69	31
Number of observations	4607					

Note: numbers are in percentages.

Figure 17. PPBR versus PC (small-house problem).



Note: The vertical axis is the percentage of sample population who have perceived their house crowded (for data see Table 27 in Appendix I).

7.6 Research design

In this chapter, the impact of household crowding and density on residential satisfaction is estimated. In line with the last chapter's research design, this chapter estimates the following equation:[63]

$$y_i = \alpha_0 + \alpha_1 . x_i \;\; + \alpha_2 . x_i^2 \;\; + \alpha_3 . (\text{Control variables})_i + \varepsilon_i \qquad (22)$$

where the dependent variable (y_i) is individual i's residential satisfaction. x_i is a vector of our variables of interest, namely household crowding and density, which are both measures of households' absolute position; x_i^2 is the squared term of the absolute measures; α_1 and α_2 are the parameters of interest; and ε_i is the residuals. Control variables are explained in the following subsection (7.6.1).

The dependent variable (y_i) follows a Bernoulli distribution; that is, the dependent variable is a dummy. Thus, the estimated values derived from the right-hand side of the equation should be modified so that the outcome is between 0 and 1. To do this, as discussed in Section 6.6, a categorical dependent variable transform function will be served. Two common functions are logit and probit. In this study, a logit transform function is used.

This study takes a parametric approach. Manski (1987) noted the difficulties of interpreting the full nonparametric discrete choice models. The semiparametric approach is a bridge between the more robust nonparametric estimators and the fragile but more informative parametric estimators. However, as the nonparametric approach does not provide us with any equally efficient counterpart to White's heteroskedasticity-consistent estimator (White, 1980), the treatment of heteroskedasticity in discrete choice models needs to be addressed parametrically (Greene & Hensher, 2010).

It is necessary to deal with the endogeneity issue in the estimation of equation (23). It may be the case that a higher level of crowding or density will be desirable for some people based on their backgrounds. Therefore, this chapter employs a two-stage least squares (2SLS) method to account for the endogeneity (the existence of endogeneity must first be checked by using statistical tests, such as Durbin and Wu–Hausman tests). To do this, in order to maximise the chance of the exogeneity of instruments, I used the observation of the first cross-section of data as a control group. Hence, the endogenous variables for individual i at year 2008 are regressed on the list of instruments, z_{i2008}. The first stage of 2SLS is as follows:

$$x_{it} = \theta_0 + \theta_1 . z_{it} + u_{it}, t = 2008 \qquad (23)$$

The list of instruments includes the variables that represent the own-group attraction characteristics, such as the number of generations in New Zealand. Then, based on the estimated parameter, $\widehat{\theta_1}$, the

[63] The functional form has been chosen by a comparison between using the model with and without the squared term.

probability of living in a crowded household ($\widehat{\text{PPBR}}$) and the probability of living in a dense neighbourhood ($\widehat{\text{DENSA}}$) for the observations of 2010 and 2012 will be estimated as follows:

$$\hat{x}_{it} = \widehat{\theta_0} + \widehat{\theta_1}.z_{it}, t = 2010, 2012 \tag{24}$$

The dependent variable would then be regressed on the predicted values ($\hat{x}_{it}$). Hence, the variables of interest in equation (23) will be substituted with their predicted values:

$$y_i = \beta_0 + \beta_1.\hat{x}_i + \beta_2.\hat{x}_i^2 + \beta_3.(\text{Control variables})_i + \psi_i, \tag{25}$$

$$\begin{bmatrix} \psi_i \\ \left(\sum_{t=2008}^{2012} u_t\right)_i \end{bmatrix} \sim \text{NID}\left(0, \begin{bmatrix} \sigma_\psi^2 & 0 \\ 0 & \sigma_u^2 \end{bmatrix}\right)$$

where β_1 and β_2 are the parameters of interest and ψ_i is the disturbance term, which follows a normal distribution with a mean equal to zero and a variance equal to σ^2, i.e. $\psi_i \sim N(0, \sigma_\psi^2)$. The instruments ($z_{it}$), included in equation (25), need to be exogeneous; that is, not correlated with ψ_i[64]. The basis for the choice of instruments is presented in Section 7.2.1. All of the variables included in Table 21, except for ethnicity (which has a high correlation with ψ_i) will be used as instruments. Therefore, z_{it} in equation (25) is a vector of 15 binary variables.

The variables of interest include both an endogenous effect and an exogenous effect. The exogenous effect can be thought of as a measure of undesirable crowding. Assuming that I have captured the endogeneous effect precisely, the exogenous (or unexpected) effect is equal to the impact of the variable of interest that has not captured by the endogenous effect. The exogenous effect will be estimated based on the following equation:

$$y_i = \gamma_0 + \gamma_1.\hat{x}_i + \gamma_2.\hat{x}_i^2 + \gamma_3.(x_i - \hat{x}_i) + \gamma_4.(x_i^2 - \hat{x}_i^2) + \gamma_5.(\text{Control variables})_i + \varepsilon_i \tag{26}$$

where γ_1 to γ_4 are the parameters of interest. The methodology used for measuring the impact of expected versus unexpected levels of the variables of interest is very similar to the technique employed by Barro and Gordon (1981). While this is in a very different context from our study (regarding the government's incentives to surprise the private sector with unexpected inflation), the distinction it draws about expected versus unexpected components is highly relevant to our application.

Based on the model specifications in this section, models will be categorised into three general specifications: Specification I, which does not take into account the selection bias; specification II,

64 As stated earlier, u_{it} follows a binomial distribution. However, the binomial distribution approximates to the normal distribution as the number of observations tend to infinity. Therefore, we can assume an asymptotic normal distribution for u_{it} (Menard, 2002).

which accounts for the selection issue; and specification III, which measures the effect of the exogenous impact. The first specification includes the estimation of equation (23), and the second and third specifications are estimations of equations (26) and (27), respectively.

7.6.1 Control variables

Two ratios are calculated to take into account individuals' own group characteristics in their neighbourhood. The first is the ratio of the number of people living in individual i's MB[65] who have the same ethnicity as the individual to the population of the individual i's MB; that is:

$$\text{METHN} = \frac{\text{The number of people with the same ethnicity as individual i living in his or her neighbourhood}}{\text{Population of individual i's neighbourhood}}$$

This variable measures the desirability of living next to people with the same ethnicity. Five different ethnicities and five dummies are included in the models.[66] A similar measure is defined based on people's occupations; specifically, the OCCN variable is the ratio of people living in individual i's MB and have the same occupation. For this variable I have nine dummies derived from nine different occupation categories.[67]

In order to account for the characteristics of neighbourhoods, METHT is defined as follows:

$$\text{METHT} = \frac{\text{The number of people of a specific ethnicity}}{\text{Population of individual i's MB}}$$

This measure accounts for the desirability of living in areas with different combinations of ethnicities. METHT consists of five dummies derived from five categories of the ethnicity variable. In the same way, MOCCT is constructed as the ratios of different occupations. This measure consists of nine dummy variables.

The other control variables included in the model have been explained in the previous sections. The dummies are gender, marital/partnership status, age groups, length of stay in New Zealand, ethnicity groups, education groups, household income groups, health satisfaction groups, satisfaction with abilities and skills groups, housing and neighbourhood problems, free time groups, socialising groups, recycling behaviour categories, water and energy use categories, council services categories, access to

[65] Different neighbourhood boundaries will be defined. Here, the definition is presented based on mesh-blocks' boundaries.

[66] The name of variable followed by any of the following numbers shows the dummy variable for that specific ethnicity group: 1 - European, 2 - Maori, 3 - Pacific, 4 - Asian, 5 - Other.

[67] Occupation is categorised into 9 aggregate groups: 1 - managers, 2 - professionals, 3 - technicians and trades workers, 4 - community and personal service workers, 5 - clerical and administrative workers, 6 - sales workers, 7 - machinery operators and drivers, 8 - labourers, 9 - residual categories (operational codes only).

facilities,[68] green spaces and coastlines, the state of coastlines and green spaces, and also household size.

The importance of these variables is discussed in Chapter 2. Any of these variables may affect the impact of the variables of interest on residential satisfaction through different channels; that is, not controlling for these variables may lead to omitted relevant variables bias. Here, I briefly explain the importance of controlling for these variables (for details see Section 2.4.4).

The impact of gender should be taken into account as the level of satisfaction for men is usually different from that of women. Males may have a different understanding of the desirable level of crowding, largely because they usually spend less time at home than women, overall. Marital status is important because a couple's use of house is different from singles. The use of space is different amongst age groups. For example, elderly people may require less space than young people due to their lower level of activity (or it may be the other way around). Different ethnicity groups have varying definitions of the desirability of their living environment. The level of education affects residential satisfaction as highly educated people usually have a wider social network and, therefore, are usually less concentrated in their own neighbourhood. Income is the main factor in the ability to choose a desirable living environment. The impact of crowding on residential satisfaction may be affected by household size, as the need for space is different for different households depending on their size.

Satisfaction with health status and satisfaction with abilities and skills are both controls for over-optimism (as discussed in Section 2.2). The range of problems with housing and neighbourhood control for the impact of issues other than crowding that may affect residential satisfaction. As homeowners are more settled and more involved in loss and gains of their dwellings compared to renters, owners may state a higher level of satisfaction with their houses (for details see Section 2.3.2). People with more free time are more likely to use their living environment and may therefore suffer more from an undesirable neighbourhood. Socialising more is usually correlated with more connection with neighbours and therefore being more involved in the living environment. This may lead to more social comparisons and, therefore, less (or more) satisfaction with own level of crowding. Voting is a measure of the importance of the living environment to an individual. Recycling and water/energy use behaviours are measures of the importance of the living space to an individual. People with better access to green-spaces and coastlines may substitute living in smaller dwellings with using green spaces and coastlines. It is also important to account for the quality of green-spaces and coastlines as some people may have access to them but not use them because of low quality.

---

[68] The access to facilities variable takes the proximities to the desired school (school zones) into account. In the GSS's questionnaire, respondents answer the question about the ease of accessing facilities "such as shops, schools, post shops, libraries, medical services and so on".

7.7 Models and results

A range of density measures is used to capture people's understanding of their neighbourhood's density, using residential and population density definitions. For each of these measures, neighbourhood is defined at three different scales: MB, AU and TLA levels. This results in six measures for density. Keeping all other variables, including PPBR, the same, there are six versions of the first specification. Based on the GOF measure (AIC) amongst the six equations, those including the area density measure generally perform better than those containing the residential density measure. Also, amongst the three equations containing area density defined at different geographic scales, the one at the AU scale (DENSA) provides the highest prediction power for RS. In other words, when people evaluate the density of their neighbourhood, they consider the area that is conventionally known as their suburb the most. The details of our estimations for choosing the best measure of density are presented in Appendix ii.

Based on the discussion in Section 7.5, the results of the estimation of the first (non-instrumented) specification is shown on the first column of Table 23, which only presents the outcomes of interest. As shown in Appendix ii, area density at area unit level (DENSA) is chosen as the measure of density because it provides the highest predictive power for RS.[69] Based on the first column, after controlling for a wide range of effects, both household crowding and area density at the AU level (DENSA) lower RS significantly.

PPBR and its squared term are jointly significant at better than the 1 per cent significance level. According to the descriptive statistics, presented in Appendix i, PPBR has a minimum of 0.105 and a maximum of 2 (for detailed description of PPBR see Section 6.5). The impact of a higher crowding is negative at any crowding level; however, this negative impact is lower when the level of crowding is high. The next variable of interest is density (DENS). Both density and its squared term have a negative effect on RS. Their effect is jointly significant at better than the 5 per cent significance level.

In the first model, the null hypothesis of the Durbin and Wu–Hausman tests is that the variable under consideration can be treated as exogenous. With a high significance the null hypothesis of exogeneity is rejected; that is, CG should be treated as endogenous. Also, in the second model, the null hypothesis of the Stock and Yogo test that the set of instruments is weak is rejected at a bias of less than 5 per cent.

The second column of the table presents the results of the estimation of the second specification, based on equation (26). In contrast to the first specification, the impact of PPBR on RS is now

[69] DENSA is compared to DENSM, which is the area density measure constructed based on the mesh-blocks' areas, DENST, which is the area desnity measure constructed based on the TLA area, and DENSD, which is the area desnity measure constructed based on the developable land area.

positive. Therefore, a higher level of expected crowding level raises RS. This may suggest that when people can reach the level of household crowding that they expect to have, they are happier with their residential environment. For instance, people of certain ethnicities may prefer to live in a 'crowded' environment that reflects their cultural norms. However, a higher level of expected area density does not affect RS significantly.

The third column of the table demonstrates the estimated results for equation (27). In order to measure the impact of the difference between the actual and expected levels of the variables of interest, the expected level needs to be controlled for. For example, in order to measure the impact of having a household crowding higher than the expected crowding level ($\mathrm{PPBR} - \widehat{\mathrm{PPBR}}$), it is also necessary to control for the effect of the expected crowding level ($\widehat{\mathrm{PPBR}}$). As depicted, holding all other variables constant, a higher level of exogenous (unexpected) crowding level lowers residential satisfaction. Also, a higher level of exogenous area density lowers residential satisfaction.

In the last column of the table, the third specification is estimated in the presence of perceived crowding (PC). As concluded in the previous chapter, PC has a very high predictive power for RS. Therefore, its presence in the estimation is a check on the robustness of the results. Results show stability in the negative impact of the unexpected household crowding measure and in the unexpected density measure, as well as an increase in the GOF; that is, a decrease in the AIC.

For each model, the joint significance of different pairs of variables are reported at the bottom of the table. The variables included in this joint significance test are symbolised by '†' and '††'. For example, in the first model the joint significance of the crowding variable (PPBR) and its squared form (PPBR^2) are reported at the bottom of the table in the 'Crowding measures†' row.

Table 23. Results of the estimations of the impact of expected and unexpected crowding and density levels.

Model number	1	2	3	4
	Specification I	Specification II	Specificatio n III	Specification III + PC
PPBR	0.171***†			
	(0.113)			
$PPBR^2$	1.553†			
	(0.456)			
$\widehat{PPBR}$		8.369**†	3.706	3.520
		(5.562)	(3.135)	(3.121)
$\widehat{PPBR}^2$		4.983*†	1.136	1.113
		(4.240)	(0.891)	(7.006)
$PPBR - \widehat{PPBR}$			0.168***†	0.233**††
			(0.113)	(0.156)
$PPBR^2 - \widehat{PPBR}^2$			1.555†	1.551††
			(0.464)	(0.462)
DENSA	0.800**††			
	(0.086)			
$DENSA^2$	1.031††			
	(0.019)			
$\widehat{DENSA}$		0.300††	0.288	0.254
		(0.493)	(0.474)	(0.438)
$\widehat{DENSA}^2$		1.104††	1.080	1.112
		(0.297)	(0.289)	(0.313)
$DENSA - \widehat{DENSA}$			0.804**††	0.798**††
			(0.083)	(0.087)
$DENSA^2 - \widehat{DENSA}^2$			1.031*††	1.029††
			(0.018)	(0.019)
PC				0.210***
				(0.050)

Observations	4607	4607	4607	4607
Population size	904053	904053	904053	904053
R^2	0.190	0.185	0.191	0.206
AIC	5046.57	5073.23	5043.27	4961.07
Joint significance (Wald test):				
Crowding measures†	<0.001	0.06	<0.001	<0.001
Density measures††	0.092	0.52	0.023	0.026

Reported coefficients are exponentiated. Standard errors are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

† The significance level for PPBR and its squared (Squared PPBR) variables is reported jointly on the Crowding Wald test row. †† The significance level for the density measure and its squared term is reported jointly on the density Wald test row. Note that the reported joint significance is reported for different coefficients in different models. The reported AICs are for the models that do not take jackknife replications into account.

7.8 Conclusion

The aim of this chapter was to estimate a causal relationship between absolute residential position measures, such as household crowding and density, and residential satisfaction. In order to do this, the measures of absolute residential position must first be chosen appropriately. As discussed in the previous chapter, the chosen measure of household crowding is PPBR, although I have added perceived crowding (PC) as a robustness check in my estimation.

Because density is an area-level measure, I needed to have a correct definition for the area's boundary. In this chapter I checked people's understanding of their neighbourhood's boundaries, in terms of the geographic scale that people feel the density consequences the most, by comparing the GOF of the regression of different measures of density, defined at different geographic levels, on the satisfaction of people with their residential environment.

People choose their living environment so that it suits them the best; the best residential environment for an individual is one that provides him or her with the amenities and facilities that he or she needs. The need is defined based on each individual's background. For example, the needs of someone whose parents are New Zealanders may be defined differently from someone whose parents are foreigners. Therefore, this study uses a two-stage least squares approach to estimate the level of household crowding and the level of density that an individual may expect based on his or her background. Of interest here is the impact of the difference between the actual absolute residential position measures and the expected values on residential satisfaction; that is, the difference between actual and expected levels is an indicator of the levels of crowding and density that an individual does not expect to experience.

Table 24 summarises the results derived from the equations shown in Table 23. The crowding measure (PPBR) has a very significant negative effect. The impact of the expected household crowding on RS is positive; that is, $\widehat{\mathrm{PPBR}}$ raises RS. This shows that when people can satisfy their expectations, they are happy with their residential environment. For example, a member of a four-person household, whose background favours children sharing a bedroom, may choose to live in a dwelling with two bedrooms. This person might consider such a living environment as a positive. This positive significant effect confirms the role of norms in individuals' satisfaction with their residential environment. The unexpected level of crowding lowers RS, which suggests that once the level of crowding is more than the level that an individual may expect to have, he or she is more likely to be unsatisfied with his or her living environment.

Table 24. A summary of the results of chapter 7.[70]

Category	Result	Significance
Household crowding (PPBR)	A higher household crowding level lowers RS	Yes
Household crowding derived from backgrounds ($\widehat{\mathrm{PPBR}}$)	A higher household crowding stemming from 'own group' factors raises RS	Yes
Exogenous household crowding ($\mathrm{PPBR} - \widehat{\mathrm{PPBR}}$)	A higher unexpected level of crowding lowers RS	Yes
Absolute area density (DENSA)	A higher absolute density lowers RS	Yes
Absolute area density derived from backgrounds ($\widehat{\mathrm{DENSA}}$)	A higher area density level stemming from 'own group' attraction does not affect RS	No
Exogenous area density ($\mathrm{DENSA} - \widehat{\mathrm{DENSA}}$)	A higher level of unexpected area density lowers RS	Yes
Perceived crowding (PC)	A higher PC level affects RS negatively	Yes

The comparisons between different measures of our other variable of interest, density, show that area density is the best measure for residential satisfaction in terms of its predictive power. After taking a range of individual and area characteristics into account, a higher area density level (DENSA), measured at the area unit or suburb level, has a significant negative impact on RS. Before taking endogeneity into account, this impact confirms my hypothesis that once the positive impact of agglomeration will be controlled for, people (in Auckland) do not like to live in a densely populated area. When I account for the endogeneity of the area density measure, I find that a higher exogenous area density level lowers RS. Consistent with this study's hypothesis, people choose to live in an area

[70] In addition to the listed results in this table, I also assessed the impact of residential density on RS. The results suggest that residential density does not affect RS significantly.

that has preferred density which is not annoying to them; that is they endogenise their location choice. However, if the area is denser than they expect, due to other constraints, they will not be happy with it.

People's PC is associated with lower RS. Including PC in the equation does not affect the significant negative effects of unexpected household crowding or density.

Amongst the measures of density, the area density measure, which is based on population per area unit (suburb), has the highest predictive power for RS (as presented in Table 28). Moreover, the residential density measure has no significant effect. To choose the best density measure that should be included in the analyses of Table 23, I conducted an analysis of the best fit models. Density measures were constructed based on the ratios and variations at different geographic scales, namely mesh-blocks, Area units and Territorial authorities. The best-fit model contains DENSA, which is the area density of area units. Therefore, people rely on the conventional definition of their suburbs in their evaluations of the living environment.

In relation to previous studies, such as Rodgers (1982), I have expanded the knowledge regarding the importance of socioeconomic factors, especially exogenous ones, by taking the endogeneity of the variables of interest into account. This chapter has controlled for a wide range of variables and has taken spatial proximities into account. The sample population used for the current study is much larger than that in previous studies. To the best of my knowledge, this is the first study to distinguish between expected levels of household position and the unexpected levels of it.

Our results on the impact of density are similar to the results of Song and Knaap's (2003) study, which indicated a lower willingness to pay for houses in denser areas and in areas with more commercial, multifamily and public uses.

Appendix I. Descriptive statistics for chapters 7, 8 and 9

Table 24 describes the continuous variables. For example, DENSA takes values between a minimum of 0.01 and a maximum of 7.73. As depicted in the P_value column, the mean difference amongst years is not statistically significant, at better than the 1 per cent level, and the overall mean is equal to 2.535 (derived from the sample population of 4607 people).

All of the variables in Table 26 are categorical variables. For example, 'Gender' is a binary variable with 0 for females and 1 for males. The mean of this variable is equal to 0.491, indicating that 49.1 per cent of respondents are male and the rest are female.[71] On the far right column, the equality of the mean of each variable between three cross-sections of data – namely 2008, 2010 and 2012 – is tested. The significance level of less than 0.245 for the 'Gender' variable does not reject the null hypothesis that there is no change in the mean of this variable across the three surveys.

Table 25. Descriptive statistics for sample population used in chapters 7, 8 and 9 – continuous variables.

	Mean	P_value	Min	Max	Std error
DENSM	3.695	<0.001	0.002	154.905	0.12
DENSA	2.535	<0.001	0.01	7.73	0.018
DENST	767.694	<0.001	37.814	1631.407	6.681
MADENS	0.002	<0.001	0	0.023	0
MTDENS	0.008	<0.001	0	0.2	0
ATDENS	4.753	<0.001	0.021	70.473	0.092
MTVARDENS	-1827.06	<0.001	-6323.2	-2.79	31.779
ATVARDENS	10428.4	<0.001	-2195.97	114966.1	200.447
MAVARDENS	48550.65	<0.001	-8554.97	19338442	14037.06
PPBR	1.016	<0.001	0.2	2	0.003
$\widehat{PPBR}$	1.028	<0.001	0.78	1.674	0.002
CM	1.025	<0.001	0.644	2.471	0.004
CA	1.027	<0.001	0.798	1.921	0.003
CT	1.031	<0.001	0.889	1.139	0.001
MC	0.9	<0.001	0.101	3.105	0.007
AC	0.888	<0.001	0.112	2.505	0.006
TC	0.876	<0.001	0.092	2.249	0.006
MVARC	-0.279	<0.001	-13.194	7.374	0.021
AVARC	-0.371	<0.001	-11.149	7.829	0.026
TVARC	-0.756	<0.001	-13.719	15.828	0.058
THVARC	-0.514	<0.001	-16.661	22.175	0.086
THVARC95	-0.514	<0.001	-16.661	22.175	0.086
AHVARC	-0.299	<0.001	-8.498	10.492	0.023
AHVARC95	-0.299	<0.001	-8.517	10.515	0.023
MHVARC	-0.189	<0.001	-9.604	9.604	0.019

[71] In a comparison between the statistics reported in this table and Table 2, which describes the data for the full sample population available for Auckland, the mean for Gender has changed from 0.485 to 0.491.

CVARS	-0.283	<0.001	-13.828	7.721	0.022
$\widehat{\text{CVARS}}$	-0.197	<0.001	-3.726	2.249	0.011
$\text{CVAR} - \widehat{\text{CVARS}}$	-0.086	<0.001	-11.583	6.803	0.019
$\text{PPBR} - \widehat{\text{PPBR}}$	-0.044	<0.001	-1.109	1.434	0.005
$\text{DENSA} - \widehat{\text{DENSA}}$	0.054	0.002	-4.087	5.64	0.017
$\widehat{\text{ATVARDENS}}$	10666.98	<0.001	-11072.6	58836.92	100.772
$\text{ATVARDENS} - \widehat{\text{ATVARDENS}}$	-238.577	0.189	-33123.3	99807.66	181.406
DEP	990.147	<0.001	866	1436	1.417
RDEP	-19.922	<0.001	-423	447	1.86
DUMMY	0.414	<0.001	0	1	0.008
$\widehat{\text{RDEP}}$	-19.922	<0.001	-423	447	1.86
$\widehat{\text{DEP}}$	986.572	<0.001	873.955	1236.42	0.81

Note: Reported statistics account for jackknife resampling weights.

Table 26. Descriptive statistics for sample population used in chapters 7, 8 and 9 – dummy variables.

Variables	Description	Group	Mean	Mean difference significance
Gender	0 = 'female', 1 = 'male'	-	0.491	0.245
Partner	0 = 'Non-partnered', 1 = 'Partnered'	-	0.638	0.216
Age	= 15–19	i	0.041	0.718
	= 20–24	ii	0.107	
	= 25–34	iii	0.196	
	= 35–54	iv	0.389	
	= 55–59	v	0.074	
	= 60–64	vi	0.057	
	= 65–69	vii	0.049	
	= 70–74	viii	0.035	
	>75	ix	0.052	
Length of living in NZ	< 4 years	i	0.046	0.762
	= 4–10 years	ii	0.127	
	= 10–25 years	iii	0.143	
	>25 years	iv	0.684	
Ethnicity	European	i	0.64	0.189
	Maori	ii	0.08	
	Pacific	iii	0.08	
	Asian	iv	0.15	
	Other	v	0.05	
METHT	European	i	0.582	0.036
	Maori	ii	0.119	
	Pacific	iii	0.129	
	Asian	iv	0.181	
	Other	v	0.099	
METHN	European	i	0.417	0.043

	Maori	ii	0.017	
	Pacific	iii	0.024	
	Asian	iv	0.046	
	Other	v	0.004	
MOCCT	Managers	i	0.123	0.082
	Professionals	ii	0.203	
	Technicians and trades workers	iii	0.122	
	Community and personal service workers	iv	0.073	
	Clerical and administrative workers	v	0.135	
	Sales workers	vi	0.101	
	Machinery operators and drivers	vii	0.057	
	Labourers	viii	0.075	
	Residual category	ix	0.057	
MOCCN	Managers	i	0.041	0.056
	Professionals	ii	0.027	
	Technicians and trades workers	iii	0.021	
	Community and personal service workers	iv	0.005	
	Clerical and administrative workers	v	0.007	
	Sales workers	vi	0.009	
	Machinery operators and drivers	vii	0.003	
	Labourers	viii	0.003	
	Residual category	ix	0.002	
Education	= No qualification	i	0.104	0.076
	= Certificates	ii	0.479	
	= Degree	iii	0.417	
Household income	=Zero	i	0.005	0.034
	=$1–$5,000	ii	0.003	
	=$5,001–$50,000	iii	0.281	

	=$50,001–$70,000	iv	0.092	
	=$70,001–$150,000	v	0.417	
	=$150,001 or more	vi	0.202	
Household size	= One person	i	0.266	0.482
	= Two people	ii	0.214	
	= Three people	iii	0.236	
	= Four people	iv	0.114	
	= Five people	v	0.048	
	= Six people	vi	0.017	
	= Seven people	vii	0.005	
	>= Eight people	viii	0.004	
Health perception	= Very dissatisfied	i	0.019	0.015
	= Dissatisfied	ii	0.089	
	= No feeling either way	iii	0.244	
	= Satisfied	iv	0.366	
	= Very satisfied	v	0.282	
Ability perception (includes individual's feeling about his/her knowledge, skills and abilities)	= Very dissatisfied	i	0.002	0.024
	= Dissatisfied	ii	0.042	
	= No feeling either way	iii	0.07	
	= Satisfied	iv	0.628	
	= Very satisfied	v	0.258	
Residential Satisfaction	= Satisfied	i	0.365	0.275
	= Dissatisfied	ii	0.635	
Free time (Individual's feeling about having enough free time)	= Too much free time	i	0.438	0.012
	= The right amount of free time	ii	0.474	
	= Not enough free time	iii	0.088	
Socialising (Frequency of meeting friends)	= Every day	i	0.089	0.119
	= Around 3–6 times a week	ii	0.157	

	= Around 1–2 times a week	ii	0.388	
	= Around once a fortnight	iii	0.215	
	= At least once in the last four weeks	iv	0.151	
Local political involvement	0 = 'No', 1 = 'Yes'	-	0.365	0.401
Recycling (How much does the household recycle)	= None	i	0.01	0.382
	= A little	i	0.018	
	= Some	ii	0.111	
	= Most	iii	0.561	
	= All	iv	0.301	
Council services (Individual's feeling about the quality of council services)	= Very dissatisfied	i	0.027	0.011
	= Dissatisfied	ii	0.101	
	= No feeling either way	iii	0.13	
	= Satisfied	iv	0.606	
	= Very satisfied	v	0.136	
Green space access (Access to local green spaces including bushes, forests, nature reserves)	= Never want or need to go	i	0.008	0.122
	= None or a few of them	ii	0.052	
	= Some of them	iii	0.15	
	= Most of them	iv	0.449	
	= All of them	v	0.341	
Green space state	= Not been to	i	0.008	0.072
	= Very dissatisfied	ii	0.002	
	= Dissatisfied	iii	0.031	
	= No feeling either way	iv	0.087	
	= Satisfied	v	0.648	
	= Very satisfied	vi	0.224	
Coastline access (Access to local lakes, rivers, harbours and oceans)	= Never want or need to go	i	0.012	0.289
	= None or a few of them	i	0.05	
	= Some of them	i	0.146	
	= Most of them	ii	0.437	
	= All of them	iii	0.355	
Coastline state (Household's feeling about	= Not been to	i	0.007	0.149
	= Very dissatisfied	ii	0.005	

the state of coastlines)	= Dissatisfied	iii	0.083	
	= No feeling either way	iv	0.107	
	= Satisfied	iv	0.642	
	= Very satisfied	v	0.156	
Water use (How often does the household try to minimise water use?)	= Never	i	0.085	0.601
	= A little of the time	ii	0.11	
	= Some of the time	iii	0.268	
	= Most of the time	iv	0.397	
	= All of the time	iv	0.14	
Energy use (How often does the household try to minimise energy use?)	= Never	i	0.022	0.011
	= A little of the time	ii	0.083	
	= Some of the time	ii	0.303	
	= Most of the time	iii	0.474	
	= All of the time	iii	0.118	
Facilities access (including shops, schools, post shops, libraries and medical services)	= Never want or need to go to any of them	i	0.001	0.566
	= None of them	ii	0.004	
	= A few of them	ii	0.021	
	= Some of them	iii	0.057	
	= Most of them	iii	0.294	
	= All of them	iv	0.624	
Year	= 2008	-	0.335	
	= 2010	-	0.344	
	= 2012	-	0.321	
Homeownership status	= Owned	i	0.521	0.261
	= Not owned	ii	0.321	
	= Family trust	iii	0.159	
Perceived crowding	Being too small is a major problem with the person's house/flat? 0 = 'No', 1 = 'Yes'		0.106	0.696
Bad street access	Being hard to get to from the street is a major problem with the person's house/flat? 0 = 'No', 1 = 'Yes'		0.017	0.037
Poor condition	Being in poor condition is a major problem with the person's house/flat? 0 = 'No', 1 = 'Yes'		0.058	0.292

Damp dwelling	Being damp is a major problem with the person's house/flat? 0 = 'No', 1 = 'Yes'	0.104	0.716
Difficult to heat	Being too cold, or difficult to heat/keep warm is a major problem with the person's house/flat? 0 = 'No', 1 = 'Yes'	0.146	0.85
Having pests	Having pests such as mice or insects is a major problem with the person's house/flat? 0 = 'No', 1 = 'Yes'	0.06	0.316
Expensive house	Being too expensive is a major problem with the person's house/flat? 0 = 'No', 1 = 'Yes'	0.074	0.167
Observations	4607		

Note: Reported statistics account for jackknife resampling weights.

Table 27. PPBR versus PC (small-house problem).

PPBR	PC: Small-house problem	
	No	Yes
$0 < PPBR \leq 0.3$	98	2
$0.3 < PPBR \leq 0.6$	96	4
$0.6 < PPBR \leq 0.9$	96	4
$0.9 < PPBR \leq 1.3$	87	13
$1.3 < PPBR \leq 2$	78	22
$2 < PPBR$	65	35
Number of observations	4607	

Note: Numbers are in percentages. Reported statistics account for jackknife resampling weights.

Appendix II. Table of estimates using different measures of density

Table 28 illustrates the estimation results using different measures of density. In Table 28, DENSM is the area density measure constructed based on the mesh-blocks' areas, DENST is the area density measure constructed based on the TLA area and DENSD is the area density measure constructed based on the developable land area. Model 1 has a better GOF than models 4 and 5. As the developable land area's dataset does not match well with our dataset, the GOF of this model is not comparable with other models. As DENSA has the best predictive power for RS, it is used in all of the estimations presented in Table 23. For a description of control variables, see Section 7.6.1.

Table 28. Estimations including different measures of area density.

	(1) specification I DENSA	(2) specification II DENSA	(3) specification III DENSA	(4) specification I - DENSM	(5) specification I - DENST	(6) specification I - DENSD
Residential Satisfaction (dummy)						
PPBR	0.170919207***			0.165544940***	0.167312002***	0.139650258***
	(0.113)			(0.112)	(0.113)	(0.104)
$PPBR^2$	1.552999033			1.581001528	1.571703218	1.711340281
	(0.456)			(0.473)	(0.469)	(0.556)
DENSA	0.799654432**					
	(0.086)					
$DENSA^2$	1.031303329					
	(0.019)					
$\widehat{PPBR}$		8.369763777**	3.706583671			
		(5.562)	(3.135)			
$\widehat{PPBR}^2$		4.982979046*	1.13696949**			
		(4.240)	(0.891)			
$\widehat{DENSA}$		0.300399760	0.254155396			
		(0.493)	(0.438)			
$\widehat{DENSA}^2$		1.104201966	1.112			
		(0.297)	(0.313)			
$PPBR - \widehat{PPBR}$			0.168705366***			
			(0.113)			
$PPBR2 - \widehat{PPBR}^2$			1.555431389			
			(0.464)			
$DENSA - \widehat{DENSA}$			0.803773952**			
			(0.083)			
$DENSA^2 - \widehat{DENSA}^2$			1.030871293*			
			(0.018)			
$DENSM^2$				1.000000000		
				(0.000)		
DENSM				0.999984640		
				(0.000)		
$DENST^2$					1.127276136	

					(0.348)	
DENST					0.561764447	
					(0.348)	
DENSD²						1.000001699
						(0.000)
DENSD						0.997788986
						(0.002)
MOCCN_1	3.635751965	4.249670997*	4.889886470*	3.664782683	3.627137612	3.297438242
	(3.135)	(3.496)	(4.191)	(3.161)	(3.146)	(3.060)
MOCCN_2	1.297835311	1.382762733	1.498893320	1.303870352	1.285471262	1.049680758
	(0.726)	(0.769)	(0.847)	(0.728)	(0.716)	(0.665)
MOCCN_3	0.659457022	0.627774381	0.723330808	0.652953369	0.712109480	0.710557177
	(0.798)	(0.738)	(0.849)	(0.783)	(0.852)	(0.880)
MOCCN_4	1.953573400	2.388155691	4.468158730	1.973905417	1.919255617	1.392771912
	(4.256)	(5.190)	(9.759)	(4.365)	(4.180)	(3.237)
MOCCN_5	0.388314811	0.354923511	0.609135615	0.359595521	0.412257107	0.143125750
	(0.578)	(0.534)	(0.913)	(0.536)	(0.618)	(0.252)
MOCCN_6	4.152901805	3.869229132	4.987197374	3.833134181	4.271762306	3.840812411
	(6.852)	(6.199)	(8.002)	(6.329)	(6.988)	(6.749)
MOCCN_7	8.335679323	8.354049132	10.309887650	7.775911900	8.783797624	38.546514896
	(45.760)	(42.539)	(54.382)	(41.970)	(48.765)	(209.505)
MOCCN_8	1.978316981	1.809730621	1.478576504	1.445387559	1.344705686	11.748463586
	(6.191)	(5.400)	(4.698)	(4.450)	(4.093)	(41.024)
MOCCN_9	2.95694e+02	3.19279e+02	1.39175e+02	2.72004e+02	2.78622e+02	7.19904e+02
	(2163.094)	(2199.077)	(939.585)	(1977.531)	(1976.012)	(4956.089)
MOCCT_1	1.926085321	1.675664364	1.601857802	1.544173789	2.146574849	1.838310805
	(1.801)	(1.558)	(1.485)	(1.462)	(2.053)	(1.855)
MOCCT_2	0.998942126	0.772030519	0.755962637	0.849446521	1.048734464	0.678527798
	(0.610)	(0.462)	(0.457)	(0.523)	(0.713)	(0.444)
MOCCT_3	0.074741869**	0.090064154**	0.089002411**	0.082265784**	0.072540613**	0.080289441**
	(0.080)	(0.094)	(0.094)	(0.088)	(0.075)	(0.091)
MOCCT_4	0.346021439	0.309498025	0.344231514	0.371581613	0.361584579	0.237864850
	(0.392)	(0.337)	(0.373)	(0.417)	(0.411)	(0.292)
MOCCT_5	0.251676813	0.295117669	0.284803792	0.209496965	0.162917114	0.215757556
	(0.284)	(0.326)	(0.313)	(0.234)	(0.178)	(0.243)
MOCCT_6	0.196132831	0.164217068	0.144616533	0.170409417	0.162728835	0.196070428
	(0.235)	(0.190)	(0.172)	(0.206)	(0.191)	(0.228)
MOCCT_7	0.981877692	1.290403643	1.193968759	1.091647103	0.596830110	1.464810301
	(1.740)	(2.284)	(2.150)	(1.939)	(1.127)	(2.565)
MOCCT_8	0.942945294	1.267989426	1.183795266	1.041152180	1.188319396	1.280921017
	(1.046)	(1.372)	(1.314)	(1.159)	(1.297)	(1.525)
MOCCT_9	2.057128202	1.819537806	2.217721857	2.299077696	2.704034266	2.566983744
	(2.531)	(2.223)	(2.807)	(2.744)	(3.211)	(3.255)
METHN_1	1.421900934	1.514954172	1.601035482	1.561386518	1.542892673	1.379177036
	(0.995)	(1.101)	(1.208)	(1.109)	(1.082)	(1.053)
METHN_2	3.002689068	2.470783247	3.033961807	2.605116769	2.870036818	4.246274652

	(4.331)	(3.647)	(4.542)	(3.782)	(4.215)	(7.076)
METHN_3	5.494923057**	6.784408541**	6.073720445**	6.055427046**	6.083354884**	7.195654991**
	(4.448)	(5.588)	(5.099)	(4.918)	(4.892)	(6.538)
METHN_4	1.840320090	2.166090365	1.937178627	2.185490717	1.662994170	3.250379179
	(1.827)	(2.082)	(1.965)	(2.175)	(1.669)	(3.468)
METHN_5	0.061698844	0.117638810	0.085936399	0.066191679	0.079281342	0.126932672
	(0.308)	(0.574)	(0.405)	(0.325)	(0.388)	(0.633)
METHT_1	5.416796734	4.716022188	6.029686688	5.392054371	5.419821168	6.475365784
	(6.215)	(5.403)	(7.241)	(6.324)	(6.230)	(8.352)
METHT_2	0.960030660	0.759332233	0.897475806	0.987264157	0.908599041	0.614225922
	(0.965)	(0.764)	(0.916)	(0.986)	(0.894)	(0.670)
METHT_3	1.761064197	0.950325475	1.279789138	1.330618132	1.396565094	1.357491739
	(1.892)	(1.015)	(1.413)	(1.414)	(1.539)	(1.513)
METHT_4	4.199616982	2.836858175	3.782417780	3.725205296	4.590581869	2.936452666
	(4.969)	(3.332)	(4.619)	(4.378)	(5.350)	(3.632)
METHT_5	15.473994015**	13.547921057*	22.169368342**	14.704862314*	15.855534612*	11.623527950
	(21.309)	(19.149)	(31.123)	(20.778)	(22.322)	(17.600)
Gender	0.954949596	1.028896566	1.046321149	0.959144498	0.962192154	0.973061523
	(0.097)	(0.114)	(0.116)	(0.096)	(0.098)	(0.092)
Partner	1.134585307	0.404112691*	0.393852223*	1.130604322	1.132300209	1.075769045
	(0.154)	(0.209)	(0.212)	(0.152)	(0.154)	(0.163)
Age (>74)						
= 15–19	1.172607122	0.193579257	0.166641783	1.207950592	1.168938446	1.092905592
	(0.491)	(0.233)	(0.207)	(0.509)	(0.493)	(0.487)
= 20–24	0.811783301	0.222761439	0.197773198	0.827575973	0.812177047	0.737280113
	(0.249)	(0.238)	(0.218)	(0.256)	(0.257)	(0.238)
= 25–34	0.473736799***	0.209832616**	0.203751790**	0.480394056***	0.477899814***	0.468123260**
	(0.129)	(0.149)	(0.149)	(0.131)	(0.133)	(0.138)
= 35–54	0.478475277***	0.231445642**	0.229451659**	0.492776475***	0.478378619***	0.484514515***
	(0.116)	(0.135)	(0.138)	(0.118)	(0.118)	(0.121)
= 55–59	0.738406945	0.831801038	0.878754390	0.751776587	0.747310310	0.704939954
	(0.183)	(0.231)	(0.247)	(0.184)	(0.185)	(0.187)
= 60–64	0.870479860	1.181213890	1.200736814	0.892617915	0.863761914	0.967072490
	(0.241)	(0.322)	(0.329)	(0.246)	(0.240)	(0.273)
= 65–69	0.838726621	1.201194703	1.217015746	0.848830808	0.821286110	0.917687668
	(0.183)	(0.285)	(0.294)	(0.186)	(0.182)	(0.212)
= 70–74	0.824273586	1.002933765	0.993403458	0.828832151	0.812672294	0.870759601
	(0.192)	(0.250)	(0.252)	(0.194)	(0.191)	(0.215)
Length of stay in NZ (<4 years)						
= 4–10 years	1.033471266	1.261406549	1.352398673	1.033653538	1.038744067	1.283715530
	(0.338)	(0.452)	(0.499)	(0.339)	(0.341)	(0.452)
= 10–25 years	0.858821798	1.171254880	1.236286602	0.857779095	0.851088372	1.074407529
	(0.267)	(0.497)	(0.552)	(0.268)	(0.266)	(0.377)
>25 years	1.238066228	2.144757672*	2.421837892*	1.233560166	1.221880060	1.678231002
	(0.388)	(0.971)	(1.159)	(0.382)	(0.381)	(0.606)

Ethnicity (European)						
= Maori	1.204509792	0.780491437	0.759608541	1.324918939	1.307735918	1.094504177
	(0.663)	(0.495)	(0.484)	(0.736)	(0.717)	(0.657)
= Pacific	0.763378937	0.297846082	0.276063916	0.783098889	0.780640320	0.666497929
	(0.386)	(0.316)	(0.297)	(0.405)	(0.397)	(0.378)
= Asian	0.754449812	0.601295173	0.603945069	0.763112422	0.813621343	0.712254894
	(0.421)	(0.461)	(0.472)	(0.431)	(0.458)	(0.430)
= Other	1.944031388	1.420414244	1.394268986	2.046293479	1.970817562	2.004188535
	(1.331)	(1.179)	(1.153)	(1.393)	(1.330)	(1.411)
Education (No qualification)						
= Diploma	0.989813741	1.399385027	1.438235945	0.986125971	0.998809611	1.219941961
	(0.203)	(0.458)	(0.480)	(0.199)	(0.204)	(0.251)
= Degree	0.898732647	1.377245979	1.438491920	0.903232316	0.915833523	1.036654117
	(0.189)	(0.462)	(0.494)	(0.187)	(0.193)	(0.220)
Household income ($150,001 or more)						
= Zero	0.808236327	2.917274980	3.173691801	0.858183832	0.851716962	1.008156907
	(0.436)	(2.496)	(2.820)	(0.447)	(0.447)	(0.668)
= $1–$5,000	3.127370653	7.014648393	6.854952779	3.273003247	3.248023312	1.115267226
	(3.116)	(9.350)	(9.190)	(3.722)	(3.378)	(1.078)
= $5000–$50,001	0.786016987	1.232412541	1.331761565	0.784620768	0.785416282	0.813624462
	(0.125)	(0.423)	(0.474)	(0.126)	(0.126)	(0.142)
= $50,001–$70,000	0.844559731	0.992224914	1.067155743	0.845955733	0.837966132	0.854582161
	(0.157)	(0.248)	(0.269)	(0.156)	(0.152)	(0.168)
= $70,001–$150,000	0.857290141	0.993017521	1.017194087	0.850910482	0.845022037	0.913404853
	(0.117)	(0.169)	(0.180)	(0.116)	(0.115)	(0.135)
Health (Very dissatisfied)						
= Dissatisfied	2.504733850**	3.439067802***	3.566468692***	2.497386667**	2.437511646**	2.036836428*
	(0.953)	(1.396)	(1.482)	(0.944)	(0.906)	(0.798)
= No feeling either way	2.971713597***	4.889062121***	5.212151293***	2.993588207***	2.906664301***	2.433360349**
	(0.987)	(1.932)	(2.104)	(0.987)	(0.928)	(0.915)
= Satisfied	3.404267173***	6.050565676***	6.430429690***	3.417736813***	3.329445006***	2.765032474***
	(1.133)	(2.524)	(2.765)	(1.126)	(1.067)	(1.027)
= Very Satisfied	4.326584159***	7.093314066***	7.610454928***	4.381918976***	4.256439436***	3.615575074***
	(1.520)	(2.873)	(3.194)	(1.524)	(1.449)	(1.372)
Abilities (Very dissatisfied)						
= Dissatisfied	1.093366735	0.674272248	0.444520092	1.103938766	0.991375521	3.394859650
	(1.635)	(1.075)	(0.648)	(1.615)	(1.449)	(9.636)
= No feeling either way	1.265957686	1.109089473	0.695665739	1.225687108	1.150545186	3.379484258
	(1.827)	(1.685)	(0.960)	(1.732)	(1.621)	(9.684)
= Satisfied	1.697713627	1.621296353	1.001174876	1.672167500	1.545982745	4.718926338
	(2.446)	(2.414)	(1.353)	(2.357)	(2.174)	(13.427)

= Very Satisfied	2.691930402	3.113831636	1.968463196	2.647275883	2.415553617	7.860290705
	(3.916)	(4.641)	(2.649)	(3.767)	(3.427)	(22.326)
Bad street access	0.551248788	0.593016667	0.646861946	0.558235692	0.552274314	0.634465725
	(0.277)	(0.303)	(0.326)	(0.277)	(0.274)	(0.334)
Poor condition	0.319117890***	0.278613184***	0.303599799***	0.317346486***	0.323693088***	0.324248582***
	(0.100)	(0.083)	(0.096)	(0.100)	(0.100)	(0.106)
Damp dwelling	0.546931090***	0.549648608***	0.595536560**	0.543316109***	0.545432474***	0.552810385***
	(0.113)	(0.119)	(0.135)	(0.112)	(0.111)	(0.117)
Difficult to heat	0.458085770***	0.413698859***	0.410263228***	0.451308271***	0.458487721***	0.435284792***
	(0.090)	(0.087)	(0.087)	(0.087)	(0.089)	(0.095)
Having pests	1.401349676	1.119804418	1.068830313	1.385966925	1.415533139	1.346880388
	(0.383)	(0.384)	(0.377)	(0.379)	(0.390)	(0.405)
Expensive house	0.563152347**	0.579988168**	0.593658450**	0.570279422**	0.584451711**	0.612383217*
	(0.140)	(0.147)	(0.148)	(0.139)	(0.141)	(0.155)
Work distance	1.069742514	0.980797022	1.079916818	1.074684160	1.062194779	1.151109615
	(0.260)	(0.298)	(0.340)	(0.259)	(0.253)	(0.278)
Far from facilities	0.582375154	0.499210763	0.491893667	0.590436504	0.552397850	0.513886994
	(0.251)	(0.264)	(0.265)	(0.256)	(0.240)	(0.258)
Unsafe neighbourhood	0.758427271	0.935346231	1.022475908	0.749636441	0.752560078	0.755646685
	(0.218)	(0.438)	(0.497)	(0.215)	(0.218)	(0.240)
Noise and vibration	0.612201448***	0.533600114*	0.549390659*	0.622286836***	0.616608059***	0.635427855***
	(0.098)	(0.189)	(0.195)	(0.100)	(0.098)	(0.108)
Air polluted	0.944198392	1.132939859	1.134387917	0.957103157	0.942435168	1.104499258
	(0.230)	(0.336)	(0.353)	(0.230)	(0.232)	(0.272)
Tenure status (Owner)						
= Renter	0.664270158***	0.408393459**	0.395196564**	0.665878248***	0.655667919***	0.646952825***
	(0.079)	(0.156)	(0.156)	(0.079)	(0.080)	(0.081)
= Family trust	1.176281806	1.174143448	1.176582129	1.170495839	1.157067861	1.065855197
	(0.132)	(0.145)	(0.147)	(0.132)	(0.133)	(0.137)
Free Time (Not enough free time)						
= The right amount of free time	1.257514145**	1.324495797***	1.250759937**	1.259708086**	1.268901403**	1.235657173*
	(0.126)	(0.140)	(0.133)	(0.126)	(0.127)	(0.132)
= Too much free time	1.233235443	1.211696724	1.143208367	1.225242826	1.249887957	1.070089456
	(0.246)	(0.240)	(0.229)	(0.243)	(0.247)	(0.234)
Socialising (Every day)						
= At least once in The last four weeks	0.858148852	0.583474018**	0.560702075**	0.861272791	0.858469767	0.908389489
	(0.147)	(0.130)	(0.125)	(0.148)	(0.147)	(0.172)
= Around once a fortnight	0.823411921	0.721379176*	0.697880144*	0.830476841	0.840491338	0.821189250
	(0.117)	(0.141)	(0.141)	(0.116)	(0.118)	(0.123)
= Around 1–6 times a week	0.899135270	0.945191646	0.928888727	0.900378135	0.909056177	0.925373311
	(0.088)	(0.096)	(0.096)	(0.088)	(0.089)	(0.103)
Local political involvement	0.956203883	0.820447942*	0.835959336	0.964085193	0.959008541	1.065887680

	(0.105)	(0.094)	(0.099)	(0.106)	(0.105)	(0.120)
Recycle==None/A little	2.035175348	0.862182044	0.898597972	2.114486955	2.021286387	1.869912206
	(1.235)	(0.597)	(0.634)	(1.290)	(1.200)	(1.380)
Recycle== Some	0.817902823	0.612595587	0.599124569	0.847868550	0.816132483	0.835074310
	(0.297)	(0.233)	(0.223)	(0.314)	(0.304)	(0.337)
Recycle==Most	0.592842312***	0.601199913***	0.591185334***	0.599014341***	0.604036936***	0.676316736**
	(0.109)	(0.113)	(0.109)	(0.109)	(0.110)	(0.128)
Water use (Never)						
= A little of the time	1.105309914	1.364743511*	1.377123799*	1.086663557	1.112158188	1.141655740
	(0.157)	(0.230)	(0.242)	(0.156)	(0.159)	(0.170)
= Some of the time	1.106780664	1.319108696*	1.301865531	1.092672516	1.125661955	1.118804728
	(0.159)	(0.208)	(0.214)	(0.155)	(0.164)	(0.172)
= Most of the time	0.976016277	1.140948309	1.167301070	0.971701097	0.982992322	0.949540830
	(0.108)	(0.149)	(0.158)	(0.107)	(0.110)	(0.105)
Energy use (Never)						
= A little of the time	0.642346953**	0.584994289	0.566190054*	0.633847157**	0.625255671**	0.560537971***
	(0.116)	(0.197)	(0.190)	(0.115)	(0.114)	(0.119)
= Most of the time	0.842478420	0.764672012	0.780646087	0.840674072	0.838744153	0.766401798**
	(0.092)	(0.124)	(0.129)	(0.092)	(0.091)	(0.089)
Council services (Very satisfied)						
= Very Dissatisfied	0.910476672	0.751932571	0.714281102	1.000195588	0.890630107	0.988113899
	(0.349)	(0.867)	(0.861)	(0.374)	(0.347)	(0.436)
= Dissatisfied	0.510589717***	0.572568653	0.565181210	0.532942597***	0.508660335***	0.541366556***
	(0.090)	(0.320)	(0.328)	(0.095)	(0.092)	(0.105)
= No feeling either way	0.404739720***	0.559537982	0.551388045	0.411613998***	0.408343022***	0.398249420***
	(0.068)	(0.227)	(0.229)	(0.070)	(0.069)	(0.075)
= Satisfied	0.439321782***	0.487692000***	0.485092471***	0.446628031***	0.438476595***	0.454933325***
	(0.056)	(0.108)	(0.108)	(0.057)	(0.057)	(0.059)
Facilities access (All of them)						
= Never been to	0.024037858	0.049588135	0.054793201	0.022031515	0.023144126	0.021532189
	(0.096)	(0.198)	(0.215)	(0.089)	(0.092)	(0.086)
= A few of them	0.219940046	0.235272444	0.209804745	0.211849161	0.218656170	0.293391706
	(0.346)	(0.379)	(0.352)	(0.350)	(0.363)	(0.499)
= None of them	0.838890136	0.510076280	0.475421799	0.869338108	0.838021372	0.832452200
	(0.432)	(0.295)	(0.268)	(0.462)	(0.437)	(0.478)
Green space access (Few of them)						
= Most of them	0.445486692	0.216943628	0.232137274	0.445465067	0.458266818	0.402876900
	(0.356)	(0.202)	(0.220)	(0.362)	(0.369)	(0.335)
= All of them	1.000905936	1.006489174	0.995188642	1.014448109	1.021199379	0.980440178
	(0.291)	(0.290)	(0.295)	(0.292)	(0.298)	(0.304)
Green space state (Very satisfied)						
= Not been to	0.267565738*	0.598448367	0.562506091	0.262005077*	0.256423539*	0.465916417
	(0.196)	(0.487)	(0.469)	(0.194)	(0.187)	(0.341)
= Very Dissatisfied	1.490578341	0.778715116	0.718325255	1.483871576	1.424739617	2.025255590
	(2.340)	(1.319)	(1.185)	(2.432)	(2.316)	(2.553)

= Dissatisfied	0.884208427	0.920853785	0.908847637	0.871281925	0.863390355	0.779908100
	(0.271)	(0.284)	(0.285)	(0.263)	(0.263)	(0.254)
= No feeling either way	1.023269143	0.918358575	0.918512067	1.032309242	1.007799844	0.998741947
	(0.237)	(0.248)	(0.250)	(0.239)	(0.231)	(0.242)
Coastline access (Few of them)						
= Most of them	1.991461855	2.293063877	2.471738316	1.972695547	1.925709666	1.816146208
	(1.075)	(1.283)	(1.431)	(1.061)	(1.046)	(0.986)
= All of them	1.152375465	0.806574943	0.805339677	1.158315145	1.137050831	1.359113265
	(0.320)	(0.287)	(0.299)	(0.321)	(0.317)	(0.386)
Coastline state (Very satisfied)						
= Not been to	2.173437921	2.652625465	2.451193622	2.143006489	2.107189596	1.748720295
	(1.560)	(2.091)	(1.968)	(1.526)	(1.495)	(1.289)
= Very Dissatisfied	1.553958627	1.935186117	1.917147045	1.473633297	1.588034087	1.487283342
	(1.128)	(1.387)	(1.359)	(1.054)	(1.161)	(1.125)
= Dissatisfied	0.983825910	1.161620576	1.198738448	0.979886190	0.994910799	0.993536387
	(0.195)	(0.272)	(0.280)	(0.194)	(0.196)	(0.214)
= Satisfied	0.880371574	1.045597700	1.027145552	0.866918178	0.882352524	0.836646722
	(0.155)	(0.204)	(0.197)	(0.151)	(0.156)	(0.152)
Household size (>4 people)						
= Single person	0.357114572***	0.851976906	0.850859433	0.357045471***	0.366429030***	0.311232823***
	(0.118)	(0.166)	(0.164)	(0.116)	(0.119)	(0.112)
= 2 people	0.517787799**	0.891225670	0.860215790	0.528613777**	0.525866655**	0.506375756**
	(0.130)	(0.157)	(0.151)	(0.130)	(0.129)	(0.140)
= 3 people	0.570191536***	0.765940503	0.759346537	0.573420139***	0.583324013***	0.585479765**
	(0.117)	(0.131)	(0.135)	(0.115)	(0.116)	(0.132)
= 4 people	0.720345057	0.815444173	0.824864073	0.723771771	0.730898269	0.697100898
	(0.151)	(0.156)	(0.161)	(0.150)	(0.152)	(0.155)
Constant	0.937989283	0.016827983	0.018158035	0.747222479	1.031042139	0.267021519
	(2.038)	(0.065)	(0.072)	(1.622)	(2.223)	(0.920)
Observations	4607	4607	4607	4607	4607	4091
N_pop	9.04053e+05	9.04053e+05	9.04053e+05	9.04053e+05	9.04053e+05	7.90587e+05
AIC	5046.57	5073.23	5043.27	5051.198	5073.229	5092.918

Reported coefficients are exponentiated ones. Standard errors are reported in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The reported AICs are for the models that do not take into account jackknife replications.

The base group for each variable is shown in parentheses. For example, for the Age variable, the 'more than 74 years old' category is supposed as the base group.

8 Residential Satisfaction, Relative Crowding and Relative Density: Evidence over Different Geographic Scales

Abstract

This study investigates the impact of relative residential positions on the satisfaction of residents. In order to do this, it was necessary to find an optimal reference group and an optimal set of subjective measures for social comparisons. Concentrating on objective measures of an individual household's situation may be insufficient when dealing with the determinants of satisfaction with residential environment. There is a long history of relative measurement in the economics literature, especially in welfare studies, in which 'keeping up with the Joneses' may play an important role in people's evaluations of their different life circumstances. Conceptually, the reference group for the relative measurement refers to people who affect the individual the most. Neighbours, as a group who live shoulder-to-shoulder with the person, may be the relevant reference group for some situations. In this study, I analyse the role of relative residential positions, in relation to neighbours, in determining the satisfaction of residents. Reference groups are considered as neighbours living over different geographic scales, neighbours who are the symbols of success over different geographic scales, and neighbours who may be reached within a 15-minute walk. The results indicate that the most relevant reference group for comparisons amongst individuals' household crowding levels consists of people who may be reached within a 15-minute walk. By taking into account the endogeneity of the location choice, the results indicate that locating in a denser neighbourhood decreases residential satisfaction. People like to locate in relatively denser areas but they do not care about their relative household crowding.

8.1 Introduction

While one of the major goals of the urban development schemes is to increase residential satisfaction, there is little evidence regarding the effectiveness of such schemes in terms of delivering increased satisfaction. Residential satisfaction consists of a range of objective and subjective factors. People usually compare themselves to a reference group. In addition to any objective surrounding environmental characteristics, people's subjective evaluations often depend on their relative position within a reference group.

As concluded in Chapter 1, amongst three measures of crowding – perceived crowding (PC), Canadian National Occupation Standard (CNOS) and people per bedroom (PPBR) – the one comprising individuals' perceptions (PC) is the best predictor of their residential satisfaction. This claim considers the importance of subjective evaluations, which depends on the relative position of an individual. Nevertheless, both CNOS and PPBR perform almost as well.

It is generally accepted amongst policy makers that people have imaginary privacy circles around themselves. An individual may be annoyed when someone else stands too close to him or her. In the policy environment, the definition of 'too close' depends on an individual's background, in terms of where and how he or she was brought up. It is commonly believed that, for an estimation of the radius of their privacy circles, people look at the conventional distance between other similar people. If so, the desirability of a neighbourhood may be related both to the density of a neighbourhood and the composition of people in the neighbourhood.

Despite the long history of relative measurement in the economics literature (Duesenberry, 1949; Frank, 1985; Veblen, 1899), the reference group is often not clearly specified in theory. The word 'reference group' can be understood as those people who affect an individual the most. Neighbours, as a group who live in close proximity to the person, may comprise the relevant reference group for some issues (see, for example, Clark, Westergård-Nielsen, & Kristensen, 2009). Thus, in this study, I am interested in the impact of relative measures that measure the relative position of an individual, as well as the impact of the absolute measures, on residential satisfaction. I have based my relative indicators on objective measures of the neighbourhood by comparing the individual's crowding and population (area) density with the average crowding and the average area density of his/her neighbourhood.

In his seminal study on the economics of happiness, Easterlin (1974) emphasised the role of people's relative positions as a factor that negatively affects their happiness. He suggested that when people subjectively evaluate the position of another and proceed to judge themselves according to that subjective evaluation, they feel less satisfied than they would have done if they had been in a position to make an objective assessment. Easterlin presented the following statement from Karl Marx as a motivation for his study:

> 'A house may be large or small; as long as the surrounding houses are equally small it satisfies all social demands for a dwelling. But if a palace rises beside the little house, the little house shrinks into a hut.' (As quoted in Lipset (1960))

Following Easterlin's study, the recent wave of well-being studies has emphasised the role of relative positions. Relative income is often considered as the variable of interest. The idea is that people feel satisfied/dissatisfied as a result of comparing their income with the incomes of others, who are known as the reference group. A large body of the literature confirms the existence of a negative correlation between satisfaction and the relative income level. Therefore, in a neighbourhood, a person's utility falls once her neighbour's income increases.

To extend these studies, it is necessary to have a better understanding of their two main components, which are the reference group and the variable of interest. As Clark (2012) noted, people may compare themselves with others including their household members, people with similar characteristics (peer group), neighbours, colleagues or with themselves in the past. Comparisons with others are known as social comparison and comparisons with oneself in the past are known as adaptation. The possible impact of social comparisons on saving and consumption behaviours has been discussed by Duesenberry (1949). Knight and Gunatilaka (2010) claimed that a change in the reference group was the main reason for the lower levels of well-being amongst Chinese rural-urban migrants.

In a Chinese survey used by Kingdon and Knight (2007), respondents were asked about their reference group. The majority (68 per cent) stated that they compare themselves with other individuals in their own village, which is a sign of the importance of geographic proximities. Thus, in the present study, the validity of the neighbours' reference group at different geographic scales should be checked on, which largely depends on people's understanding of the boundaries of their surrounding neighbourhood. A comparison group may consist of people living 'over the road' or people living within the boundaries of an individual's larger neighbourhood. The other potential reference group consists of local people who are perceived as being more successful.

The second component of well-being studies to be considered is the variable of interest, which is often constructed based on a nominal variable, such as income. For instance, by taking a ranking approach, Clark, Westergård-Nielsen and Kristensen (2009) argued for the effect that relative income has on economic satisfaction in small neighbourhoods. They introduced neighbours as the 'right' reference group and discussed the impact of comparisons amongst neighbours. They also emphasised the importance of the availability of data at the local level.[72] Consequently, the richness of data plays a prominent role in this study. Since the output variable in the current study, residential satisfaction, depends mostly on the physical aspects, I used real variables in when constructing the relative measures. Luttmer (2005) studied the impact that higher income of neighbours has on individual's happiness by using individual-level panel data. His results suggest that, controlling for an individual's own income, a higher level of neighbours' income is associated with a lower level of self-reported happiness.

Pingle and Mitchell (2002) referred to substitution between work hours and leisure and assessed the role that 'keeping up with the Jones' leisure' plays in ones' 'keeping up with the Jones' income' behaviour. They concluded that the effort that a person puts into keeping up with others' income may be partly due to the effort that he or she puts in keeping up with others' leisure; that is, follower behaviour. Solnick and Hemenway (2005) used two surveys with a total of 226 respondents to study the differences in positional concerns amongst different domains. They concluded that consumption goods, such as house size, are more positional than safety and health. Their results suggest that the goods provided in an individual's community (or country), such as national parks, are more positional than goods provided for an individual's personal use, such as the number of rooms in a home.

[72] That study used the ECHP survey data, which is a Danish sample of European Community Household Panel with eight waves from 1994–2001.

As concluded in Section 6.8, PC has the highest marginal effect amongst predictors of residential satisfaction, which confirms the importance of the crowding factor in people's evaluations of the condition of their residential environment. The marginal effects of raw crowding measures are less than that of the subjective measure. This difference may derive from the relative positions of individuals, which are not captured by a raw crowding measure. Therefore, a relative crowding measure constructed based on PPBR, which is an objective measure of crowding, may account for comparisons amongst people. The construction of the relative factors will be discussed in Section 8.3.2.

The opportunities to access green spaces are not equal amongst different neighbourhoods. People located in business centres often benefit less from public green spaces. Anderson and West (2006) studied the impact that proximity to open space has on sale prices and concluded that proximity to open space is valued more highly in denser areas located nearby the CBD. Barbosa et al. (2007) studied the distance of different social sectors, including symbols of success, happy families and blue collars, from public green spaces. They concluded that the elderly and most deprived social groups enjoy the highest access to public green spaces. Therefore, even though people may not live in a crowded dwelling, they may think of a greener neighbourhood in their evaluations of the condition of their living environment.

For a review of the differences between crowding and density, see Section 2.3.1. The importance of a clear definition of neighbourhood boundaries in identifying the effects of neighbourhoods on the residents is discussed in Section 2.4, which also included a discussion on the average approach that will be taken when constructing relative measures. For a review of the studies on the impact of individual and household socio-demographic characteristics on residential satisfaction, see Section 2.4.3. The present chapter has used the same variables as the previous chapter to control for individual's predisposition (see Section 2.4.4). For a review of the definition of residential satisfaction and measurement errors in identifying the factors of residential satisfaction, see Section 2.2.

The relative measures are calculated based on a comparison with neighbours at different geographic boundaries, a comparison with successful people at different geographic boundaries and a comparison with people who may be visited within a 15-minute walk around the living area.[73] The following section presents the hypotheses of the current chapter. The construction of relative variables constructed for the purpose of this study will be explained in Section 8.3.2 and described in Section 8.4. The research design is described in Section 8.5. Section 8.6 presents the estimated results. The last section, 8.7, has our concluding remarks, considering the best-fit model.

8.2 Hypothesis development

In the previous chapter, I concluded that a household crowding higher than the level that an individual expects will lower that person's residential satisfaction. In addition to a comparison with norms, individuals may compare their residential position with others in their neighbourhood. This relative position, like the

[73] We will check on different walking distances for a range of 5–20 minutes.

absolute one, is likely to be endogenised by the person; that is, an individual, based on his or her background, chooses where to live and, therefore, influences what crowding level to have. Therefore, I hypothesise that:

H1: An increase in the relative household crowding on top of an individual's expected crowding level, based on her norms, lowers her residential satisfaction.

As concluded in the last chapter, an increase in the exogenous (unexpected) area density lowers residential satisfaction. However, when an area's density level is higher than its surrounding areas, the individual may be better off by using the agglomeration benefits provided in a relatively dense area. In other words, relative density may be a proxy for amenity availability concentrated in most dense places.[74] Therefore, I hypothesise that:

H2: An increase in the relative density level of an individual's neighbourhood will raise that person's residential satisfaction.

Summary

Despite the long history of the relative economic positions in the economics literature, the role of the relative position of an individual in terms of her housing situation has not been assessed before. This chapter asks two main questions: Does living in an area denser than others improve residential satisfaction? Does an increase in the relative household crowding on top of an individual's expected crowding level, based on his or her norms, lower his or her residential satisfaction?

To answer these questions, I need to recognise the group that people compare themselves with in their evaluations of their relative housing situation. Among the groups I consider in this study are: neighbours living over different geographic scales, neighbours who are the symbols of success over different geographic scales, and neighbours who may be visited in a 15-minute walk.

8.3 Data and sample

Like other chapters, this chapter uses three series of Statistics New Zealand's General Social Survey (NZGSS), (April 2008, October 2010, and April 2012). This dataset is merged with the Census 2006. For more information on data, see Section 7.3.

In order to take into account the geographic proximity of mesh-blocks, four geospatial datasets are used, namely, mesh-blocks' dataset, address points' dataset, roads' and tracks' datasets. After adopting the

[74] In other words, a dense area may provide more amenities or it may be the case that areas with more amenities tends to be relatively dense.

appropriate projection system and editing the datasets, by using a network analysis toolbox, a matrix of spatial proximity is derived (for more details see Section 8.3.2).

8.3.1 The geographic scales

In order to have a reliable definition of the neighbourhood I should rely on the residents' perceptions of their neighbourhood scales that can be derived from face-to-face interviews (Witten et al., 2003). A typical definition of neighbourhood consists of a certain number of nearest neighbours. This definition is very similar to the mesh-block's definition as an area consisting of dwellings that are located close to each other. This is defined by SNZ as the very small neighbourhood, which consists of a few houses located next to each other, with an average population of 110 residents. As I have assumed that the reference group consists of people who live shoulder-to-shoulder with an individual; the MB definition may be an approximate basis for defining the reference group.

In the context of New Zealand, larger neighbourhoods are defined as area units (AUs), akin to suburbs. Most people account for themselves as a member of their suburbs, which is not based on administrative boundaries. Hence, to take perceptions into account, area units may also be considered as a good definition of neighbourhoods. To take a step forward into the perceived definition of neighbourhoods, I consider the boundaries of territorial authorities (TLAs), which emphasise the importance of the political/administrative community of interest.

For a review of the geographic scales, see Section 7.4.2. In this chapter, MB, AU and TLA are depicted by M, A and T, respectively.

8.3.2 Relative measurement

A key distinction of this study is to consider the relative measures, as well as the raw measures, by considering the ratio (and variation) of the absolute measure of an individual, absolute crowding and absolute area density, to (from) the average value of that measure in her neighbourhood. To the best of my knowledge, no previous study has attempted to generate spatially referenced measures of relativity in order to interpret satisfaction responses.

Because of the different calculation approaches used and the use of different neighbourhood boundaries, the naming of the relative variables is complicated. A range of relative variables need to be constructed for our relative measures. A comprehensive description of the methods used for the construction of the relative measures of this chapter is presented in Appendix i. To presage the results, CVARS and ATVARDENS are the variables found to be the best variables in terms of prediction. CVARS is equal to the crowding level of an individual ($\mathrm{C_i}$) minus the average crowding in the areas that he or she may visit in a 15-minute walk around his or her MB ($\mathrm{CSM_i}$). This is stated as a ratio of the standard deviation of the CSM measure ($\sigma(\mathrm{CSM})$):

$$\mathrm{CVARS_i} = \frac{\mathrm{C_i} - \mathrm{CSM_i}}{\sigma(\mathrm{CSM})} \tag{27}$$

ATVARDENS is the population density of an individual's AU ($\mathrm{DENSA_i}$) minus the density of the TLA that the individual is located in ($\mathrm{DENST_i}$). This is stated as a ratio of the standard deviation of the average density amongst area-units ($\sigma(\mathrm{DENST})$):

$$\mathrm{ATVARDENS_i} = \frac{\mathrm{DENSA_i} - \mathrm{DENST_i}}{\sigma(\mathrm{DENST})} \tag{28}$$

The glossary included in the end of this thesis briefly describes the measures constructed in this study.

8.4 Descriptive statistics

Table 29 contains the descriptive statistics of the relative variables of interest that will be used in this chapter's analyses. For the description of other relative variables, see Appendix i.3.

As shown in Table 29, the CVARS variable has a mean of -0.283. As depicted in the P_value column, the mean difference amongst years is not statistically significant, at better than the 1 per cent level. The minimum and maximum values are reported in the 'Min' and 'Max' columns, respectively.

Table 29. Descriptive statistics of the constructed relative measures.

Relative measures	Mean	P_value[75]	Min	Max	Std error
ATVARDENS	10428.4	<0.001	-2195.97	114966.1	200.447
CVARS	-0.283	<0.001	-13.828	7.721	0.022

Note: Reported statistics account for jackknife resampling weights.

For a review of the relationship between PPBR versus household size see Table 22 and for the relationship between PBBR versus PC see Table 27.

8.5 Research design

The methodology employed in this chapter is the same as that in Chapter 7.

The variables of interest include both an endogenous effect and an exogenous effect. The exogenous effect can be thought of as a measure of undesirable crowding level and also undesirable density level. Assuming that I have captured the endogeneous effect precisely, the exogenous (or unexpected) effect is equal to the impact of the variable of interest that has not captured by the endogenous effect. The exogenous effect will be estimated based on the following equation:

$$\mathrm{y_i} = \gamma_0 + \gamma_1.\hat{\mathrm{x}}_\mathrm{i} + \gamma_2.\hat{\mathrm{x}}_\mathrm{i}^2 + \gamma_3.(\mathrm{x_i} - \hat{\mathrm{x}}_\mathrm{i}) + \gamma_4.(\mathrm{x_i^2} - \hat{\mathrm{x}}_\mathrm{i}^2) + \gamma_5.\hat{\mathrm{r}}_\mathrm{i} + \gamma_6.(\mathrm{r_i} - \hat{\mathrm{r}}_\mathrm{i}) + \gamma_7.(\text{Control variables})_\mathrm{i} + \varepsilon_\mathrm{i} \tag{29}$$

[75] The mean difference amongst years is not statistically significant, at better than the 1 per cent level.

where the dependent variable (y_i) is individual i's residential satisfaction. x_i includes the measures of households' absolute position, including household's crowding and area density. x_i^2 is the squared term of the absolute measures and r_i measures the relatives, including relative crowding and relative density. ε_i measures the residuals. γ_1 to γ_5 are the parameters of interest. $(r_i - \hat{r}_i)$ is a measure of the level of the relative crowding level of an individual that is beyond (or below) his or her expected crowding level. γ_5 measures of the impact of the exogenous relative term[76] (for further discussion on the complexity of the model specifications, see Appendix ii). As in the previous chapter, the dependent variable (y_i) follows a Bernoulli distribution. As discussed in the previous chapters, a logit transform function is used.

As in the previous chapter, this chapter uses the resampling weights (for a review of resampling by using replicate weights, see Section 3.1.3).

8.6 Models and results

As discussed in Section 8.3.2, there are 14 relative crowding measures and six relative area density measures.[77] To find the model with the highest predictive power for our dependent variable, residential satisfaction (RS), I should account for all combinations of the relative crowding measures and the relative density measures, which derives 84 equations.[78]

By using different relative crowding and also different measures, the best measures are chosen based on the GOF measures. As noted in Section 8.3.2, the results indicate that amongst relative crowding measures, the equations with CVARS has the highest predictive power for RS.[79] Both of these measures are constructed using the variation method. Amongst the relative density measures, ATVARDENS has the highest predictive power. Therefore, the estimations containing combinations of CVARS and

[76] To check on the sensitivity of results to the average crowding of areas – that is, to check on any plausible bias in the estimations caused by heterogeneity amongst neighbourhoods, on top of all estimations based on equations (38) and (30) – a synthetic panel approach is used. The Mundlak correction is used to take into account the intrinsic features of the area (Mundlak, 1978). The results show that the impact of the variables of interest are not affected by the heterogeneity issue.

[77] The 14 relative crowding measures are MC, AC, TC, MVARC, AVARC, TVARC, MHC, AHC, THC, MHVARC, AHVARC, THVARC, MCS, CVARS. The six relative crowding measures are MADENS, MTDENS, ATDENS, MAVARDENS, MTVARDENS, ATVARDENS. In addition to these, the comparison with symbols of success is done at the 95th income percentile as well.

[78] I tried both the 15- and five-minute walking distance weighting matrices and looked for the change in the coefficients in presence of year interactions. However, these provided no extra information and are therefore not presented here.

The same approach is taken in regard to the relative density measure; that is, if the density of the geographic scale in which the individual lives is compared with the density over a larger geographic scale, the smallest geographic scale's absolute density level (DENSM, DENSA or DENST) and its squared form ($DENSM^2$, $DENSA^2$ or $DENST^2$) are included in the equations. For example, in an equation that contains ATDENS, which is the density level of individuals' area units as a proportion of the density level at the TLA that the area unit is located in, the density of AU and its squared form are included.

[79] The second-best relative crowding measure is AHVARC95.

ATVARDENS will be presented in this section. Appendix iii presents the table of the GOF (AIC), by including different combinations of relative measures.

Table 30 summarises the results derived from the estimation of equation (30) (for results including the control variables, see Appendix iv). [80] The joint significance of the related terms, such as $\widehat{\text{PPBR}}$ and its squared term, are reported in the bottom of the table. Results indicate a positive impact for the relative density measure; that is, an unexpected relative density raises RS. As in Chapter 7, higher levels of unexpected crowding and density have a significant negative impact on RS. As presented in the last column of the results table, the significant effects of unexpected crowding and density on RS persist with the presence of PC in the models.

[80] The results presented here will be discussed more thoroughly in Chapter 9's models and results section (9.7).

Table 30. Results of estimation of the impact of relative crowding and density measures.

Model number	1	2
	Equation (30)	Equation (30) + PC
Residential Satisfaction (dummy)		
$\widehat{\text{PPBR}}$	1.485‡	1.487‡
	(2.655)	(3.284)
$\widehat{\text{PPBR}}^2$	1.136‡	0.953‡
	(0.789)	(7.006)
$\text{PPBR} - \widehat{\text{PPBR}}$	0.151**†	0.210**†
	(0.111)	(0.155)
$\text{PPBR}^2 - \widehat{\text{PPBR}}^2$	1.352†	1.343†
	(0.490)	(0.500)
$\widehat{\text{DENSA}}$	0.951‡‡	0.785‡‡
	(1.959)	(1.678)
$\widehat{\text{DENSA}}^2$	0.821‡‡	0.840‡‡
	(0.242)	(0.258)
$\text{DENSA} - \widehat{\text{DENSA}}$	0.950††	0.965††
	(0.183)	(0.193)
$\text{DENSA}^2 - \widehat{\text{DENSA}}^2$	0.936††	0.931††
	(0.050)	(0.051)
$\widehat{\text{CVARS}}$	2.884¡	2.531¡
	(3.332)	(3.017)
$\text{CVAR} - \widehat{\text{CVARS}}$	1.160¡	1.164¡
	(0.153)	(0.152)
$\text{ATV}\widehat{\text{ARD}}\text{ENS}$	1.00004**¡¡	1.00002**¡¡
	(0.000)	(0.000)
$\text{ATVARDENS} - \text{ATV}\widehat{\text{ARD}}\text{ENS}$	1.00004**¡¡	1.00004**¡¡
	(0.000)	(0.000)
PC		0.211* *
		(0.050)
Observations	4607	4607
Population size	904053	904053
R^2	0.192	0.196
AIC	5033.65	5025.655
Joint significance (Wald test):		
Expected crowding measures‡	Insig.	Insig.
Expected density measures‡‡	Insig.	Insig.
Unexpected Crowding measures†	<0.01	<0.001
Unexpected Density measures††	<0.01	<0.01
Relative crowding measures¡	Insig.	Insig.
Relative density measures¡¡	<0.05	<0.05

Reported coefficients are exponentiated ones. Standard errors are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

‡The significance levels for $\widehat{\text{PPBR}}$ and its squared form (Squared $\widehat{\text{PPBR}}$) are reported jointly on the expected crowding measures Wald test row. ‡‡ The significance levels for $\widehat{\text{DENSA}}$ and its squared term are reported jointly on the expected density measures Wald test row. † The significance levels for $\text{PPBR} - \widehat{\text{PPBR}}$ and its squared term are reported jointly on the unexpected crowding

measures Wald test row. †† The significance levels for $\mathrm{DENSA} - \widehat{\mathrm{DENSA}}$ and its squared term are reported jointly on the unexpected density measures Wald test row. ⁱ The significance levels for $\widehat{\mathrm{CVARS}}$ and $\mathrm{CVAR} - \widehat{\mathrm{CVARS}}$ variables are reported jointly on the relative crowding measures Wald test row. ⁱⁱ The significance levels for $\mathrm{ATV}\widehat{\mathrm{ARD}}\mathrm{ENS}$ and $\mathrm{ATVARDENS} - \mathrm{ATV}\widehat{\mathrm{ARD}}\mathrm{ENS}$ variables are reported jointly on the relative density measures Wald test row.

Note that the joint significance tests are reported for different coefficients in different models. The reported AICs are for the models that do not take jackknife replications into account.

8.7 Conclusion

Relative measures are constructed at different geographic scales, namely mesh-blocks, Area units and Territorial authorities. In the construction of the relative measures, three reference groups were defined: neighbours at different geographic scales, neighbours who are the symbols of success at different geographic scales and neighbours who live within a 15-minute walking distance.

Amongst the relative crowding measures, CVARS – which is a variation measure constructed based on the difference between an individual's crowding level and the crowding level of the areas that he or she may visit within 15-minutes walking distance around his or her neighbourhood – provides the highest predictive power for residential satisfaction. This measure has been used in this study to study the impact of relative crowding on residential satisfaction.

An expected crowding level higher than others, which a person may see within a 15-minute walk around his or her neighbourhood, does not affect that person's satisfaction. This suggests that people do not really choose their location based on a comparison of their household crowding level with others' crowding level.

In a comparison between the results of this chapter with those of the previous chapter, an insignificant effect of expected crowding level in the presence of the relative terms highlights the importance of social comparisons in people's location choice. In other words, when people choose their location, they compare the characteristics of their living environment with others rather than thinking of the characteristics of the location per se. This result is compatible with those of previous studies on the impact that relative income has on happiness, namely that people care more about their relative position than the absolute position (Easterlin, 1974; Easterlin, 1995; Grimes & Reinhardt, 2015).

However, in line with the results from the last chapter, locating in a household with a higher-than-expected crowding level – that is, after taking the location choice into account – lowers residential satisfaction remarkably.

Amongst the measures of relative crowding, ATVARDENS, which compares area units' area density with the density level at the territorial authorities by using a variation method, is the best predictor of residential satisfaction. Using this measure, I infer that if people locate in a relatively dense area they will be happier with their neighbourhood.

Based on Table 31, the impact of the measures of relative density measures are as follows: a higher level of unexpected area density lowers RS; a higher level of relative density raises RS; a higher level of expected relative density raises RS; a higher level of unexpected relative density raises RS. Thus, the relative measures show that greater density than one's neighbouring areas raises RS.[81] The relative measures are zero-sum games – that is, one person wins and a neighbour loses from this measure, whereas, according to the last chapter's results, the absolute measures accrue to all. Therefore, higher density overall reduces RS but, keeping the overall density constant, living in a relatively dense area benefits residents since that is where the greatest amenities of the city are likely to be.

Table 31. A summary of the results of chapter 8.

Category	Result	Significance
Household crowding derived from backgrounds ($\widehat{\mathrm{PPBR}}$)	A higher household crowding stemming from 'own group' does not affect RS	No
Exogenous household crowding ($\mathrm{PPBR} - \widehat{\mathrm{PPBR}}$)	A higher unexpected level of crowding lowers RS	Yes
Absolute area density derived from backgrounds ($\widehat{\mathrm{DENSA}}$)	A higher area density level stemming from 'own group' attraction does not affect RS	No
Exogenous area density ($\mathrm{DENSA} - \widehat{\mathrm{DENSA}}$)	A higher level of unexpected area density lowers RS	Yes
Expected relative crowding ($\widehat{\mathrm{CVARS}}$)	A higher level of expected relative crowding does not affect RS	No
Unexpected relative crowding ($\mathrm{CVAR} - \widehat{\mathrm{CVARS}}$)	A higher level of unexpected relative crowding does not affect RS	No
Expected relative density ($\mathrm{ATV}\widehat{\mathrm{ARD}}\mathrm{ENS}$)	A higher level of expected relative density raises RS	Yes
Unexpected relative density ($\mathrm{ATVARDENS} - \mathrm{ATV}\widehat{\mathrm{ARD}}\mathrm{ENS}$)	A higher level of unexpected relative density raises RS	Yes
Perceived crowding (PC)	A higher PC level affects RS negatively	Yes

Density is a good provided in an individual's community, while crowding is a good directly related to a specific individual. Hence, the results in Chapters 7 and 8 suggest that the goods provided in the community are more positional than the goods provided for an individual's use. This is similar to the results of Solnick

[81] According to the results in Chapter 7 and in the results table presented in Appendix IV, a higher (non-relative) density reduces RS.

and Hemenway (2005), but contrary to Galbraith and Galbraith's (1984) hypothesis that the goods provided for a community are less positional than the goods for personal use.

Luttmer (2005) showed that an increase in neighbours' income lowers the happiness of an individual. This result is not similar to the finding from the present study that locating in a relatively denser area raises residential satisfaction and that people do not compare their household crowding with others. A reason for the differences between our results is the difference between our variables of interest and our outcome variables. As discussed, relative crowding is more of a private good (than a good provided in the community), and is therefore more comparable with relative income, which is Luttmer's variable of interest. From a conceptual point of view, relative density represents an individuals' benefit from locating nearby amenities and facilities provided in her area and is therefore comparable with the relative income variable that Luttmer is interested in. The difference between my results and Luttmer's suggest that people attain satisfaction from neighbouring affluent others as long as it benefits them.

Sloan and Morrison (2015) studied the post-move satisfaction of people moving within New Zealand and concluded that moving to larger urban areas is associated with lower housing satisfaction. Their result is compatible with the findings from Chapters 7 and 8 of this thesis on the negative impact of living in an area with a higher density level on residents' satisfaction with their living environment.

Appendix I. Construction of the relative measures

Section 8.3.2 presented the two relative measures that best predict RS. This appendix explains the construction of relative measures in details. A range of relative variables need to be constructed for our relative measures. A glossary included in the end of this thesis briefly describes the measures constructed in this study.

People's subjective evaluations, such as those derived from the perceived crowding measure, depend on a range of factors apart from their relative position. Thus, in order to account for the relative position of people it is necessary to have a measure that can be gauged objectively and expressed relatively. As discussed in the literature review section, the PPBR measure is our only crowding variable with such a specification.

Density is abbreviated to DENS and crowding is abbreviated to C. The names of the variables are chosen such that N needs to be substituted for the definition of neighbourhood boundaries: M (mesh-block), A (area unit), T (territorial authority) or S (spatial – walking distance). If the name of variable begins with the geographic scale's letter, it means that it is a relative measure. For example, the variable named 'AC' is the crowding level of an individual compared to the average crowding level in his or her area unit (A). If the name ends with a letter, this means that the measure is an absolute measure. For example, DENSA is the absolute area density at the area unit (A) geographic scale.[82]

I first explain the construction of relative crowding levels by using the ratio method.[83] As depicted in equation (31), the relative crowding ratio (NC_i), which is an ix1 column vector, is defined as the ratio of the

---

[82] The ratio of the crowding of the observation i to the average of its mesh-block (j) will be called MC_{ij}, where the subscript shows the element of the MC matrix that we are referring to. Since the number of people living in a MB is small, the relative measure at this level may be very sensitive to small changes in the surrounding environment.

$$MC_{ij} = \frac{\text{PPBR of the observation i}}{\text{Average PPBR of observation i's mesh block}} = \frac{C_i}{CM_j}$$

For example, assume that individual 1 lives in a dwelling with a crowding equal to 1 PPBR, which is located in neighbourhood A with an average crowding of 0.5ppr. Individual 2 with a crowding equal to 2 PPBR lives in neighbourhood B with an average crowding of 2 PPBR. The (relative) crowding ratio for individual 1 is then equal to 2 and for individual 2 it is equal to 1. At a bigger scale, we compare the individual with the average of its area unit, which derives AC_{ij}; or to the average of its territorial authority TC_{ik}.

[83] For constructing the measures, we first need to know whether an individual lives in neighbourhood N or not. Therefore, for each individual (i) I have j dummies that show whether individual i lives in neighbourhood j or not. This is shown in the following equation, where the final matrix (D) consists of D_{ij}, which contains the dummy values for i individuals and j neighbourhoods.

$$D = [D_{ij}], \qquad D_{ij} = 1 \text{ (if individual i lives in neighbourhood j); 0 (otherwise)}$$

$$i = 1, \ldots N; j = 1 \ldots, J$$

crowding level of their household (C_i)[84] to the average crowding level of their neighbourhoods' defined at geographic scale N (CN_i).

$$NC = [NC_i], NC_i = \frac{C_i}{CN_i} \tag{30}$$

I should account for the relative density of surrounding areas in the same way as I did for the relative crowding of individuals. The same approach as the relative crowding measure construction will be taken for deriving the relative density measure. However, as density is an area-level variable (and not an individual–level one), individuals' values are equal to the values assigned to their neighbourhoods. Therefore, for naming the relative density measures, two indicators of geographical scales are required.[85] For example, if I assume that people evaluating their living environment compare the density level of the area unit (A) that they are located in with the density level at their territorial authority (T), the constructed relative density variable will be called ATDENS.

The relative density level (NRDENS), by employing the ratio method, is the ratio of the density of individual i's neighbourhood ($DENSN_i$), defined at the geographic scale N, to the density of his or her reference group's neighbourhood ($DENSR_i$):

$$NRDENS = [NRDENS_i], NRDENS_i = \frac{DENSN_i}{DENSR_i} \tag{31}$$

For example, MTDENS is the ratio of the density at mesh-blocks divided by the density at TLAs; that is, MTDENS assumes that people compare the density of their mesh-blocks with the density of the territorial authority in which they are located. Similarly, the area density of mesh-block j ($DENSM_j$) can be expressed as a ratio of the area density of the area unit k ($DENSA_k$), which derives $MADENS_{jk}$. Similarly, MTDENS

In order to determine the value of crowding in individual i's neighbourhood, the matrix of average crowding of neighbourhoods (CN_j) for j neighourhoods is multiplied by the matrix containing the dummy values for individual i (D_{ij}).

$$CN = [CN_i], \qquad CN_i = \sum_{j=1}^{J} D_{ij}.CN_j$$

[84] C_i is equal to individual i's PPBR.

[85] The density level for individual i ($DENS_i$) at her neighbourhood level (N), as depicted in the following equation, is a multiplication of the neighbourhoods' absolute density level ($DENSN_j$) by D.

$$DENSN = [DENSN_i], \qquad DENSN_i = \sum_{j=1}^{J} D_{ij}.DENSN_j$$

In the same way, the density level of the area that individual i compares himself or herself with (DENSR) derives from the multiplication of the absolute density at the relative area's geographic scale by the dummy variable.

and ATDENS are the ratios of the area density of MBs to TLAs', and AUs' density level to TLAs', respectively.

Besides the ratio approach, another way of measuring relativities is to state the relative measures in terms of a distance from the average of the neighbourhood. This variation can be stated as a ratio of the standard deviation (σ) of the average crowding amongst neighbourhoods to account for the relative magnitude of the variation. 'VAR' is added to the name of the variables in order to distinguish between the variables constructed by using the variation method and the measures constructed based on the ratio method. For example, the relative crowding based on a comparison with the average crowding in mesh-blocks constructed by using the variation method is equal to the crowding level of an individual (C_i) minus the average crowding in his or her MB (CM_i) divided by the standard deviation of the average crowding of MBs ($\sigma(CM)$):

$$\mathrm{MVARC} = [\mathrm{MVARC_i}], \mathrm{MVARC_i} = \frac{\mathrm{C_i} - \mathrm{CM_i}}{\sigma(\mathrm{CM})} \quad (32)$$

Also, the relative density measure constructed based on a comparison between the density level at an individual's MB and the density level at the TLA in which he or she is located by using the variation method (MTVARDENS) is equal to:

$$\mathrm{MTVARDENS} = [\mathrm{MTVARDENS_i}], \mathrm{MTVARDENS_i} = \frac{\mathrm{DENSM_i} - \mathrm{DENST_i}}{\sigma(\mathrm{DENST})} \quad (33)$$

In this study, I consider both the ratio and the relative variation measures. Alternatively, I can compare the crowding ratios of observations by taking a decile approach and defining the reference group as the 'symbol of success' group. There are a number of criteria that characterise successful people, such as their wealth and reputation. The most common feature is the level of income. Therefore, I compared the individual's crowding with the crowding of people with an income at the 90^{th} and 95^{th} percentiles.

In order to compare the individuals with the symbols of success group as the reference group, I construct the AHC measure, which compares an individual's PPBR with the average crowding of individuals with an income at the 90^{th} percentile in his or her AU. Similarly, THC is the ratio of the individuals' crowding to the average crowding of their TLA's 90^{th} percentile of income. Amongst the constructed relative measures using the symbols of success as their reference group, AHVARC95 has the highest predictive power for RS. AHVARC95 is the crowding level of an individual (C_i) minus the average crowding level of his or her AU's 95^{th} income percentile ($HA95_i$) divided by the standard deviation of HA ($\sigma(HA95)$):

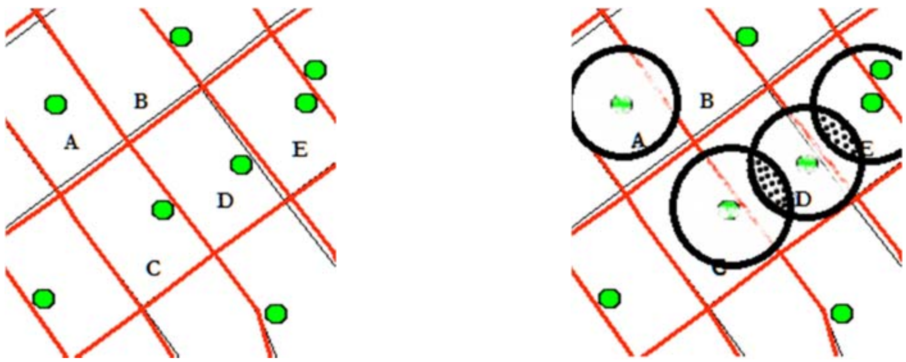
$$\mathrm{AHVARC95} = [\mathrm{AHVARC95_i}], \mathrm{AHVARC95_i} = \frac{\mathrm{C_i} - \mathrm{HA95_i}}{\sigma(\mathrm{HA95})} \quad (34)$$

I may also consider a walking distance definition for the boundaries of the neighbourhoods in my definition of the referenced neighbourhoods' geographic scales. This will be discussed further in Appendix i.1. A summary of the constructed relative measures will be presented in Appendix i.2. The descriptive statistics of the constructed relative measures will be depicted in Appendix i.3.

Appendix I.1 The walking distance approach[86]

A complementary approach to relative measurement is to use Network Spatial analysis to account for walking distance from mesh-blocks' centroids (for a review of spatial weight matrix construction, see Appendix i.1). This method derives polygons of, for example, 15 minutes walking distance around the centroids of areas. Figure 18 depicts areas A, B, C, D and E located next to each other. Lines depict the boundaries and the solid circles are the centroids of each area. On the right-hand side of the figure, the circles around the centroids depict a 15-minute walking distance from the centre of each area.

Figure 18. An example of the walking distance method.



Source: Torshizian (2017).

This derives some areas intersecting one another. For example, area D's polygon has an overlap with areas C's and E's polygons. The overlap of an area's polygon layer divided by its total intersecting area with other areas can be considered as the areas' standardised weights.

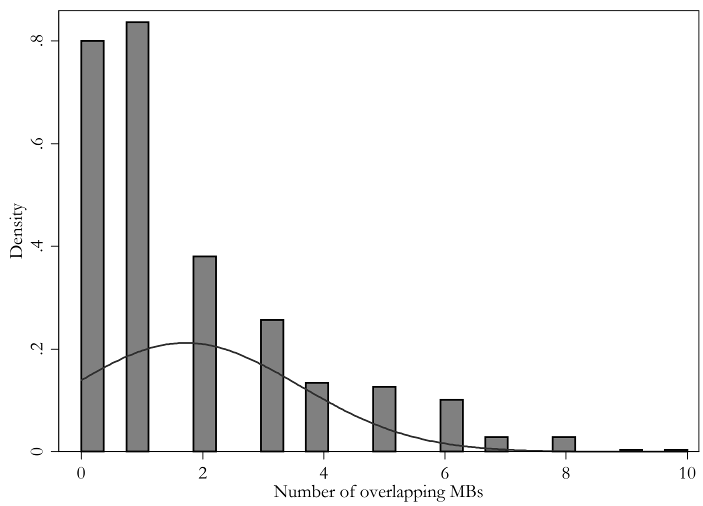
The overlap of a mesh-block's polygon layer divided by its total intersecting area with other mesh-blocks plus the MB's average crowding/area density value is considered as the reference group's absolute crowding/area density value. This will be used in the construction of the ratio and variation measures. This method takes into account the importance of walking distances in social comparisons. Since I am not aware of the exact location of individuals in the mesh-blocks, the centroid of dwellings' doors is considered as a mesh-block's centre, which helps to decrease the probability of the improper attribution of the relative measures to the people who live closer to the edges of MBs.

The final weight matrix in our sample is a 2992 x 2992 matrix; that is, it contains the weights for 2992 MBs. This matrix is a sparse one, which means that it contains a large proportion of zeros. A total of 1182 MBs are not affected by any other MB in terms of walking distance. As illustrated in Figure 19, 41 percent of the

[86] This approach will be discussed in details in chapter 10.

remaining 1810 MBs have only one neighbouring MB, 45 per cent have 2–4 adjacent neighbouring MBs and 14 per cent have between 5–18 neighbouring MBs. The maximum number of adjacent MBs is 18 with an average of 2.5 adjacent areas.

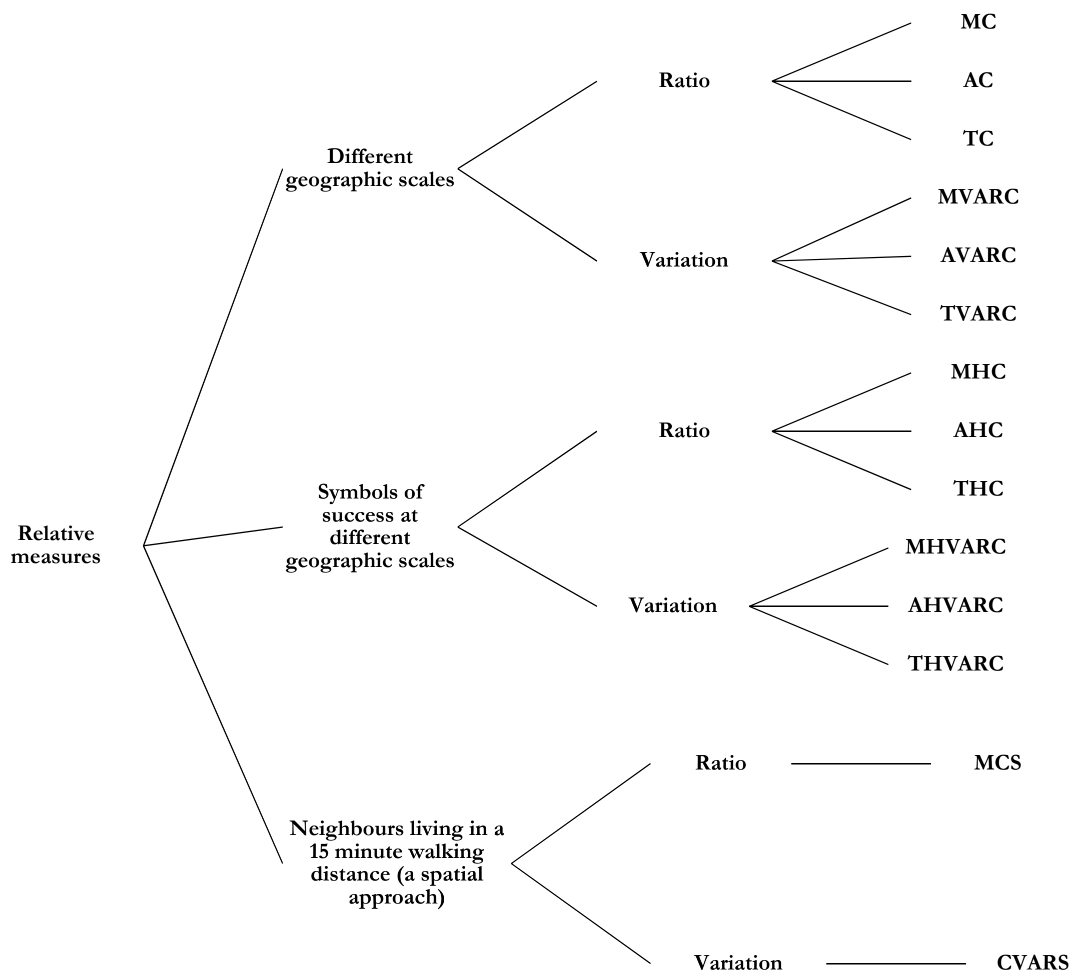
Figure 19. The spatial weights derived from the walking distance approach.



I have assumed that an individual's own MB affect him or her the most. Therefore, the diagonal of the spatial weight matrix is equal to one. The multiplication of this matrix by the crowding of MBs is the reference group's absolute crowding value, named CMS, which is used in the construction of a relative measure as a ratio (named MCS) and a variation (called CVARS).

Appendix I.2 Quick review of the relative crowding measures

In the diagram below, the second nodes from the left-hand side represent the reference group considered in the construction of the relative crowding measure. The third nodes from the left-hand side illustrate the method used in the construction of the measures. The final nodes on the right-hand side contain the name of the measure at three levels: mesh-block, if the name starts with 'M'; Area unit, if the name starts with 'A'; and Territorial Authority, if the name starts with 'T'. For a description of the variation measures, see equations (33) and (34). For a description of relative measures, see equations (31) and (32).


Relative measures
Different geographic scales
Ratio
MC
AC
TC
Variation
MVARC
AVARC
TVARC
Symbols of success at different geographic scales
Ratio
MHC
AHC
THC
Variation
MHVARC
AHVARC
THVARC
Neighbours living in a 15 minute walking distance (a spatial approach)
Ratio
MCS
Variation
CVARS

Appendix I.3 Descriptive statistics of the relative measures

Following the naming strategy discussed in Section 8.3.2, CM, CA and CT are the absolute values of crowding at MBs, AUs and TLAs. MC is the crowding level of individuals as a ratio of their MB's average crowding level. The same approach is taken in the construction of AC and TC, which are at AU and TLA levels. MVARC is the crowding level of individuals compared with their MB constructed by using the variation method, as presented in equation (33) in Appendix i. MHC is the crowding level of individuals compared with the average crowding of the symbols of success living in the same MB, constructed based on the ratio method. The same measure constructed based on the variation method is called MHVARC. DENSM, DENSA and DENST are the absolute density values at MB, AU, TLA levels. MADENS is the ratio of DENSM to DENSA. MAVARDENS is the same measure as MADENS, but constructed by using the variation method. MTDENS is the ratio of the density level at MBs to the density level at TLAs. MTVARDENS is the same measure for a comparison between density levels at MBs and at TLAs by using a variation method. CMS, as discussed, is the absolute crowding level of the neighbourhoods defined by using the 15-minute walking distance weighting matrix.

Table 32. Geographic units and individual-level descriptive statistics.

Relative measures	Mean	P_value[87]	Min	Max	Std error
MC	0.9	<0.001	0.101	3.105	0.007
AC	0.888	<0.001	0.112	2.505	0.006
TC	0.876	<0.001	0.092	2.249	0.006
MVARC	-0.279	<0.001	-13.194	7.374	0.021
AVARC	-0.371	<0.001	-11.149	7.829	0.026
TVARC	-0.756	<0.001	-13.719	15.828	0.058
THVARC	-0.514	<0.001	-16.661	22.175	0.086
THVARC95	-0.514	<0.001	-16.661	22.175	0.086
AHVARC	-0.299	<0.001	-8.498	10.492	0.023
AHVARC95	-0.299	<0.001	-8.517	10.515	0.023
MHVARC	-0.189	<0.001	-9.604	9.604	0.019
ATVARDENS	10428.4	<0.001	-2195.97	114966.1	200.447
CVARS	-0.283	<0.001	-13.828	7.721	0.022

Note: Reported statistics account for jackknife resampling weights.

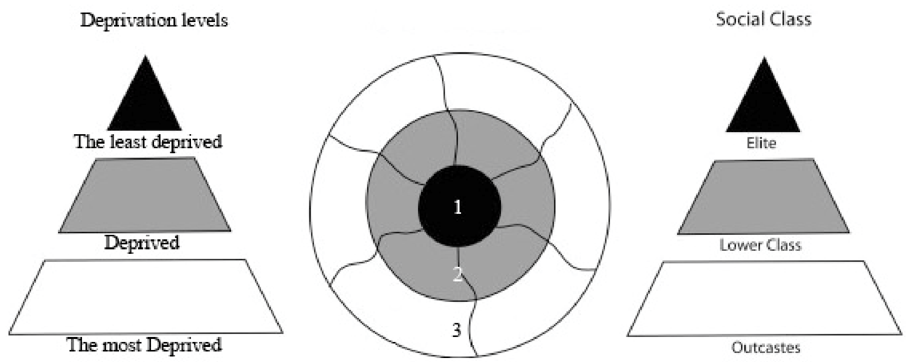
Appendix I.4 A review of spatial weights matrix

This appendix explains the method used in this study for the construction of spatial weight matrices. To fix ideas about the method of accounting for spatial proximities, assume a simple model of neighbourhood, which is compatible with Sjoberg's model of a pre-industrial city illustrated in Figure 20. In this figure, the circle in the middle shows three urban areas; the triangle on the right shows Sjoberg's classification; and the triangle on the left-hand side shows, for example, different density levels. To make it more comprehensible, I combined Sjoberg's explanations and made the following assumption about density levels. The elites who

[87] The mean difference amongst years is not statistically significant, at better than 1 per cent level.

are usually from high social levels, which probably have a lower density level than other social levels, and work with governmental, administrative or religious sectors, are located in the centre of the city, which might be a compact area as well. After elites, middle-class people, with a medium density level, live in the second layer. In the third layer, people with a lower social level and higher density level are located far away from the city centre. Although this exact situation is rare in reality, as an example I have assumed that each of these layers is an urban area with a specific density level (D_1, D_2 and D_3). Furthermore, each of these layers has many residents with a specific residential satisfaction level derived from many factors.

Figure 20. Assumed neighbourhood model inspired by Sjoberg's idealised model of a preindustrial city.



I have assumed a simple linear-in-parameters cross-sectional model with residential satisfaction as the dependent variable on the left-hand side of the equation and density levels plus some control variables as explanatory variables on the right-hand side of the equation. In order to estimate the relationship amongst urban areas, spatial effects should be assessed. In this case as I am just interested in the average-neighbour values of the independent variable. The Spatial Durbin Model (SDM) can be used as a basis as follows (Viton, 2010):

$$\mathbf{y_i = X'_i\beta + X'_i.(W \times X'_i)^{-1}.\theta + u_i, \quad u_i \sim N(0, \sigma^2 I), \quad i = neighbourhoods} \tag{35}$$

where n × 1 vector y is the dependent variable, X is an n × (k + 1) matrix and represents the explanatory variables, β is the vector of regression coefficients, the scalar θ is a spatial autoregressive parameter that reflects the spatial dependence inherent in the relationship, W is the spatial weights matrix and u is the error term. I included control variables on the right-hand side of this equation to control for the effect of other factors such as demographics. The SDM is a more accurate model than simple OLS in cases where there is a correlation between explanatory variables and omitted variables (Gerkman, 2011). Some results derived from the example are as follows.

Table 33. The neighbour relations based on Figure 1. [88]

Region	Neighbours
1	2
2	1,3
3	2

As Anselin (1988) noted, W is an exogenous matrix that is based on the theoretical assumptions of research, and it can derive from the distances between the observations or from contiguity. Therefore, based on the contiguity approach in Table 33, the row standardized weights matrix is as follows:

Equation 1

$$W = \begin{pmatrix} 0 & 1 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \end{pmatrix}, D = \begin{pmatrix} D_1 \\ D_2 \\ D_3 \end{pmatrix}$$

where D is the explanatory variable of interest (density levels). By multiplying the weight matrix and density levels (D), I derive a relative density level matrix as follows:

Equation 2

$$W \times X' = W \times \begin{pmatrix} D_1 \\ D_2 \\ D_3 \end{pmatrix} = \begin{pmatrix} D_2 \\ \frac{D_1}{2} + \frac{D_3}{2} \\ D_2 \end{pmatrix}$$

$$X'.(W \times X')^{-1} = \begin{pmatrix} D_1 \\ D_2 \\ D_3 \end{pmatrix} . \begin{pmatrix} D_2^{-1} \\ (\frac{D_1}{2} + \frac{D_3}{2})^{-1} \\ D_2^{-1} \end{pmatrix} = \begin{pmatrix} D_1 D_2^{-1} \\ D_2(\frac{D_1}{2} + \frac{D_3}{2})^{-1} \\ D_3 D_2^{-1} \end{pmatrix}$$

The resulting variable is clearly the density level of each area relative to its adjacent areas.

Appendix II. The complexity and necessity of a more advanced spatial econometric method

The modelling method of the current study is the simple Spatial explicit model (SLX), which accounts for the spatial distribution of the variable of interest.[89] The robustness of the results has been checked using

[88] Region 1 is the neighbour of district 2; region 2 is surrounded by districts 1 and 3; and region 3 is next to neighbourhood 2.

[89] The main model of this study, described in equation (38), can be written in the following form, which is an SLX model:

$$y_i = \alpha_0 + \alpha_1.x_i + \alpha_2.x_i^2 + \alpha_3.W.x_i + \alpha_4.W.x_i^2 + \alpha_5.(Control\ variables)_i + \varepsilon_i$$

Moran's I test, which is a measure of spatial autocorrelation. The results of the test show that at better than 5 per cent significance level I can reject the null hypothesis that residuals are independent and identically distributed. However, when there is little spatial dependence in the data, this test may only be considered as an appropriate estimator of the strength of the spatial dependence parameter. There are some alternatives available; for example, see Li, Calder, and Cressie (2007). For a credible analysis, however, a more complex spatial modelling approach, such as the Spatial Durbin Model (SDM) or the Spatial Durbin Error Model (SDEM), needs to be used and to be compared with our SLX model in terms of goodness of fit.

Few studies have considered spatial dependence in discrete choice models. One reason for this could be the complications associated with taking the full spatial information into account. For example, the problem with a maximum-likelihood estimation of a standard limited dependent variable model, such as binary probit model, is that, for a sample population of n, an n-dimensional integration is needed in order to account for off-diagonal terms of the variance-covariance matrix.

The estimation of discrete choice models in the presence of spatial dependence is discussed thoroughly by Fleming (2004). Table 34 summarises the techniques that can be used as an alternative for the SLX model in this study:

Table 34. Differences between estimators using different approaches.

Approach	Computational burden	Requires Calculation of n by n determinant	Provides spatial parameter standard error	solution for n-dimensional integration
Pinkse & Slade (SAE)	high	yes	yes	no
Non-Linear Least Squares (SAL) low	low	no	yes	not necessary
Non-Linear Least Squares (SAE)	moderate	no	no	not necessary
EM Algorithm (SAL)	higher	yes	no	yes
EM Algorithm (SAE)	higher	yes	no	yes
RIS Simulator (SAL)	highest	yes	yes	yes
RIS Simulator (SAE)	highest	yes	yes	yes
Gibbs Sampler (SAL)	higher	yes	yes	yes
Gibbs Sampler (SAE)	higher	yes	yes	yes

Source: (Fleming, 2004)

Table 34 compares the computational burden of different estimators. A comparison between these estimators shows that, considering the large sample population of this study, the use of other estimators

Where W is the spatial weight matrix that is explained in Chapter 10.

may be either not be feasible or not informative enough. Therefore, this thesis does not provide a more complex modelling approach.[90]

Appendix III. Goodness of fit of the relative measures

In order to identify the best relative crowding and relative density in terms of their predictive power for RS, I estimate the following equation:

$$y_i = \alpha_0 + \alpha_1 . x_i + \alpha_2 . x_i^2 + \alpha_3 . r_i + \beta . Controls_i + \varepsilon_i \qquad (36)$$

Where the dependent variable (y_i) is individual i's residential satisfaction. x_i includes the measures of households' absolute position, including household's crowding and area density. x_i^2 is the squared term of the absolute measures and r_i measures the relatives, including relative crowding and relative density. ε_i measures the residuals.

The combinations of the variables of interest, including both the relative and the absolute measures, are depicted in Appendix iii.1. Table 35 presents the predictive power (AIC) of the different combinations of relative crowding measures, depicted in the first column, and the relative density measures, shown in the first row. As shown in Table 30, the best predictive power derives from the model that includes ATVARDENS, as the relative density measure, and CVARS, as the relative crowding measure. This model has an AIC equal to 5033.65, which is smaller than the AIC values of all other models depicted in Table 35.

Table 35. Goodness of fit (AIC) of the models, including different combinations of relative measures.

	MADENS	MTDENS	ATDENS	MAVARDENS	MTVARDENS	ATVARDENS
MC	5080.156	5080.658	5066.605	5069.444	5079.746	5068.222
AC	5075.246	5077.835	5061.843	5066.653	5075.986	5060.677
TC	5082.171	5085.64	5070.944	5073.535	5085.305	5070.997
MVARC	5082.972	5083.243	5069.002	5073.056	5081.906	5069.885
AVARC	5077.917	5080.06	5064.672	5067.756	5076.937	5065.277
TVARC	5080.033	5082.657	5070.377	5071.906	5079.206	5068.74
MHC	5080.208	5082.618	5070.497	5071.686	5080.655	5068.526
AHC	5083.114	5082.554	5071.462	5073.678	5083.455	5072.202
THC	5081.01	5082.712	5070.667	5071.284	5081.761	5069.426
AHVARC95	5079.381	5080.893	5066.309	5071.126	5074.841	5062.403
CVARS	5050.157	5052.265	5035.592	5045.258	5053.891	5033.652

Note: Smaller AIC values are better when comparing the models. The values for the measures derived from the walking distance approach are presented in Table 30.

[90] Use of a more sophisticated method is not always associated with a more reliable result. For example, in studying happiness, which is a categorical variable, the use of an ordered probit method, which takes the ordinality of the dependent variable into account, does not enhance the results derived from an OLS estimation, which does not take the ordinality into account (Ferrer-i-Carbonell & Frijters, 2004).

Appendix III.1 Models derived from the combinations of relative crowding and density measures

The following table presents the combinations of relative crowding measures and the relative density measures.

Table 36. Combinations of crowding and density measures included in estimations of equation (38).

Models	Relative C	Relative Density	Raw and relative crowding and density measures included in each model
1	MC	MADENS	PPBR, $PPBR^2$, MC, DENSM, $DENSM^2$, MADENS
2		MTDENS	PPBR, $PPBR^2$, MC, DENSM, $DENSM^2$, MTDENS
3		ATDENS	PPBR, $PPBR^2$, MC, DENSA, $DENSA^2$, ATDENS
4		MAVADENS	PPBR, $PPBR^2$, MC, DENSM, $DENSM^2$, MAVARDENS,
5		MTVADENS	PPBR, $PPBR^2$, MC, DENSM, $DENSM^2$, MTVARDENS
6		ATVADENS	PPBR, $PPBR^2$, MC, DENSA, $DENSA^2$, ATVARDENS
7	AC	MADENS	PPBR, $PPBR^2$, AC, DENSM, $DENSM^2$, MADENS
8		MTDENS	PPBR, $PPBR^2$, AC, DENSM, $DENSM^2$, MTDENS
9		ATDENS	PPBR, $PPBR^2$, AC, DENSA, $DENSA^2$, ATDENS
10		MAVADENS	PPBR, $PPBR^2$, AC, DENSM, $DENSM^2$, MAVARDENS
11		MTVADENS	PPBR, $PPBR^2$, AC, DENSM, $DENSM^2$, MTVARDENS
12		ATVADENS	PPBR, $PPBR^2$, AC, DENSA, $DENSA^2$, ATVARDENS
13	TC	MADENS	PPBR, $PPBR^2$, TC, DENSM, $DENSM^2$, MADENS
14		MTDENS	PPBR, $PPBR^2$, TC, DENSM, $DENSM^2$, MTDENS
15		ATDENS	PPBR, $PPBR^2$, TC, DENSA, $DENSA^2$, ATDENS
16		MAVADENS	PPBR, $PPBR^2$, TC, DENSM, $DENSM^2$, MAVARDENS
17		MTVADENS	PPBR, $PPBR^2$, TC, DENSM, $DENSM^2$, MTVARDENS
18		ATVADENS	PPBR, $PPBR^2$, TC, DENSA, $DENSA^2$, ATVARDENS
19	MVARC	MADENS	PPBR, $PPBR^2$, MVARC, DENSM, $DENSM^2$, MADENS
20		MTDENS	PPBR, $PPBR^2$, MVARC, DENSM, $DENSM^2$, MTDENS
21		ATDENS	PPBR, $PPBR^2$, MVARC, DENSA, $DENSA^2$, ATDENS
22		MAVADENS	PPBR, $PPBR^2$, MVARC, DENSM, $DENSM^2$, MAVARDENS
23		MTVADENS	PPBR, $PPBR^2$, MVARC, DENSM, $DENSM^2$, MTVARDENS
24		ATVADENS	PPBR, $PPBR^2$, MVARC, DENSA, $DENSA^2$, ATVARDENS
25	AVARC	MADENS	PPBR, $PPBR^2$, AVARC, DENSM, $DENSM^2$, MADENS
26		MTDENS	PPBR, $PPBR^2$, AVARC, DENSM, $DENSM^2$, MTDENS
27		ATDENS	PPBR, $PPBR^2$, AVARC, DENSA, $DENSA^2$, ATDENS
28		MAVADENS	PPBR, $PPBR^2$, AVARC, DENSM, $DENSM^2$, MAVARDENS
29		MTVADENS	PPBR, $PPBR^2$, AVARC, DENSM, $DENSM^2$, MTVARDENS
30		ATVADENS	PPBR, $PPBR^2$, AVARC, DENSA, $DENSA^2$, ATVARDENS
31	TVARC	MADENS	PPBR, $PPBR^2$, TVARC, DENSM, $DENSM^2$, MADENS
32		MTDENS	PPBR, $PPBR^2$, TVARC, DENSM, DENSM2, MTDENS
33		ATDENS	PPBR, $PPBR^2$, TVARC, DENSA, $DENSA^2$, ATDENS

34		MAVADENS	PPBR, PPBR², TVARC, DENSM, DENSM², MAVARDENS
35		MTVADENS	PPBR, PPBR², TVARC, DENSM, DENSM², MTVARDENS
36		ATVADENS	PPBR, PPBR², TVARC, DENSA, DENSA², ATVARDENS
37	MHC	MADENS	PPBR, PPBR², MHC, DENSM, DENSM², MADENS
38		MTDENS	PPBR, PPBR², MHC, DENSM, DENSM², MTDENS
39		ATDENS	PPBR, PPBR², MHC, DENSA, DENSA², ATDENS
40		MAVADENS	PPBR, PPBR², MHC, DENSM, DENSM², MAVARDENS
41		MTVADENS	PPBR, PPBR², MHC, DENSM, DENSM², MTVARDENS
42		ATVADENS	PPBR, PPBR², MHC, DENSA, DENSA², ATVARDENS
43	AHC	MADENS	PPBR, PPBR², AHC, DENSM, DENSM², MADENS
44		MTDENS	PPBR, PPBR², AHC, DENSM, DENSM², MTDENS
45		ATDENS	PPBR, PPBR², AHC, DENSA, DENSA², ATDENS
46		MAVADENS	PPBR, PPBR², AHC, DENSM, DENSM², MAVARDENS
47		MTVADENS	PPBR, PPBR², AHC, DENSM, DENSM², MTVARDENS
48		ATVADENS	PPBR, PPBR², AHC, DENSA, DENSA², ATVARDENS
49	THC	MADENS	PPBR, PPBR², THC, DENSM, DENSM², MADENS
50		MTDENS	PPBR, PPBR², THC, DENSM, DENSM², MTDENS
51		ATDENS	PPBR, PPBR², THC, DENSA, DENSA², ATDENS
52		MAVADENS	PPBR, PPBR², THC, DENSM, DENSM², MAVARDENS
53		MTVADENS	PPBR, PPBR², THC, DENSM, DENSM², MTVARDENS
54		ATVADENS	PPBR, PPBR², THC, DENSA, DENSA², ATVARDENS
55	MHVARC	MADENS	PPBR, PPBR², AVARC, DENSM, DENSM², MADENS
56		MTDENS	PPBR, PPBR², AVARC, DENSM, DENSM², MTDENS
57		ATDENS	PPBR, PPBR², AVARC, DENSA, DENSA², ATDENS
58		MAVADENS	PPBR, PPBR², AVARC, DENSM, DENSM², MAVARDENS
59		MTVADENS	PPBR, PPBR², AVARC, DENSM, DENSM², MTVARDENS
60		ATVADENS	PPBR, PPBR², AVARC, DENSA, DENSA², ATVARDENS
61	AHVARC	MADENS	PPBR, PPBR², TVARC, DENSM, DENSM², MADENS
62		MTDENS	PPBR, PPBR², TVARC, DENSM, DENSM2, MTDENS
63		ATDENS	PPBR, PPBR², TVARC, DENSA, DENSA², ATDENS
64		MAVADENS	PPBR, PPBR², TVARC, DENSM, DENSM², MAVARDENS
65		MTVADENS	PPBR, PPBR², TVARC, DENSM, DENSM², MTVARDENS
66		ATVADENS	PPBR, PPBR², TVARC, DENSA, DENSA², ATVARDENS
67	THVARC	MADENS	PPBR, PPBR², MHC, DENSM, DENSM², MADENS
68		MTDENS	PPBR, PPBR², MHC, DENSM, DENSM², MTDENS
69		ATDENS	PPBR, PPBR², MHC, DENSA, DENSA², ATDENS
70		MAVADENS	PPBR, PPBR², MHC, DENSM, DENSM², MAVARDENS

71		MTVADENS	PPBR, $PPBR^2$, MHC, DENSM, $DENSM^2$, MTVARDENS
72		ATVADENS	PPBR, $PPBR^2$, MHC, DENSA, $DENSA^2$, ATVARDENS
73	MCS	MADENS	PPBR, $PPBR^2$, AHC, DENSM, $DENSM^2$, MADENS
74		MTDENS	PPBR, $PPBR^2$, AHC, DENSM, $DENSM^2$, MTDENS
75		ATDENS	PPBR, $PPBR^2$, AHC, DENSA, $DENSA^2$, ATDENS
76		MAVADENS	PPBR, $PPBR^2$, AHC, DENSM, $DENSM^2$, MAVARDENS
77		MTVADENS	PPBR, $PPBR^2$, AHC, DENSM, $DENSM^2$, MTVARDENS
78		ATVADENS	PPBR, $PPBR^2$, AHC, DENSA, $DENSA^2$, ATVARDENS
79	CVARS	MADENS	PPBR, $PPBR^2$, THC, DENSM, $DENSM^2$, MADENS
80		MTDENS	PPBR, $PPBR^2$, THC, DENSM, $DENSM^2$, MTDENS
81		ATDENS	PPBR, $PPBR^2$, THC, DENSA, $DENSA^2$, ATDENS
82		MAVADENS	PPBR, $PPBR^2$, THC, DENSM, $DENSM^2$, MAVARDENS
83		MTVADENS	PPBR, $PPBR^2$, THC, DENSM, $DENSM^2$, MTVARDENS
84		ATVADENS	PPBR, $PPBR^2$, THC, DENSA, $DENSA^2$, ATVARDENS

Appendix IV. Table of results

Table 37 presents the estimated results in presence of all control variables. As mentioned in Section 6.6, the research design of this chapter is similar to the methodology of Chapter 7. Hence, the first model specification is:

$$y_i = \alpha_0 + \alpha_1 . x_i + \alpha_2 . x_i^2 + \alpha_3 . r_i + \beta . \text{Controls}_i + \varepsilon_i \tag{37}$$

Where the dependent variable (y_i) is individual i's residential satisfaction. x_i includes the measures of households' absolute position, including household's crowding and area density. x_i^2 is the squared term of the absolute measures and r_{in} measures the relatives, including relative crowding and relative density. ε_i measures the residuals. The combinations of the variables of interest, including both the relative and the absolute measures, are depicted in Appendix iii.1.

Similar to the last chapter, the dependent variable (y_i) follows a Bernoulli distribution. As discussed in the previous chapters, a logit transform function is used.

The results derived from the estimation of the first model estimation are presented in the first column of Table 37. Accordingly, holding all other factors constant, a higher level of absolute crowding lowers residential satisfaction. After controlling for all other factors, an increase in relative crowding does not affect RS. However, a higher level of relative density is associated with higher RS. As in the previous chapter, a higher area density affects RS negatively.

A higher crowding or density level may be desirable for some people based on their backgrounds. The endogeneity issue will be addressed, by employing a two-stage least squares method, as I did in the previous chapter; see Section 7.6. The same approach is taken to estimate the relative crowding measures' expected values; that is:

$$r_{it} = \theta_0 + \theta_1 . z_{it} + u_{it}, t = 2008 \tag{38}$$

and, by using the same instrument list as in Chapter 7, the estimated expected relative measures will be equal to

$$\hat{r}_{it} = \widehat{\theta_0} + \widehat{\theta_1} . z_{it}, t = 2010, 2012 \tag{39}$$

Therefore, the following equation estimates the impact of the variables of interest, after taking their endogeneity into account:

$$y_i = \beta_0 + \beta_1 . \hat{x}_i + \beta_2 . \hat{x}_i^2 + \beta_3 . \hat{r}_i + \beta_4 . (\text{Control variables})_i + \psi_i, \begin{bmatrix} \psi_i \\ \left(\sum_{t=2008}^{2012} u_t \right)_i \end{bmatrix} \sim NID \left(0, \begin{bmatrix} \sigma_\psi^2 & 0 \\ 0 & \sigma_u^2 \end{bmatrix} \right) \tag{40}$$

where β_1, β_2 and β_3 are the parameters of interest and ψ_i is the disturbance term, which follows a normal distribution with a mean equal to zero and a variance equal to σ^2, i.e. $\psi_i \sim N(0, \sigma_\psi^2)$. The instruments ($z_{it}$), included in equation (39), need to be exogeneous; that is, not correlated with ψ_i.[91] The basis for the choice of instruments is presented in Section 7.2.1. All of the variables included in Table 21 will be used as instruments, except for ethnicity, which has a high correlation with ψ_i. Therefore, z_{it} in equation (25) is a vector of 15 binary variables.

The results of the second model specification, presented in the second column of Table 37, indicate an insignificant effect of absolute values when the endogeneity is taken into account. Compared with the last chapter, the impact of the expected absolute values has become less significant. As in the first model specification, relative crowding does not affect RS. The expected relative density, after taking its endogeneity into account, still has a significant positive impact on RS.

The last two columns in Table 37are the result of the estimations presented in Section 8.6.

[91] As stated earlier, u_{it} follows a binomial distribution. However, the binomial distribution approximates to the normal distribution as the number of observations tend to infinity. Therefore, we can assume an asymptotic normal distribution for u_{it} (Menard, 2002).

Table 37. Results table including the control variables and step by step estimations.

	(1) Specification I	(2) Specification II	(3) Specification III	(4) Specification III + PC
PPBR	0.163***†			
	(0.109)			
$PPBR^2$	1.581†			
	(0.492)			
$\widehat{PPBR}$		2.773†	1.485	1.487
		(4.15)	(2.655)	(3.284)
$\widehat{PPBR}^2$		0.593*†	1.136	0.953
		(0.489)	(0.789)	(7.006)
$PPBR - \widehat{PPBR}$			0.151**†	0.210**†
			(0.111)	(0.155)
$PPBR^2 - \widehat{PPBR}^2$			1.352†	1.343†
			(0.490)	(0.500)
DENSA	0.799**††			
	(0.086)			
$DENSA^2$	1.0313††			
	(0.019)			
$\widehat{DENSA}$		1.049††	0.951	0.785
		(2.1)	(1.959)	(1.678)
$\widehat{DENSA}^2$		0.881††	0.821	0.840
		(0.248)	(0.242)	(0.258)
$DENSA - \widehat{DENSA}$			0.950	0.965
			(0.183)	(0.193)
$DENSA^2 - \widehat{DENSA}^2$			0.936	0.931
			(0.050)	(0.051)
CVARS	0.997†			
	(0.082)			
$\widehat{CVARS}$		2.856†	2.884	2.531
		(3.269)	(3.332)	(3.017)
$CVAR - \widehat{CVARS}$			1.160	1.164
			(0.153)	(0.152)
ATVARDENS	1.00004**††			
	(0.000)			
$ATV\widehat{ARD}ENS$		1.00007***	1.00007**††	1.00002**††
		(0.000)	(0.000)	(0.000)
$ATVARDENS - ATV\widehat{ARD}ENS$			1.00004**††	1.00004**††
			(0.000)	(0.000)
PC				0.211***
				(0.050)
MOCCN_1	3.773642243	2.586082432	2.288175958	3.422709890
	(3.268)	(4.727)	(4.299)	(6.321)

MOCCN_2	1.278103784	2.430534312	1.979118138	2.053454788
	(0.711)	(2.029)	(1.707)	(1.824)
MOCCN_3	0.691624657	2.103846483	2.031634654	1.972512813
	(0.828)	(3.616)	(3.517)	(3.484)
MOCCN_4	2.143895471	1.33932e+02	58.036843711	1.00684e+02
	(4.704)	(558.002)	(242.527)	(421.892)
MOCCN_5	0.413031548	0.230408952	0.279046596	0.461324784
	(0.614)	(0.351)	(0.437)	(0.720)
MOCCN_6	4.473675114	2.992881300	3.135888576	3.953920715
	(7.310)	(4.820)	(5.138)	(6.460)
MOCCN_7	8.570066573	10.001271548	10.320781102	17.769041973
	(47.745)	(53.009)	(57.377)	(101.270)
MOCCN_8	1.640808231	0.625864891	0.473589636	0.474178776
	(5.115)	(2.007)	(1.580)	(1.643)
MOCCN_9	2.58271e+02	9.807547248	3.590172765	1.862517121
	(1885.324)	(71.599)	(21.855)	(11.290)
MOCCT_1	2.242744338	1.860878692	2.406224185	1.754714729
	(2.105)	(5.464)	(7.163)	(5.097)
MOCCT_2	1.106675225	1.069582274	1.926339148	1.077907542
	(0.679)	(5.549)	(10.137)	(5.556)
MOCCT_3	0.078217351**	0.056900689	0.049579210	0.066340783
	(0.083)	(0.163)	(0.144)	(0.193)
MOCCT_4	0.388132229	4.258947808	3.603216827	1.973276666
	(0.443)	(24.943)	(21.226)	(11.688)
MOCCT_5	0.187797774	0.325867823	0.249899459	0.322949921
	(0.208)	(0.775)	(0.570)	(0.730)
MOCCT_6	0.177589037	0.104159076	0.134096439	0.095616946
	(0.217)	(0.308)	(0.416)	(0.289)
MOCCT_7	0.709176477	3.743966686	2.558858162	2.710210144
	(1.335)	(8.462)	(5.830)	(5.983)
MOCCT_8	1.044027523	2.617917499	1.618422844	1.637298279
	(1.140)	(3.565)	(2.324)	(2.391)
MOCCT_9	2.761879759	3.141858220	2.644352286	1.928101766
	(3.369)	(14.228)	(11.863)	(8.578)
METHN_1	1.465169044	0.666566749	0.600101343	0.679785429
	(1.040)	(0.774)	(0.721)	(0.854)
METHN_2	2.887964425	2.245274768	2.871932767	3.093138502
	(4.235)	(5.634)	(7.323)	(7.831)
METHN_3	5.837036815**	7.776422610	6.421038429	5.057199153
	(4.623)	(14.605)	(12.381)	(9.849)
METHN_4	1.689916985	2.410598309	2.118291384	2.024820121
	(1.704)	(2.263)	(2.089)	(2.071)
METHN_5	0.058347226	1.743631572	2.571215044	1.628606547
	(0.286)	(8.915)	(11.486)	(7.178)
METHT_1	5.293257575	0.810074432	1.792869953	3.175135440
	(5.951)	(2.348)	(5.200)	(9.309)

METHT_2	0.802277152	0.432205721	0.712228066	0.590662041
	(0.789)	(1.203)	(2.041)	(1.612)
METHT_3	1.483764891	0.605895828	0.990160809	0.985488937
	(1.579)	(1.655)	(2.738)	(2.650)
METHT_4	4.352149173	0.330547331	1.180669429	1.391014473
	(4.919)	(0.821)	(3.029)	(3.421)
METHT_5	15.567025445**	4.566291289	8.316806876	13.772387294
	(20.880)	(7.767)	(13.930)	(23.201)
Gender	0.956682974	0.936839712	0.960599613	0.962970276
	(0.097)	(0.119)	(0.122)	(0.122)
Partner	1.135192406	0.517079654	0.538913005	0.520274937
	(0.155)	(0.305)	(0.317)	(0.316)
Age (>74)	1.194494522	0.286997044	0.290543529	0.228661968
= 15–19	(0.508)	(0.427)	(0.437)	(0.349)
= 20–24	0.824798003	0.276366476	0.289657232	0.242993639
	(0.258)	(0.312)	(0.330)	(0.284)
= 25–34	0.483480636**	0.197183441*	0.225410872*	0.188780772*
	(0.135)	(0.164)	(0.191)	(0.163)
= 35–54	0.481669147***	0.237942287**	0.240414307**	0.217590836**
	(0.119)	(0.165)	(0.170)	(0.157)
= 55–59	0.752728932	0.690741772	0.693570990	0.729115071
	(0.189)	(0.292)	(0.306)	(0.321)
= 60–64	0.865697137	0.914985009	0.872489388	0.901399156
	(0.245)	(0.393)	(0.390)	(0.405)
= 65–69	0.837508287	0.922394470	0.898618630	0.944607836
	(0.187)	(0.407)	(0.404)	(0.430)
= 70–74	0.816596093	0.855717213	0.837380751	0.840639750
	(0.194)	(0.274)	(0.276)	(0.276)
Length of stay in NZ (<4 years)				
= 4–10 years	1.039816422	0.974974972	0.985657532	1.034551682
	(0.339)	(0.424)	(0.426)	(0.456)
= 10–25 years	0.852639806	1.042213838	0.912870406	1.080941598
	(0.264)	(0.751)	(0.661)	(0.787)
>25 years	1.234276908	1.495834326	1.403923881	1.623127006
	(0.383)	(0.979)	(0.921)	(1.088)
Ethnicity (European)				
= Maori	1.249026923	0.351148884	0.349249303	0.294727113
	(0.699)	(0.330)	(0.334)	(0.284)
= Pacific	0.759100468	0.083602224	0.111738559	0.078987065
	(0.385)	(0.133)	(0.178)	(0.127)
= Asian	0.780896801	0.331357610	0.335578756	0.317526977
	(0.441)	(0.335)	(0.342)	(0.332)
= Other	1.956692511	0.519321804	0.504390995	0.444153664
	(1.336)	(0.499)	(0.476)	(0.424)
Education (No qualification)				

= Diploma	0.995446768	1.779593527	1.611906068	1.883909714
	(0.206)	(0.806)	(0.742)	(0.872)
= Degree	0.909233385	1.699479170	1.553932481	1.804915715
	(0.193)	(0.713)	(0.660)	(0.779)
Household income ($150,001 or more)				
= Zero	0.836064260	1.468139044	1.671231974	1.649606051
	(0.459)	(2.099)	(2.397)	(2.380)
=$1–$5,000	8.731098076	3.046482063	5.653841620	1.619442466
	(3.030)	(2.619)	(4.448)	(3.216)
=$5000–$50,001	1.375959962	1.422914049	0.963548704	0.963548704
	(0.126)	(0.460)	(0.573)	(0.506)
=$50,001–$70,000	1.028828247	1.081627525	0.916915592	0.916915592
	(0.154)	(0.350)	(0.387)	(0.379)
=$70,001–$150,000	1.040985427	1.049665906	0.878567774	0.878567774
	(0.117)	(0.236)	(0.267)	(0.249)
Health (Very dissatisfied)				
= Dissatisfied	2.455015282**	3.759123436***	3.418319962***	3.723971359***
	(0.929)	(1.650)	(1.498)	(1.670)
= No feeling either way	2.910252475***	5.483980423***	4.967610705***	5.679479396***
	(0.946)	(2.601)	(2.342)	(2.701)
= Satisfied	3.342035972***	6.607942945***	6.078842816***	6.869102359***
	(1.082)	(3.352)	(3.057)	(3.491)
= Very Satisfied	4.255019083***	8.005834580***	7.277856728***	8.518748053***
	(1.463)	(4.010)	(3.631)	(4.299)
Abilities (Very dissatisfied)				
= Dissatisfied	1.009716047	0.513580950	0.532649075	0.280710383
	(1.529)	(1.011)	(1.049)	(0.506)
= No feeling either way	1.173608533	0.869238828	0.846236219	0.475434346
	(1.715)	(1.427)	(1.373)	(0.687)
= Satisfied	1.588961702	1.265272994	1.202520229	0.667721042
	(2.320)	(2.029)	(1.909)	(0.944)
= Very Satisfied	2.485082999	2.263822667	2.123743029	1.214532584
	(3.663)	(3.677)	(3.421)	(1.736)
Bad street access	0.545932765	0.584555244	0.585106998	0.690808590
	(0.274)	(0.362)	(0.371)	(0.431)
Poor condition	0.318739979***	0.248905899***	0.263458372***	0.273386724***
	(0.100)	(0.076)	(0.080)	(0.088)
Damp dwelling	0.552504592***	0.531845853***	0.532762006***	0.574989746**
	(0.114)	(0.118)	(0.114)	(0.127)
Difficult to heat	0.462896264***	0.448736078**	0.466085246**	0.468008338**
	(0.091)	(0.150)	(0.155)	(0.154)
Having pests	1.418835083	1.018307504	1.020988614	0.932830813
	(0.392)	(0.400)	(0.410)	(0.385)

Expensive house	0.571759966**	0.580061287*	0.563056461*	0.567763293*
	(0.142)	(0.190)	(0.188)	(0.185)
Work distance	1.056730311	1.052047702	1.041266647	1.162549253
	(0.254)	(0.360)	(0.364)	(0.411)
Far from facilities	0.556737332	0.582820089	0.521716651	0.503441798
	(0.244)	(0.416)	(0.374)	(0.367)
Unsafe neighbourhood	0.748798452	0.554649470	0.674642333	0.615401140
	(0.221)	(0.347)	(0.426)	(0.392)
Noise and vibration	0.606096076***	0.462520172	0.510262517	0.468369908
	(0.096)	(0.295)	(0.328)	(0.295)
Air polluted	0.950848159	0.891128649	1.001897667	0.948742175
	(0.236)	(0.313)	(0.363)	(0.356)
Tenure status (Owner)				
= Renter	0.654823638***	0.349872972**	0.410398217*	0.352010594**
	(0.080)	(0.166)	(0.198)	(0.171)
= Family trust	1.161032457	1.347403494**	1.288995792*	1.338522699*
	(0.133)	(0.195)	(0.191)	(0.205)
Free Time (Not enough free time)				
= The right amount of free time	1.263658984**	1.339463830***	1.341334332***	1.271816667**
	(0.127)	(0.139)	(0.137)	(0.131)
= Too much free time	1.245311877	1.317763602	1.334058157	1.253256997
	(0.245)	(0.269)	(0.270)	(0.257)
Socialising (Every day)				
= At least once in the last four weeks	0.857802963	0.716343343	0.709917828	0.717243094
	(0.147)	(0.194)	(0.194)	(0.196)
= Around once a fortnight	0.833154200	0.756876823	0.737764711	0.759492887
	(0.118)	(0.216)	(0.215)	(0.223)
= Around 1–6 times a week	0.907010901	1.002974462	1.002428120	1.025601833
	(0.089)	(0.126)	(0.128)	(0.134)
Local political involvement	0.958519474	0.953390210	0.934866546	0.964220817
	(0.105)	(0.144)	(0.143)	(0.151)
Recycle== A little	1.988799174	0.999051563	1.139272680	0.919141552
	(1.170)	(1.319)	(1.517)	(1.226)
Recycle== Some	0.808795523	0.806604043	0.735105239	0.777830719
	(0.299)	(0.405)	(0.380)	(0.388)
Recycle== Most	0.596528053***	0.550281826**	0.553508346**	0.517804574**
	(0.109)	(0.152)	(0.151)	(0.141)
Water use (Never)				
= A little of the time	1.105957108	1.402989998	1.429928440	1.455450772
	(0.158)	(0.406)	(0.412)	(0.410)
= Some of the time	1.119201116	1.258687399	1.275896889	1.249189807
	(0.162)	(0.384)	(0.393)	(0.380)
= Most of the time	0.984439571	1.029135244	1.052461141	1.046410509
	(0.108)	(0.276)	(0.283)	(0.280)

Energy use (Never)				
= A little of the time	0.629928725**	0.539993528	0.581579456	0.529996551
	(0.114)	(0.216)	(0.234)	(0.212)
= Most of the time	0.841085059	0.805310340	0.808769226	0.809380483
	(0.091)	(0.153)	(0.156)	(0.157)
Council services (Very satisfied)				
= Very Dissatisfied	0.894523658	1.347661891	1.037714645	1.091088153
	(0.344)	(2.000)	(1.564)	(1.676)
= Dissatisfied	0.504886024***	0.698455132	0.592769133	0.650659518
	(0.090)	(0.460)	(0.395)	(0.440)
= No feeling either way	0.405143533***	0.598503606	0.553084579	0.595611575
	(0.069)	(0.267)	(0.249)	(0.271)
= Satisfied	0.432196215***	0.510649329**	0.479669947***	0.502690071**
	(0.056)	(0.132)	(0.126)	(0.132)
Facilities access (All of them)				
= Never been to	0.024118657	0.066679978	0.057970140	0.098423228
	(0.096)	(0.280)	(0.242)	(0.405)
= A few of them	0.225732715	0.256833528	0.222170287	0.209288408
	(0.368)	(0.410)	(0.367)	(0.349)
= None of them	0.836009886	0.581577080	0.563717878	0.569360367
	(0.442)	(0.343)	(0.332)	(0.327)
Green space access (Few of them)				
= Most of them	0.449320037	0.144453991*	0.185844472	0.160850463*
	(0.356)	(0.155)	(0.202)	(0.175)
= All of them	1.016827818	0.789371889	0.867569696	0.812589641
	(0.300)	(0.286)	(0.321)	(0.306)
Green space state (Very satisfied)				
= Not been to	0.261113192*	0.505868715	0.410772513	0.432628425
	(0.193)	(0.527)	(0.447)	(0.462)
= Very Dissatisfied	1.408579712	0.589484942	0.744065203	0.476074740
	(2.220)	(1.124)	(1.500)	(0.929)
= Dissatisfied	0.867353054	0.907250675	0.884077781	0.888043958
	(0.265)	(0.272)	(0.273)	(0.278)
= No feeling either way	1.009716379	0.950099010	0.958992567	0.908370835
	(0.237)	(0.397)	(0.411)	(0.385)
Coastline access (Few of them)				
= Most of them	1.958772418	2.914129304*	3.011910557*	3.212369616*
	(1.055)	(1.676)	(1.732)	(1.951)
= All of them	1.136615003	0.857135323	0.925315849	0.830027456
	(0.319)	(0.514)	(0.554)	(0.503)
Coastline state (Very satisfied)				
= Not been to	2.178122111	2.094293328	2.249440645	1.977561191
	(1.570)	(1.725)	(1.929)	(1.701)
= Very Dissatisfied	1.658746619	2.278697002	2.432189463	2.648618928

	(1.222)	(2.255)	(2.503)	(2.719)
= Dissatisfied	0.991591567	1.161626122	1.132177116	1.206458849
0	(0.196)	(0.360)	(0.342)	(0.366)
= Satisfied	0.870593612	0.945512962	0.942342140	0.955508427
	(0.157)	(0.321)	(0.324)	(0.327)
Household size (>4 people)				
= Single person	0.351681354***	0.825917140	0.383275727***	0.514557613*
	(0.117)	(0.162)	(0.129)	(0.176)
= 2 people	0.513393990**	0.881068466	0.498458728***	0.595085336**
	(0.131)	(0.155)	(0.130)	(0.155)
= 3 people	0.572115095***	0.764551561	0.556897571***	0.631442110**
	(0.118)	(0.131)	(0.119)	(0.137)
= 4 people	0.723248470	0.813310155	0.706661458	0.766309033
	(0.152)	(0.154)	(0.151)	(0.165)
Constant	1.000547681	0.034438153	0.289088220	0.031745377
	(2.151)	(0.286)	(2.465)	(0.270)
Observations	4607	4607	4607	4607
Population size	904053	904053	904053	904053
R^2	0.193	0.195	0.192	0.196
AIC	5033.852	5067.771	5033.65	5025.655
Joint significance:				
Crowding measures†	<0.001	0.204	0.005	<0.001
Density measures††	0.002	0.816	0.022	0.033

Reported coefficients are exponentiated ones. Standard errors are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

† The significance level for PPBR and its squared (Squared PPBR) variables is reported jointly on the Crowding Wald test row. Note that the joint significance tests are reported for different coefficients in different models. The reported AICs are for the models that do not take jackknife replications into account.

9 Negative Envy or Positive Amenity Effects? A Neighbourhood Study of Aucklanders' Residential Satisfaction

Abstract

Residents of a deprived area who are located next to an affluent area may enjoy the amenities provided by the rich area, or they may suffer from comparing themselves with the rich people. In order to distinguish between these effects, the endogeneity of residential characteristics should be taken into account; that is, the individual's location and housing choices may be correlated with idiosyncratic factors that affect an individual's satisfaction with her residential environment. My results indicate that people are negatively affected by locating in an area with a deprivation level higher than that which are expected to experience. Furthermore, I find that residents of poor (rich) areas are positively (negatively) affected by the affluence (poverty) of neighbouring areas, which confirms the presence of a positive amenity effect for the residents of poor neighbourhoods and a negative NIMBY effect for the residents of rich neighbourhoods.

9.1 Introduction

A key objective of urban policy is to raise levels of residential satisfaction (that is, satisfaction with one's neighbourhood and housing situation). In this study, I assess how neighbourhoods' deprivation levels spill over to people's perceptions of their residential situation by considering the spatial effects that neighbourhoods have on one another. The deprivation level of a neighbourhood is affected by both the process that has led individuals to live in their current neighbourhood and by the current situation of residents.

Social deprivation of neighbourhoods may affect their residents persistently. This 'poverty trap' effect occurs when the children of poorer (richer) families have access to lower (higher) quality amenities, such as a worse (better) education environment. (For example, Maani (2004) considered the effect of living conditions on educational outcomes in New Zealand.) This situation affects the next generation's earnings opportunities, which may lead to a perpetuation of similar living condition to the parents. The persistence of poor conditions can lead to the stratification of neighbourhoods. These negative effects may be mitigated by considering policies that account for the effect of the relative deprivation positions of adjacent neighbourhoods on the welfare of their residents.

Relative deprivation refers to the deprivation level of a neighbourhood compared to its adjacent neighbourhoods. People living in more deprived areas that are proximate to more affluent areas may suffer from mental health issues caused by social comparisons (Pearson, Griffin, Davies, & Kingham, 2012). The variation of deprivation levels across neighbourhoods may lead to community fragmentation, which is associated with less socialising and less volunteering. Conversely, residents in deprived areas may benefit from higher amenities in nearby affluent areas.

From an urban planning perspective, the need for a comprehensive study of deprivation effects using a spatial approach has been noted by Robson, Lymperopoulou and Rae (2008), who studied the effect of migrating between urban areas with different deprivation levels. In the present study, I assess the determinants of RS by paying attention to the spatial effects of neighbourhoods on one another.

Pepper potting is the scheme of combining private housing areas with social (state) housings. In 2006, when the current prime minister of New Zealand, John Key, was a member of parliament, he described the pepper potting project at Hobsonville Point as 'economic vandalism'. The reason for this was that social houses were going to be built on expensive land (Key, 2006). Expensive land is generally associated with more affluent areas with good access to facilities. These facilities may help people to quit a poor condition by, for example, accessing better education facilities.

It is important for policy makers and researchers to consider the mutual effect that neighbourhoods with different deprivation levels have on one another. The literature on developing countries urban areas includes many studies on shanty towns and their effect on adjacent neighbourhoods, but there have been fewer studies of such issues for developed countries. A better understanding of the mutual effect that neighbourhood types have on each other can result in avoiding negative effects such as the formation of ghettos, which may lead to social problems. Better understanding may also result in a strengthening of positive social ties by providing a better environment for residents to interact with their neighbours. Cheshire et al. (2014) noted the lack of credible studies on neighbourhood effects. They also highlighted the possibility of a positive impact for the residents of poor areas from locating next to rich neighbours:

> '… we might ask if there are positive externalities for poor people if they could live among richer neighbours. This is the underlying logic for believing that 'mixed neighbourhoods' would produce better economic (and social?) outcomes for poor people. Remember, however, that specialised neighbourhoods allow more choice and help people live with compatible neighbours, in areas with facilities and amenities useful to them (given their tastes and incomes) so sorting may well be a source of welfare gain. This implies that 'mixed communities' will only benefit the poor if any such positive externalities poor people got from living together with other poor people were outweighed by the positive externalities and the welfare gains they would get from living with richer neighbours.'

The effect of policy interventions can be assessed by people's choices, which reflects their preferences. Some urban studies have used this traditional economic approach by paying attention to hedonic prices, which consider the subjective aspects of house price evaluations. These studies are usually interested in the regression function of the choice of consumers on the characteristics of the property (Court, 1939; Rosen, 1974). As Sheppard (1999) mentioned, this kind of pricing is used in order to investigate the effect of different incomes, different preferences and non-market constraints on the choice of households. As Grimes, Oxley and Tarrant (2012) mentioned, a monetary analysis is not sufficient to determine the effect of a policy; different aspects of life satisfaction should also be considered in the analysis. Although price represents many characteristics of the house and the neighbourhood, it is not a complete measure, especially in the case of welfare. Recently, economists have also been interested in an explicitly result-oriented assessment of policy effects, arguing that experienced utility is an appropriate concept in order to evaluate policies' influences on people's life conditions (Tella, MacCulloch, & Oswald, 2003). Consequently, residential satisfaction, which is the output of interest throughout this thesis, will again be considered as the dependent variable in this chapter.

As discussed in previous chapters, residential satisfaction is affected by a range of objective factors, such as household crowding, and subjective factors such as residents' relative positions. The subjective factors account for social comparisons, which derive from comparisons with a reference group. Conceptually, the relevant reference group refers to those people who affect an individual the most. Different groups are considered as the reference group, including neighbours living at different geographic proximities and symbols of success at different geographic proximities. The current chapter uses the two relative measures of crowding and density introduced in Chapter 8 and introduces two new measures, namely the deprivation measure and the relative deprivation measure.

9.2 Literature review

A new phase of urban redevelopment took place as a result of greater emphasis on globalisation that led to a shift from manufacturing industries to services activities (Knox & Pinch, 2009). In some cities, this phenomenon has resulted in a complex city topology without any distinct city centre (Ingersoll, 1992). Since 1930, countries such as China, the UK and the US have implemented different forms of urban redevelopment (Carmon, 1999; Meen et al., 2013). In China, for example, as a result of the policy interventions in the redevelopment process, the low income population segment had to give up its property to the high- or medium-income segments and settled into cheaper accommodation on the fringe of the city (Wu & He, 2005). After 1990, in order to promote greater sustainable development of cities in China, by solving the poor areas' problems and increasing the quality of life of the low and medium deprivation level strata, policy makers paid more attention to the social aspects of urban redevelopment policies. Different kinds of redevelopment will lead to changes in the physical and socioeconomic characteristics of neighbourhoods and societies. However, there are few studies on these effects at the neighbourhood level, which is dealt with more thoroughly in the current study.

The importance of a higher deprivation level can be explained by considering its impacts on health outcomes, such as on the rates of suicide (Burrows, Auger, Gamache, St-Laurent, & Hamel, 2011) and on the rates of smoking (Moon, Pearce, & Barnett, 2012). The persistence of this impact, after controlling for individual-level characteristics, is demonstrated in a number of studies (e.g. Abas et al., 2006; Barnett & Lauer, 2003). Depending on the field of research, the measurement of deprivation level varies amongst studies. In previous studies, deprivation levels are measured based on the mean income levels and/or the mean property values of dwellings (Nieuwenhuis, Völker, & Flap, 2013). In a number of studies, the deprivation level is measured by using neighbourhood indices, such as the New Zealand Deprivation Index (NZDep). These studies usually seek the importance of the living environment in people's life. On the other hand, because deprivation indices are an aggregate measure, they may hide exact sources of a relationship.

The social influences that individuals have on one another through relative situations can be discussed using contagion models; in other words, the behaviours of individuals may affect their neighbours through local interactions. We may expect higher satisfaction levels for neighbours with similar characteristics because of more interactions. Local interactions may arise from similarities. For example, an individual may participate in some groups because of their peers. Dissatisfaction may arise from a conflict caused by negative relationships with neighbours. Thus, a conflict is more likely to arise in a mixed neighbourhood than in an integrated neighbourhood. On the other hand, from a competition theory perspective, people may feel weak/dissatisfied in the presence of others who are similar to them because they have a feeling of competition. For example, an individual who has a sense of competition with colleagues may suffer a lower level of job satisfaction.

To account for social comparisons across neighbourhoods, Pearson et al. (2012) calculated the socioeconomic isolation quantiles based on the relative deprivation level of an area to its adjacent neighbourhoods, which is very similar to the geographic approach I have adopted in the current study. They have observed a higher rate of mental illness in areas surrounded by neighbourhoods with similar levels of deprivation, which is arguably justified by the poverty trap explanation (with alcohol and drug abuse likely to be higher in those areas). The authors mentioned the need for a more precise outcome measure and also the need to account for a causal relationship at an individual level. Clark and Morrison (2012) examined the nature of residential sorting using data from New Zealand Dynamics of Motivation and Migration Survey. They defined a matrix of social spaces defined by socio-economic status for capturing the changes in opportunities resulting from moving across neighbourhoods. They concluded that the residents of deprived areas are less likely to upgrade to a more prosperous area, particularly for those with lower incomes.

Luttmer (2005) studied the impact of a higher income level of neighbours on an individual's happiness. His results suggest that a higher income of neighbours leads to lower self-reported happiness. He also noted that the definition of happiness may differ amongst people. However, by using the outcome measures that are less prone to the definition changes, he concluded that differences in people's definition of happiness does not affect the negative correlation between neighbours' income levels and an individual's happiness. The magnitude of the decrease in a person's happiness because of an increase in his or her neighbour's income is reported to be equal to the increase in the person's own happiness level caused by an increase in their own income.

In New Zealand, there are have been few economic studies on the deprivation levels of neighbourhoods. Grimes et al. (2006) found that the ratio of house rents relative to house prices is higher in more deprived areas than in less deprived areas. They introduced some plausible reasons for this phenomenon and noted that it needs further analyses. In another study, Morrison and Nissen (2010) used the New Zealand Index of Deprivation to rank neighbourhoods at two waves of Census: 2001 and 2006. They showed that, for movers, the socio-economic rank of the origin is closely related to their destination neighbourhood. Accordingly, the residents of relatively deprived areas are more likely to move to areas with a similar level of deprivation. The authors showed strong socioeconomic patterns in people's ability to improve the socioeconomic status of their living environment.

9.3 Hypothesis development

Although the socio-spatial impacts of urban forms have been examined in a number of studies (e.g., He & Wu, 2007), there is a lack of neighbourhood-level studies. The effect of poor neighbourhoods on richer ones and vice versa could be very important when making policy and consequently for decisions that affect urban form. Morrison (2015b, p. 76), in examining the role of residential sorting, concluded that:

> '[This article] has addressed a paradox. On one hand researchers worldwide have found it extremely difficult to consistently identify the negative effects of living in poor neighbourhoods, over and above the personal difficulties faced by the residents who self-select into those neighbourhoods. On the other hand, neighbourhoods continue to matter immensely to those at the affluent end of the income spectrum. The revealed preference of high-income, high-wealth households for residing with others like themselves speaks to the substantial net benefits they expect to accrue from such decisions.'

This complexity of the proximities of poor and rich neighbourhoods and also the location choices needs to be addressed very carefully. The present study must take a wide range of control variables into consideration to account for both subjective and objective issues and also a range of identification issues, including location choice. It is likely that rich neighbourhoods will be affected negatively by being adjacent to poor neighbourhoods. However, poor neighbourhoods may be affected either positively by being adjacent to a rich neighbourhood (with better local amenities) or negatively (an envy effect). I hypothesise that:

H1: A higher deprivation gap between adjacent neighbourhoods lowers the residential satisfaction of the residents of (relatively) rich area and raises the residential satisfaction of residents of poor neighbourhoods. Thus, for poor neighbourhoods, the positive amenity effect is hypothesised to outweigh any (negative) envy effect.

9.4 Data and sample

Similar to previous chapters, Census 2006 is merged with three cross-sections of NZGSS (for more details, see Section 7.4). In order to take account of the geographic proximity of mesh-blocks, four geospatial datasets are used: mesh-blocks dataset, address points dataset, roads dataset, and tracks datasets. After using the appropriate projection system and editing the datasets, a matrix of spatial proximity is derived by using a network analysis toolbox (for more details, see Appendix i.1 in Chapter 8).

9.4.1 Index of Deprivation

New Zealand's Deprivation Index 2006 (NZDep2006) is the combination of nine variables from New Zealand's 2006 Census, which are described more in Table 38. This index covers eight dimensions of deprivation and dedicates a deprivation score to each mesh-block. NZDep2006 categorises mesh-blocks by using the first principal component score, which has been scaled up to have 1000 index points as its mean. (The ordinal value of the principal component is also available.[92] For more information, see Salmond, Crampton and Atkinson (2007)).

Table 38. Nine variables from the 2006 Census are used to measure the deprivation scores.

Dimension of deprivation	Description
Income	People aged 18–64 receiving a means tested benefit
Income	People living in equivalised[93] households with income below an income threshold
Owned home	People not living in own home
Support	People aged <65 living in a single-parent family
Employment	People aged 18–64 unemployed
Qualifications	People aged 18–64 without any qualifications
Living space	People living in equivalised households below a bedroom occupancy threshold
Communication	People with no access to a telephone
Transport	People with no access to a car

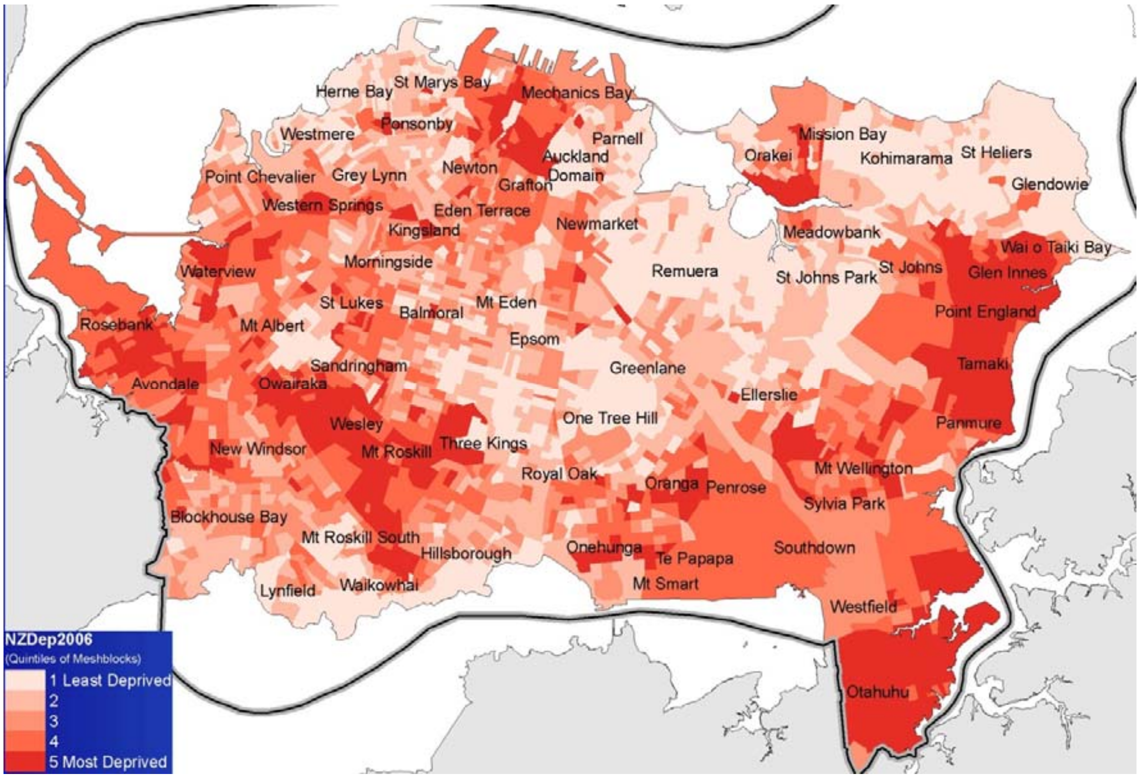
Source: Salmond, Crampton and Atkinson (2007)

[92] The deprivation index is a division of the deprivation scores into ten bins to derive an ordinal scale. However, the deprivation scale is a continuous variable. A higher level of the deprivation scale is an indication of a higher deprivation.

[93] The household composition has been controlled for.

Figure 21 illustrates the distribution of deprivation across Auckland city's territorial authority.

Figure 21. Atlas of socioeconomic deprivation levels amongst urban areas of the 'Auckland city' territorial authority of the Auckland region.



Source: White, Gunston, Salmond, Atkinson and Crampton (2008).

9.4.2 The construction of the relative deprivation measure[94]

I concluded the last chapter by stating that, amongst the relative crowding measures, those constructed based on the walking distance definition[95] of neighbourhood have the highest predictive power for RS. Following this, I have taken a walking distance approach to construct the current chapter's spatial weight matrix. This method was explained in Appendix i.1 of Chapter 8, and for more details on relative measurement, see Section 8.3.2 (relative measurement).

Living in a deprived area may not be undesirable if the area is surrounded by less deprived areas. Hence, the deprivation level of an individual's neighbourhood may be compared with the deprivation level of its adjacent neighbourhoods. For example, the relative deprivation of individuals ($RDEP_i$), which is an ix1

---

[94] For a review of spatial weight matrix construction, see Appendix i in Chapter 8.

[95] A walking distance definition of neighbourhood boundaries does not limit the boundaries of the neighbourhood to administrative or non-administrative boundaries but to an area that, for example, can be accessed by using walking path or by using public transportation.

column vector,[96] is defined as the difference between the deprivation level of their MB (DEP_i)[97] and the deprivation level of their neighbourhoods' adjacent MBs ($ADEP_i$):[98]

$$RDEP_i = DEP_i - ADEP_i \tag{41}$$

$RDEP_i$ is the deprivation level of individual i's area compared to its adjacent areas. The impact of RDEP on RS may be different for those who live in rich areas than for those who live in poor areas. The reason for attributing the average deprivation in the neighbourhood to each individual living in that area is discussed in Appendix i. I check whether the deprivation score of an individual's neighbourhood is higher or lower than the average deprivation scores of its adjacent neighbourhoods; in other words, I distinguish between people who, compared to their neighbourhoods' adjacent areas, live in relatively deprived areas and those who live in relatively affluent areas. Based on the sign of RDEP, a dummy variable (Dummy) is constructed as follows:

$$\text{Dummy} = 1\ (\text{RDEP} > 0);\ 0\ (\text{Otherwise}). \tag{42}$$

Hence, Dummy is equal to 1 when an individual resides in an area with a higher deprivation level than his or her neighbourhood's adjacent areas.

9.5 Descriptive statistics

The variables of this study have been described in Appendix I of Chapter 7. In the current section, I will only briefly describe the variables of interest. As depicted in table 39, deprivation level of each MB (DEP) is a score with a minimum of 866 and a maximum of 1436. As depicted in the P_value column, the mean difference amongst years is not statistically significant, at better than 0.1 per cent level, and the overall mean is equal to 990.147 (derived from the sample population of 4607 people).

---

[96] Vector and matrix dimensions are signified by the subscripted descriptors in all equations.

[97] Defined as follows:

$$DEP = [DEP_i], \qquad DEP_i = \sum_{j=1}^{J} MD_{ij}.\,MDEP_j$$

$$MD = [MD_{ij}], \qquad MD_{ij} = 1\ (\text{individual i lives in area j});\ 0\ (\text{otherwise})$$

[98] This derives from the multiplication of the neighbouring MBs' deprivation levels by the individuals' neighbourhood dummy (MD) and the spatial weight matrix (W). $MDEP_j$ is the deprivation of MB 'j':

$$DEP = [ADEP_i], \qquad ADEP_i = \sum_{j=1}^{J} MD_{ij}.\,W_j.\,MDEP_j$$

W is our row standardised spatial weight matrix, which is constructed by taking the k-nearest neighbour approach. Stata's –spwt- command is developed for the purpose of constructing this study's spatial weight matrices. The k-neighbour option of this command is used to derive the spatial weight matrix (W) of the nearest neighbourhoods to an individual's neighbourhood. This matrix does not include a specific number of neighbours; for example, while an area may have 20 neighbours located just on its border, another area may have only one nearest neighbour. This weight matrix is row-standarised by using the 'standardize' option of the –spwt- command.

Table 39. Descriptive statistics of the variables used in chapter 9.[99]

	Mean	Min	Max	Std error
ATVARDENS	10428.4	-2195.97	114966.1	200.447
PPBR	0.902	0.105	2	0.007
$\widehat{PPBR}$	1.028	0.78	1.674	0.002
CVARS	-0.283	-13.828	7.721	0.022
$\widehat{CVARS}$	-0.197	-3.726	2.249	0.011
DENSA	2.535	0.01	7.73	0.018
$\widehat{ATVARDENS}$	10666.98	-11072.6	58836.92	100.772
DEP	990.147	866	1436	1.417
RDEP	-19.922	-423	447	1.86
DUMMY	0.414	0	1	0.008
$\widehat{DEP}$	986.572	873.955	1236.42	0.81
$\widehat{RDEP}$	-23.497	-408.225	318.519	1.555

9.6 Research design

The methodology of this chapter is very similar to that of the previous chapter, in which I used a two-stage least squares method to account for the endogeneity of the variables of interest. These variables are treated as endogenous, since the individual's location and housing choice may be correlated with idiosyncratic factors that affect an individual's residential satisfaction. The same 2SLS approach is used in the current chapter to address the potential endogeneity of the effect of individuals' deprivation levels. It may be that a higher deprivation level would not be undesirable, or would be less undesirable, for some people based on their backgrounds. The idea is that a deprivation level equal to an individual's most likely level of deprivation, estimated based on that person's background characteristics, may be less undesirable. For example, someone from a developing country who has just migrated to a developed country may tolerate higher levels of deprivation than a citizen who has lived in the developed country for his or her entire life.

The variables of interest include both an endogenous effect and an exogenous effect. The exogenous effect can be thought of as a measure of undesirable outcomes (crowding, density and deprivation, respectively). Assuming that I have captured the endogeneous effect precisely, the exogenous (or unexpected) effect is equal to the impact of the variable of interest that is not captured by the endogenous effect. The combined (endogenous plus exogenous) effects will be estimated based on the following equation:

$$y_i = \gamma_0 + \gamma_1 . \hat{x}_i + \gamma_2 . \hat{x}_i^2 + \gamma_3 . (x_i - \hat{x}_i) + \gamma_4 . (x_i^2 - \hat{x}_i^2) + \gamma_5 . \hat{r}_i + \gamma_6 . (r_i - \hat{r}_i) + \gamma_7 . (\text{Control variables})_i + \varepsilon_i \qquad (43)$$

[99] The relative deprivation (RDEP) is constructed based on equation (42). The construction of relative deprivation measure is explained in Section 9.4.2. The construction of other variables of interest are described in Section 8.3.2.

where the dependent variable ($y_{i\times 1}$) is individual i's residential satisfaction (RS). x includes the measures of households' absolute position, DEP, and also a number of variables related to crowding and density, namely household's crowding (PPBR) and density (DENSA). $\hat{x}_{i\times 3}$ includes the expected levels of the measures of households' absolute position (x). $\hat{x}_i^2$ is the squared term of the absolute measures. $r_{i\times 3}$ includes the relative variables of interest, including relative deprivation RDEP, relative density (ATVARDENS) and relative crowding (CVARS). γ_1 to γ_5 are the parameters of interest. r_i includes the measures of the relative crowding, density and deprivation level of an individual. $(r_i - \hat{r}_i)$ is a measure of the level of the relative crowding, density and deprivation level of an individual that is beyond (or below) his or her expected crowding level. γ_6 measures of the impact of the unexpected relative term. ε_i represents the residuals.

The relative term, which is included in $\hat{r}_i$, includes the difference between an individual's deprivation level, and the deprivation level of her adjacent neighbourhoods, both of which are assumed to be endognised based on people's backgrounds. $\hat{r}_i$ also includes similar endogenised relative measures of crowding and density.

Control variables are the same as the variables used in the last chapter[100] and a logit estimation technique is again used. As concluded in Chapter 6, perceived crowding (PC), which is a subjective evaluation of living in a crowded house, accounts for a large proportion of the RS variations. In order to check on the robustness of the results in the described model specification, the PC measure is added to equation (44). Estimations will be done by using resampling weights. For more details, see Section 7.4.1.

9.7 Models and results

In addition to absolute and relative measures of crowding and density, the impacts of which were discussed in Chapters 7 and 8, the present chapter studies the impact of deprivation, both in absolute and relative terms. I have discussed our model specification in Section 9.6.

Table 40 depicts the estimated impact of the expected and unexpected level of the variables of interest on RS, based on equation (44).[101] As discussed in Section 9.6, in order to measure the impact of the difference between the actual and expected levels of the variables of interest, the expected level needs to be controlled for. For example, in order to measure the impact of locating in an area with a deprivation level higher than

[100] For a review of the list of control variables, see Section 7.6.1. Briefly, control variables are as follows: demographics, including gender, ethnicity, age and education; perceptions, including satisfaction with health status and with skills/abilities, which are included to account for any systematic over-optimism of respondents; residential problems, such as housing and neighbourhood problems; the conditions of amenities and facilities; and also social involvement. Also, a number of control variables are based on a comparison between individuals and their neighbourhoods. For example, the proportion of an individual's neighbourhood who have the same ethnicity and a similar occupation to the individual. Also, we have included control variables based on the areas' demographic composition, such as the percentages of each ethnicity in an individual's neighbourhood and the percentage of each occupation category in an individual's neighbourhood.

[101] For the complete set of results, including estimations based on equations (23) and (26), see Appendix II.

the expected level of deprivation ($\mathrm{DEP} - \widehat{\mathrm{DEP}}$), I also need to control for the effect of the expected deprivation ($\widehat{\mathrm{DEP}}$).

The '$(\mathrm{RDEP} - \widehat{\mathrm{RDEP}})\#$ Dummy0' row reports the impact of experiencing relative deprivation higher than the expected level for the residents of the (relatively) rich areas. Holding the expected level of relative deprivation constant, a higher gap between the expected and actual relative deprivation lowers RS significantly. For the residents of the (relatively) poor areas, the effect of relative deprivation higher than the expected level is reported in the '$(\mathrm{RDEP} - \widehat{\mathrm{RDEP}})\#$ Dummy1' row. Holding the expected level of relative deprivation constant, a higher gap between the expected and actual relative deprivation raises RS significantly. However, relative density and relative deprivation are both measures of relative amenity. Based on the joint significance of the measures of relative amenity, unexpectedly being located in a relatively denser and next to a richer (less deprived) neighbourhood raises residential satisfaction (the effect is significant at better than the 10 per cent significance level).[102]

A higher level of unexpected deprivation ($\mathrm{DEP} - \widehat{\mathrm{DEP}}$), together with its squared term, does not affect RS significantly (see the bottom of the table for the joint significance of the unexpected deprivation level and its squared term). A higher unexpected level of crowding ($\mathrm{PPBR} - \widehat{\mathrm{PPBR}}$), together with its squared term, lowers residential satisfaction and a higher unexpected level of density ($\mathrm{DENSA} - \widehat{\mathrm{DENSA}}$), together with its squared term, leads to lower residential satisfaction. Locating in a relatively denser area; that is, an increase in $\mathrm{ATVARDENS} - \mathrm{ATV}\widehat{\mathrm{ARD}}\mathrm{ENS}$, raises RS. However, relative crowding does not affect RS significantly. The results are compatible with those of the previous chapters. Based on AIC, including PC in the estimation improves the GOF.

As presented in the second column of the table, the PC variable is added to the estimation. Similar to the previous chapters, a perception of crowding affects RS negatively. Unlike previous chapters, the inclusion of PC does not materially increase the GOF and does not change the impact of the other variables of interest significantly, which confirms the robustness of the results.[103] Based on the results of the Moran I test, and as presented in Appendix iii, there is no spatial pattern in the distribution of the residuals.

Table 40. Results of estimations - including the impact of deprivation.

Model number	1	2
	Equation (44)	Equation (44)+PC
$\widehat{\mathrm{DEP}}$	0.9925˜	0.9925˜
	(0.046)	(0.046)
$\widehat{\mathrm{DEP}^2}$	1.00002˜	1.00002˜

[102] The p_value of the Wald test is reported at the bottom of the table in the row symbolised by "ϯ".

[103] The results do not change with the presence of the squared term of PC (therefore, I have not assumed a quadratic form for the impact of PC on RS).

	(0.000)	(0.000)
$\mathrm{DEP} - \widehat{\mathrm{DEP}}$	1.001††	1.001††
	(0.001)	(0.001)
$\mathrm{DEP}^2 - \widehat{\mathrm{DEP}}^2$	1.000005††	1.000005††
	(0.000)	(0.000)
$\widehat{\mathrm{RDEP}}$# Dummy1 (Living in poor)	1.00008†	1.00008†
	(0.000)	(0.000)
$\widehat{\mathrm{RDEP}}$# Dummy0 (Living in rich)	0.99998*	0.99998*
	(0.000086)	(0.000086)
$(\mathrm{RDEP} - \widehat{\mathrm{RDEP}})$# Dummy1 (Living in poor)	1.0074*	1.0074*
	(0.003)	(0.003)
$(\mathrm{RDEP} - \widehat{\mathrm{RDEP}})$# Dummy0 (Living in rich)	0.9952**	0.9952**
	(0.002)	(0.002)
$\widehat{\mathrm{PPBR}}$	1.492¡	1.5102¡
	(3.379)	(3.382)
$\widehat{\mathrm{PPBR}}^2$	1.1596¡	1.1596¡
	(0.982)	(0.934)
$\mathrm{PPBR} - \widehat{\mathrm{PPBR}}$	0.241***‡	0.237***‡
	(0.196)	(0.163)
$\mathrm{PPBR}^2 - \widehat{\mathrm{PPBR}}^2$	1.562‡	1.5662‡
	(0.502)	(0.502)
$\widehat{\mathrm{DENSA}}$	0.161¡¡	0.1649¡¡
	(1.328)	(1.328)
$\widehat{\mathrm{DENSA}}^2$	0.959¡¡	0.9567¡¡
	(0.293)	(0.290)
$\mathrm{DENSA} - \widehat{\mathrm{DENSA}}$	0.9351‡‡	0.9352‡‡
	(0.040)	(0.040)
$\mathrm{DENSA}^2 - \widehat{\mathrm{DENSA}}^2$	0.898**‡‡	0.9245*‡‡
	(0.112)	(0.113)
$\widehat{\mathrm{CVARS}}$	0.9102	0.9204
	(1.021)	(1.074)
$\mathrm{CVAR} - \widehat{\mathrm{CVARS}}$	0.992	0.9948
	(0.081)	(0.082)
$\mathrm{ATV}\widehat{\mathrm{ARD}}\mathrm{ENS}$	1.00015	1.00016
	(0.000)	(0.000)
$\mathrm{ATVARDENS} - \mathrm{ATV}\widehat{\mathrm{ARD}}\mathrm{ENS}$	1.00004***†	1.00004***†
	(0.000)	(0.000)
PC		0.2060***
		(0.048)
Observations	4607	4607
Population size	904053	904053
Pseudo-R^2	0.196	0.197
AIC	5032.068	5030.573
Joint significance (Wald test):		

Unexpected relative amenity measures†	<0.1	<0.1
Unexpected deprivation††	Insig.	Insig.
Unexpected crowding measures‡	<0.001	<0.001
Unexpected density measures‡‡	<0.01	<0.01
Expected deprivation measures~	<0.1	<0.1
Expected crowding measures¡	Insig.	Insig.
Expected density measures¡¡	Insig.	Insig.

Reported coefficients are exponentiated ones. Standard errors are reported in parentheses. $^{*}\ p < 0.10$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$

† The significance level for the relative deprivation term is reported on the relative deprivation test row. Note that the reported joint significance p_values are for different coefficients in different models. † The significance level for $(\mathrm{RDEP} - \widehat{\mathrm{RDEP}})\#$ Dummy1 and $\mathrm{ATVARDENS} - \widehat{\mathrm{ATVARDENS}}$ is reported jointly on the unexpected relative amenity measures' Wald test row. †† The significance level for $\mathrm{DEP} - \widehat{\mathrm{DEP}}$ and $\mathrm{DEP}^2 - \widehat{\mathrm{DEP}}^2$ is reported jointly on the unexpected deprivation measures' Wald test row. ‡ The significance level for PPBR-$\widehat{\mathrm{PPBR}}$ and its squared term is reported jointly on the unexpected crowding measures' Wald test row. ‡‡ The significance level for $\mathrm{DENSA} - \widehat{\mathrm{DENSA}}$ and its squared term is reported jointly on the unexpected density measures' Wald test row. ~ The significance level for $\widehat{\mathrm{DEP}}$ and its squared term is reported jointly on the expected deprivation measures' Wald test row. ¡ The significance level for $\widehat{\mathrm{PPBR}}$ and its squared term is reported jointly on the relative crowding measures' Wald test row. ¡¡ The significance level for $\widehat{\mathrm{DENSA}}$ and its squared term is reported jointly on the relative density measures' Wald test row. The reported AICs are for the models that do not take into account jackknife replications. For the detailed estimated results including the impact of control variables, see Appendix IIii.

9.8 Conclusion

How does locating next to others affect us? This question is of particular interest when people locate in a poor area next to a rich area, and vice versa. In this study, as in Chapter 8, a measure of relative deprivation is constructed using the 15-minute walking distance approach. People may intend (or expect) to locate next to rich areas or it may happen unintentionally (or unexpectedly). Therefore, I need to distinguish between a proximity that happens as a result of external factors and one that results from people's choice.

Living in a deprived neighbourhood may be thought of as a factor that lowers the satisfaction of residents. However, after taking a wide range of area characteristics and individual level socioeconomic indicators into account, I do not find any causal impact of the unexpected deprivation level on residential satisfaction.

In assessing the impact of a higher deprivation level, however, the role of people's expectations and also the role of social comparisons need to be taken into account. Therefore, a measure of relative deprivation is constructed based on a 15-minute walking distance from each individual's location. The appropriateness of this neighbourhood boundary definition was tested in the last chapter. At first glance, the impact of the constructed relative measure (RDEP) on residential satisfaction is not statistically significant.[104] However, as has been discussed in the poverty trap literature, people may choose their condition based on their backgrounds. In other words, in an assessment of the impact of relative deprivation on residential

[104] This effect is reported in Appendix II.

satisfaction, it is important to take into account the level of deprivation that may be expected for an individual, based on his or her background.

Similar to the previous chapters, this chapter employs a 2SLS technique to estimate the absolute and relative level of deprivation, crowding and density that an individual is expected to have based on his or her background. The residential satisfaction of an individual who lives in a relatively deprived area may be negatively affected by an envy effect or may be positively affected by the amenities provided by neighbouring amenities in an affluent area.

The results indicate that for those people who choose to live in a relatively poor area, a higher relative deprivation level raises their residential satisfaction, which confirms the presence of a positive amenity effect. People care about the amenities offered by their adjacent neighbourhoods. Relative deprivation and relative density measures jointly capture the impact of the prosperity of the own neighbourhood relative to adjacent neighbourhoods. In addition, the residents of rich areas are negatively affected by bordering on a poor area, which may be interpreted as a NIMBY effect.

I am interested in the impact on residential satisfaction of locating in an area with a higher deprivation than the level of deprivation that an individual is likely to have, which I call unexpected deprivation. A higher level of unexpected deprivation does not affect residential satisfaction significantly. This shows the importance of taking into account the impact of the likely level of deprivation for people with different backgrounds.

Adjacency to affluent areas derived by people's choice affects their residential satisfaction more materially than the adjacencies that happen incidentally. Therefore, people appreciate the affluency of adjacent areas when they aim at locating next to them but unexpected adjacency to affluent areas is less appreciated by the residents. In absolute terms, people are neither affected by a level of deprivation that they were expected to have nor by an absolute deprivation level higher than their expected level of deprivation.

The results of a previous study by Luttmer (2005) suggests that higher levels of neighbours' income are associated with a lower level of own happiness. His study is different in its components of relative measurement and the outcome of interest, as he uses income as a measure of relative position and happiness as his outcome of interest. The difference between the results of his study and the current study may be due to the fact that positive externalities of locating nearby a rich area/person may not be captured by neighbours' personal income levels but by the social benefits associated with a higher income of neighbours.

In terms of the common variables of interest, the results of this chapter are the same as in the previous chapters. In the first essay, I concluded that, amongst the measures of crowding, perceived crowding has the highest predictive power for residential satisfaction. Chapter 6 highlights the importance of perceptions compared to the objective measure of crowding (PPBR). In the current chapter, after taking into account a range of absolute and relative measures of residential positions, adding perceived crowding to equations does not improve the goodness of fit materially, but it nevertheless remains highly significant even when all expected, unexpected and relative variables are included in the regression.

Table 41. Summary of the results of chapter 9.[105]

Category	Result	Significance
Predicted deprivation level derived from backgrounds ($\widehat{\text{DEP}}$)	A higher level of predicted deprivation lowers RS	Yes
Unexpected deprivation ($\text{DEP} - \widehat{\text{DEP}}$)	A higher level of unexpected deprivation does not affect RS	No
Expected relative deprivation ($\widehat{\text{RDEP}}$)	If individual lives in a poor area: a higher relative deprivation level raises her RS	Yes
	If individual live in a rich area: a higher relative deprivation level lowers her RS	Yes
Unexpected relative deprivation ($\text{RDEP} - \widehat{\text{RDEP}}$)	A higher level of unexpected relative deprivation for the residents of poor area raises RS	Yes
	A higher level of unexpected relative deprivation for the residents of rich area lowers RS	Yes
Absolute crowding derived from backgrounds ($\widehat{\text{PPBR}}$)	Higher absolute crowding stemming from 'own group' attraction behaviour does not affect RS	No
Exogenous household crowding ($\text{PPBR} - \widehat{\text{PPBR}}$)	A higher level of unexpected crowding lowers RS	Yes
Expected area density level ($\widehat{\text{DENSA}}$)	Higher area density level stemming from individual's backgrounds does not affect RS	No
Exogenous area density ($\text{DENSA} - \widehat{\text{DENSA}}$)	A higher level of unexpected area density lowers RS	Yes
$\widehat{\text{CVARS}}$	A higher relative expected household crowding does not affect RS	No
Unexpected relative crowding ($\text{CVAR} - \widehat{\text{CVARS}}$)	A higher level of unexpected relative crowding does not affect RS	No
Expected relative density ($\widehat{\text{ATVARDENS}}$)	Living in an area with a relatively higher expected density does not affect RS	No
Unexpected relative density ($\text{ATVARDENS} - \widehat{\text{ATVARDENS}}$)	A higher level of unexpected relative density raises RS	Yes

[105] This table includes the results presented in section 9.7 and in Appendix II.

Perceived crowding (PC)	Higher PC level affects RS negatively	Yes

This table summarises the estimation results reported in Table 40.

The results of this study improve our understanding of desirable urban design in terms of maximising the satisfaction of residents. A question may arise in regard to the best policy to derive the desired neighbourhood design. Four types of shocks may lead to a change in neighbourhoods, namely policy innovations, such as regeneration schemes; exogenous innovations, such as earthquakes; technological innovations; and endogenous changes, such as migration (Meen et al., 2013). However, as Posthumus, Bolt, and van Kempen (2013) mention, an attempt to create social mixed areas by policy interventions in the lower income neighbourhoods leads to migration of a part of their population to other areas. The displaced population tends to settle into a neighbourhood in which their peers are located in[106]. This means the policy may not reach its aim of shaping neighbourhoods with desired characteristics.

Consumption of space may need to be considered over an individual's life cycle. An individual may not find her current crowding level dissatisfactory as she is temporarily living in the current situation. This especially happens if an individual thinks of accumulating wealth in the current period to use it as a collateral for her future consumption of housing.

Summary

The residents of poor areas are likely to adapt to their own living conditions and, therefore, may not be dissatisfied with living in a poorer area. However, they may benefit from locating near the amenities offered in a richer area. To analyse such effects, a relative deprivation measure needs to be constructed.

The most appropriate relative crowding measure is the one constructed based on a comparison between an individual's position and the position of people living in a 15-minute walking distance from that individual's location. The relative deprivation measure is constructed in the same way, based on a comparison between the deprivation level of an individual and the deprivation level of the areas that the individual may visit in a 15-minute walk around his or her location.

After taking account of endogenous location choice, both with respect to the level of deprivation of the individual's own neighbourhood and the level of deprivation of his or her adjacent neighbourhoods, I find that an increase in expected level of own deprivation lowers RS. However, locating in an area with a higher deprivation level than the deprivation level that an individual is expected to experience does not affect that person's residential

[106] This is confirmed by our results that people do not care about their absolute level of deprivation after accounting for their location choice and their relative deprivation.

satisfaction. Locating next to an affluent area raises the satisfaction of those people in the poor areas. Therefore, residents of a poor area enjoy a positive amenity effect when they choose to locate adjacent to a rich area. However, the residents of the rich area do not enjoy living near poor areas, which can be interpreted as a negative NIMBY effect.

Appendix I. Persistence of poor conditions

As discussed in 9.4.2, the deprivation level of an individual's neighbourhood is attributed to her. A model presented and solved by Blume and Durlauf (2001) discusses the neighbourhood aggregates derived by peer group effects. When an individual is less satisfied with her residential environment, she is more likely to move out, subject to her budget constraint. In a simplified model, which accounts for the interdependence of neighbours' decisions, I can assume the joint probability of moving out versus staying in the neighbourhood as,

$$\Pr(w_i|C_i, Z) = \prod_{i=1}^{I} \Pr(w_i|C_i, Z)$$

where, w_i represents the movement choice of individual i; C_i is the characteristics of the individual i, Z is a set of endogenous variables that may relate to the neighbourhood. If I assume a quadratic residential satisfaction function such that the function depends on the individual i's payoff which is affected by others' payoff (w_{-i}) and also by the square of the difference between the payoff of i and others, the residential satisfaction function becomes:

$$\mathbf{RS}(\mathbf{w_i}|\mathbf{C_i}) = \mathbf{I}(\mathbf{C_i})\mathbf{w_i} + \boldsymbol{\tau}_\mathbf{i}(\mathbf{w_i}) - \frac{\mathbf{1}}{\mathbf{2}}\mathbf{E}(\mathbf{S}(\mathbf{i}, -\mathbf{i})(\mathbf{w_i} - \mathbf{w_{-i}})^2) \tag{44}$$

In this equation, RS is residential satisfaction, $I(C_i)$ is the individual i's incentive to move, $I(C_i)w_i$ is the mean of the individuals' residential satisfaction, $\tau_i(w_i)$ is the random deviation from the mean and S(i,-i) represents the peer's effect derived from the individual i's movement choice versus the others' choice. The first part, $I(w_i) + \tau_i(w_i)$, can be considered as the individual component and the next part, $-\frac{1}{2}E(S(i, -i)(w_i - w_{-i})^2)$, as the social interaction component. In this study, I have used a very similar function for our estimations, as discussed in section 4. I assume a logistic distribution function for the probability of residential satisfaction (movement choice), i.e. residential satisfaction is assumed to be a dummy variable. Thus the probability function is,

$$\Pr(\tau_i(-1) - \tau_i(1) < \delta) = \frac{1}{1 + e^{-\beta(S_i)\delta}}$$

Here, by using the logistic distribution, an independent random distribution is assumed, the social interactions are omitted from the model. If I assume that individuals interact one another symmetrically based on their incentives (I), then the peer's effect for individual i derives from the sample's mean,

$$S(C_i, C_{-i}) = \frac{S}{N - 1}, N = \sum i$$

If I assume that individuals are affected by the average of their neighbourhoods, they should solve the following function, which differs from equation (45) in its social interaction component.

$$\mathbf{Max\ I(C_i)w_i + \tau_i(w_i) + Sw_iE_i(a)} \quad (45)$$

The average choice of the neighbourhood is represented by a, which is affected by the heterogeneity amongst neighbours (β) and also the impact of social interactions (S). The solution to the maximisation problem shows how peer group's effects induce neighbourhood aggregates. The solution to the maximisation problem (equation (46)) indicates that only if the individuals' incentives to move are weak, i.e. if $I(C_i)$ is small, stronger social interactions (measured by S) may lead to an undesirable social outcome, but a desirable individual behaviour. In other words, if people are not incentivized enough to move from a poor neighbourhood (not residentially dissatisfied), interacting with people with similar characteristics to themselves (neighbours) may result in a desire to stay in a poor neighbourhood (derived from a higher residential satisfaction), while this is not desirable for the society.

Appendix II. Table of results

I estimate the following equation as our baseline regression,

$$y_i = \alpha_0 + \alpha_1 . x_i + \alpha_2 . x_i^2 + \alpha_3 . r_i + \beta . (\text{Control variables})_i + \varepsilon_i \quad (46)$$

Where, the dependent variable ($y_{i\times1}$) is individual i's residential satisfaction (RS). $x_{i\times3}$ includes the measures of households' absolute position, DEP, and also a number of variables related to crowding and density, namely household's crowding (PPBR) and density (DENSA)[107]. x_i^2 is the squared term of the absolute measures with the same sign as x_i, i.e. if x_i is negative x_i^2 is negative as well. $.r_{i\times3}$ includes the relative variables of interest, including relative deprivation RDEP, relative density (ATVARDENS) and relative crowding (CVARS). α_1, α_2 and α_3 are the parameters of interest and ε_i represents the residuals.

The estimated results of the first model specification, based on equation (47), are presented in the first column of Table 42. Accordingly, the impact of RDEP on RS is statistically insignificant. The next variables demonstrated in the first column are the absolute deprivation level (DEP) and its squared term (DEP²). Due to the assumed quadratic form, the effect of absolute deprivation needs to be assessed jointly with its squared term. The p_value for the joint significance test is reported at the bottom of the table in the row symbolised by '~'. Accordingly, a higher level of deprivation does not affect RS significantly.

---

[107] The variables of interest are $i \times 1$ individual characteristics. Except for PPBR, which is an i by 1 matrix of households' crowding, the other variables of interest are derived from merging the individual dataset with i observations with the area dataset with j areas. This is done by using the neighbourhood dummy variable (MD), as follows,

$$DEP_{i\times1} = MD_{i\times j} . DEPM_{j\times1}$$

$$DENS_{i\times1} = MD_{i\times j} . DENSA_{j\times1}$$

Next I am interested in the impact of PPBR and its squared term. These two are jointly significant with better than the 1 per cent significance level. According to the descriptive statistics, presented in chapter 7's Appendix i, PPBR has a minimum of 0.105 and a maximum of 2 (for detailed description of PPBR see 6.5). The impact of a higher crowding is negative at any crowding level. However, this negative impact is lower when the level of crowding is high. The next variable of interest is density (DENSA). Both density and its squared term have a negative effect on RS. Their effect is jointly significant at better than the 1 per cent significance level.

The other two main variables of interest are relative crowding (CVARS) and relative density (ATVARDENS). Similar to the previous chapter, higher relative crowding does not affect RS significantly. Living in a suburb denser than adjacent areas (of the same TLA) is associated with higher RS. To have an understanding of the impact of an increase in density, due to opposite effects of the relative density (positive) and absolute density (negative), the joint effect of them needs to be assessed. The marginal effect of living in a relatively dense area is significant at the 10 per cent significance level. This impact is positive in the lower extreme of density, i.e. living in relatively a denser area raises RS when an individual lives in an area with a low density level. However, the marginal effect of living in a relatively dense area is negative when an individual level lives in a dense area.

The second column is similar to the first column, i.e. it follows the same model specification, but the relative deprivation's effect is split into two by interacting it with the Dummy variable (equation (43)). The impact reported here is for the case where the dummy is equal to one, i.e. when an individual lives in an area poorer than her adjacent neighbourhoods. As depicted, the effect of relative deprivation on RS is not statistically significant (the p-value for the statistical significance test is equal to 0.1042).

The endogenous variables for individual i (including both x_i and r_i) are regressed on the list of instruments (z_i) and the list of control variables. The first stage of 2SLS is as follows,

$$x_i = \theta_0 + \theta_1 . z_i + \theta_2 . (\text{Control variables})_i + u_i \tag{47}$$

The list of instruments is as in previous chapters. The validity of the instruments is confirmed by using the over-identification tests, reported with the estimates. In the next step, the dependent variable would be regressed on the likelihoods of deprivation, density and crowding ($\hat{x}_i$). The second specification is therefore as follows,

$$y_{it} = \beta_0 + \beta_1 . \hat{x}_i + \beta_2 . \hat{x}_i^2 + \beta_3 . (\text{Control variables})_i + \varepsilon_i \tag{48}$$

The relative term, which is included in $\hat{r}_i$, is the difference between an individual's deprivation level, and the deprivation level of her adjacent neighbourhoods, both of which are assumed to be endognised based on people's backgrounds. Hence, the endogenous variables for individual i (including both x_i and r_i) are regressed on the list of instruments (z_i) and the list of control variables. The first stage of 2SLS is as follows,

$$x_i = \theta_0 + \theta_1 . z_i + \theta_2 . (\text{Control variables})_i + u_i \quad (49)$$

The list of instruments is as in previous chapters. The validity of the instruments is confirmed by using the over-identification tests, reported with the estimates. In the next step, the dependent variable would be regressed on the deprivation likelihoods ($\hat{x}_i$). The second specification is therefore as follows,

$$y_{it} = \beta_0 + \beta_1 . \hat{x}_i + \beta_2 . \hat{x}_i^2 + \beta_3 . (\text{Control variables})_i + \varepsilon_i \quad (50)$$

The second model specification, presented in column 3, is an estimation of equation (51), which takes into account the endogeneity of the variables of interest. The variables of interest are estimated based on equation (50) and named as $\widehat{DEP}, \widehat{RDEP}, \widehat{PPBR}, \widehat{DENS}, \widehat{CVARS}$ and $\widehat{ATVARDENS}$. The validity of the instruments are confirmed by using the Sargan test[108]. With a high confidence I confirm the null hypothesis that the over-identifying restrictions are valid. As in the first model specification, the relationship between the estimated absolute measures and RS is assumed to be quadratic and the relationship between the estimated relative measures and RS is linear. A higher expected deprivation ($\widehat{DEP}$), which is based on the probability of living in a deprived area, does not affect RS significantly (the joint significance of $\widehat{DEP}$ and its squared term are reported at the bottom of the table in the row symbolised by '~'). This suggests that a higher level of expected deprivation level, i.e. the level of deprivation that an individual is likely to have based on her characteristics, does not affect the satisfaction with residential environment.

The other variable of interest included in column 3 is $\widehat{RDEP}$, which is equal to the expected value of $(DEP - W.DEP)$, based on equation (42). The coefficient of the expected relative term ($\widehat{RDEP}$) is reported after interacting with the Dummy variable (equation (43)). I report the impact of $\widehat{RDEP}$ for those who live in a poor area compared to their adjacent neighbourhoods ($\widehat{RDEP}$#Dummy1). As depicted in the third column, for those who live in a poor area, an increase in the relative term raises RS, i.e. people living in poorer areas like to locate next to rich areas. The estimated coefficients for the expected relative term is significant at 5 per cent level (with a p_value equal to 0.0383) and the marginal effect of it is positive. For people living in relatively an affluent area (compared to their neighbourhoods' adjacent areas ($\widehat{RDEP}$# Dummy0)), the estimated effect for them is significantly less than one, i.e. the residents of rich areas are not happy with neighbouring a poor area, suggesting the presence of a NIMBY[109] effect.

Columns 4 and 5 present the result of the estimations depicted in section 9.7, which is an estimation of equation (44) without and with including PC in the equation, respectively.

Table 42. Results table including control variables.

	(1)	(2)	(3)	(4)	(5)

[108] In the third model, the null hypothesis of the Stock and Yogo test that the set of instruments is weak is rejected at a significance level less than 5 per cent.

[109] This is an abbreviation of "Not In My Back Yard".

	Specification I	Specification I + Dummy	Specification II	Specification III	Specification III + PC
RDEP	0.999210750				
	(0.001)				
RDEP# Dummy1 (Living in poor)		1.002274925			
		(0.002)			
$\widehat{\text{RDEP}}$# Dummy0 (Living in rich)			0.9953**	0.9999845*	0.9999822*
			(0.002)	(0.000086)	(0.000086)
$\widehat{\text{RDEP}}$# Dummy1 (Living in poor)			1.0073758**	1.000080†	1.000082†
			(0.003)	(0.000)	(0.000)
$(\text{RDEP} - \widehat{\text{RDEP}})$# Dummy0 (Living in rich)				0.9952**	0.9952**
				(0.002)	(0.002)
$(\text{RDEP} - \widehat{\text{RDEP}})$# Dummy1 (Living in poor)				1.0074†	1.0074†
				(0.003)	(0.003)
DEP	0.996552288~	0.992204561~			
	(0.011)	(0.009)			
DEP^2	1.000001541~	1.000003727~			
	(0.000)	(0.000)			
$\widehat{\text{DEP}}$			0.975466103~	0.9925411~	0.9925308~
			(0.075)	(0.046)	(0.046)
$\widehat{\text{DEP}^2}$			1.000023820~	1.000015~	1.000016~
			(0.000)	(0.000)	(0.000)
$\text{DEP} - \widehat{\text{DEP}}$				1.001422††	1.001379††
				(0.001)	(0.001)
$\text{DEP}^2 - \widehat{\text{DEP}}^2$				1.000005††	1.000005††
				(0.000)	(0.000)
PPBR	0.165624057***¡	0.166006265***¡			
	(0.111)	(0.112)			
PPBR^2	1.589345763¡	1.576655208¡			
	(0.438)	(0.491)			
$\widehat{\text{PPBR}}$			1.175497¡	1.492¡	1.51025¡
			(2.695)	(3.379)	(3.382)
$\widehat{\text{PPBR}}^2$			0.6204362¡	1.1596¡	1.159617¡
			(0.490)	(0.982)	(0.934)
$\text{PPBR} - \widehat{\text{PPBR}}$				0.2414961***‡	0.2370468***‡
				(0.196)	(0.163)
$\text{PPBR}^2 - \widehat{\text{PPBR}}^2$				1.5623778‡	1.566202‡
				(0.502)	(0.502)
DENSA	0.936852088¡¡	0.939722358¡¡			
	(0.117)	(0.118)			
DENSA^2	0.933865185¡¡	0.934853382¡¡			
	(0.041)	(0.041)			
$\widehat{\text{DENSA}}$			0.542358381¡¡	0.1614049¡¡	0.1649556¡¡
			(1.247)	(1.328)	(1.328)
$\widehat{\text{DENSA}}^2$			0.845576228¡¡	0.9591839¡¡	0.9567594¡¡
			(0.236)	(0.293)	(0.290)

$DENSA - \widehat{DENSA}$				0.9351022‡‡	0.9352593‡‡
				(0.040)	(0.040)
$DENSA^2 - \widehat{DENSA}^2$				0.8987416**‡‡	0.924532*‡‡
				(0.112)	(0.113)
CVARS	0.990381262	0.995179466			
	(0.083)	(0.083)			
$\widehat{CVARS}$			1.379778447	0.9102137	0.9204471
			(1.940)	(1.021)	(1.074)
$CVAR - \widehat{CVARS}$				0.9923612	0.9948473
				(0.081)	(0.082)
ATVARDENS	1.000042096**	1.0000410**			
	(0.000)	(0.000)			
$ATV\widehat{ARD}ENS$			1.000071497*	1.000152*	1.000159
			(0.000)	(0.000)	(0.000)
$ATVARDENS - ATV\widehat{ARD}ENS$				1.000042***†	1.000043***†
				(0.000)	(0.000)
PC					0.2060302***
					(0.048)
MOCCN_1	3.410164237	3.674735836	8.648724842	12. 5176635	12.576
	(2.928)	(3.220)	(19.286)	(24.875)	(24.851)
MOCCN_2	1.221839263	1.219801519	1.747462569	1. 3863421	1.388
	(0.686)	(0.678)	(1.705)	(1.182)	(1.188)
MOCCN_3	0.709014782	0.741765248	1.029438794	0.84773521	0.778
	(0.856)	(0.888)	(1.974)	(1.311)	(1.379)
MOCCN_4	1.350795370	2.055492095	92.637146220	36.365841	34.618
	(2.950)	(4.535)	(381.881)	(160.703)	(142.893)
MOCCN_5	0.378454115	0.413059257	0.391415132	0.7375521	0.703
	(0.569)	(0.614)	(0.607)	(1.114)	(1.095)
MOCCN_6	3.643730817	4.109055216	3.870767188	5.19554758	5.192
	(5.862)	(6.611)	(6.116)	(8.414)	(8.391)
MOCCN_7	9.847091702	8.046827316	46.042045617	49.2146347	49.222
	(53.262)	(44.573)	(264.188)	(271.540)	(283.606)
MOCCN_8	1.751766244	1.687394651	1.082361243	0.6675157	0.686
	(5.614)	(5.198)	(3.302)	(2.326)	(2.287)
MOCCN_9	1.69039e+02	3.39764e+02	19.913601831	19.6371744	20.431
	(1093.879)	(2487.054)	(150.929)	(144.850)	(149.875)
MOCCT_1	1.479076732	1.915557140	0.309726779	0.246715323	0.235
	(1.379)	(1.812)	(1.101)	(0.854)	(0.721)
MOCCT_2	0.934460178	1.040404998	0.039383640	0.026855557	0.029
	(0.599)	(0.695)	(0.255)	(0.171)	(0.162)
MOCCT_3	0.075355908**	0.081922776**	0.340639601	0.416631704	0.397
	(0.083)	(0.088)	(1.241)	(1.504)	(1.246)
MOCCT_4	0.263582210	0.437580694	0.082153556	0.054987380	0.033

	(0.321)	(0.508)	(0.607)	(0.402)	(0.205)
MOCCT_5	0.242848698	0.152909392*	1.108386369	1.269057542	0.739
	(0.258)	(0.170)	(3.169)	(3.617)	(1.819)
MOCCT_6	0.178231459	0.161291378	0.021995863	0.016651972	0.021
	(0.218)	(0.201)	(0.077)	(0.058)	(0.068)
MOCCT_7	1.203463308	0.735584395	5.577821751	2.2546551	2.278
	(2.044)	(1.389)	(12.815)	(5.778)	(5.196)
MOCCT_8	0.973268854	1.275273153	3.482718550	2.01553089	2.191
	(1.092)	(1.368)	(4.750)	(3.154)	(3.117)
MOCCT_9	1.261380565	2.630551031	0.135950391	0.116653506	0.196
	(1.683)	(3.252)	(0.772)	(0.654)	(0.938)
METHN_1	1.220539620	1.444318557	1.078971278	1.212926276	1.461
	(0.889)	(1.019)	(1.398)	(1.634)	(1.740)
METHN_2	3.028774829	2.585946660	0.694869142	0.711777598	0.690
	(4.404)	(3.882)	(1.945)	(1.973)	(1.848)
METHN_3	5.203885781*	5.763385927**	3.247557517	2.522953549	1.860
	(4.375)	(4.601)	(7.840)	(6.108)	(3.838)
METHN_4	2.152724831	1.793929724	3.491256247	2.43139087	2.509
	(2.109)	(1.816)	(3.448)	(2.334)	(2.624)
METHN_5	0.137626183	0.054363662	0.694811705	0.11258042	0.112
	(0.607)	(0.264)	(3.414)	(0.681)	(0.547)
METHT_1	9.200486370*	5.329809073	7.109964272	12.433284814	21.295
	(11.220)	(6.068)	(26.268)	(45.765)	(65.256)
METHT_2	1.281633607	0.881457852	0.093633324	0.097929728	0.137
	(1.272)	(0.853)	(0.301)	(0.301)	(0.399)
METHT_3	1.815758763	1.581911693	0.162432146	0.191789652	0.339
	(2.089)	(1.717)	(0.515)	(0.592)	(0.980)
METHT_4	7.064521945	4.826997311	0.283470111	0.391184080	1.152
	(8.524)	(5.490)	(0.699)	(0.927)	(2.997)
METHT_5	19.943288135**	14.611192288**	8.951284804	15.446963658	18.299*
	(27.853)	(19.687)	(15.579)	(27.250)	(30.408)
Gender	0.965750553	0.957537243	0.882262474	0.903470699	0.904
	(0.098)	(0.097)	(0.135)	(0.136)	(0.125)
Partner	1.138204277	1.139099586	0.365346614	0.352550863	0.337*
	(0.154)	(0.155)	(0.243)	(0.244)	(0.201)
Age (>74)	1.192546241	1.201082976	0.110611160	0.084900846	0.071*
= 15-19	(0.498)	(0.505)	(0.191)	(0.150)	(0.111)
= 20-24	0.808530574	0.808777995	0.151913629	0.123904228*	0.114*
	(0.248)	(0.250)	(0.181)	(0.154)	(0.136)
= 25-34	0.477597127***	0.476666687***	0.129274585**	0.109570984**	0.107**
	(0.133)	(0.133)	(0.115)	(0.099)	(0.095)
= 35-54	0.484247690***	0.480511943***	0.182954165**	0.160624247**	0.146**
	(0.118)	(0.118)	(0.130)	(0.117)	(0.108)

= 55-59	0.748728221	0.746396222	0.931494597	0.940660076	0.947
	(0.188)	(0.189)	(0.552)	(0.554)	(0.420)
= 60-64	0.873666114	0.851241603	1.076669408	1.099985690	1.023
	(0.246)	(0.242)	(0.517)	(0.533)	(0.441)
= 65-69	0.845689561	0.825914482	1.141801595	1.196622877	1.162
	(0.193)	(0.188)	(0.585)	(0.619)	(0.520)
= 70-74	0.834850733	0.810589495	0.877442025	0.878215776	0.808
	(0.198)	(0.192)	(0.277)	(0.276)	(0.257)
Length of stay in NZ (<4 years)					
= 4-10 years	1.042774399	1.032457518	1.247973047	1.286232899	(0.577)
	(0.344)	(0.338)	(0.608)	(0.645)	2.361
= 10-25 years	0.848262621	0.838891102	2.159906008	2.347358429	(1.963)
	(0.262)	(0.260)	(2.281)	(2.502)	2.827
>25 years	1.207618168	1.217178342	2.384445263	2.678709185	(1.878)
	(0.377)	(0.380)	(1.949)	(2.258)	1.000
Ethnicity (European)					
= Maori	1.110374559	1.273987219	0.222795720	0.208013390	0.207
	(0.631)	(0.710)	(0.218)	(0.208)	(0.208)
= Pacific	0.746773827	0.763763108	0.043243629**	0.034928744**	0.041*
	(0.387)	(0.388)	(0.068)	(0.056)	(0.066)
= Asian	0.654380474	0.762868426	0.340234141	0.327664731	0.373
	(0.373)	(0.429)	(0.341)	(0.337)	(0.372)
= Other	1.598913555	1.966320747	0.496247431	0.453554411	0.594
	(1.031)	(1.315)	(0.471)	(0.432)	(0.552)
Education (No qualification)					
= Diploma	0.978923692	0.997171766	3.231471622	3.530300075	3.668**
	(0.204)	(0.206)	(2.502)	(2.750)	(2.279)
= Degree	0.885730520	0.915769684	3.237512299	3.491400157	3.828**
	(0.191)	(0.196)	(2.636)	(2.875)	(2.338)
Household income ($150,001 or more)					
= Zero	0.828283897	0.813275873	1.268478318	1.349187706	1.895
	(0.430)	(0.437)	(1.819)	(1.941)	(2.82)
=$1-$5,000	2.839625878	2.805208152	2.839625878	0.259431061	0.875
	(3.216)	(2.765)	(0.555)	(0.515)	(2.61)
=$5000 - $50,001	0.802812970	0.799815011	0.791483932	0.408662677	0.429
	(0.126)	(0.126)	(0.431)	(0.440)	(0.355)
=$50,001 - $70,000	0.863328591	0.862965123	0.846791890	0.491310574	0.499
	(0.158)	(0.157)	(0.404)	(0.443)	(0.279)
=$70,001 - $150,000	0.869860567	0.866503925	0.855390318	0.565982386	0.574
	(0.119)	(0.117)	(0.314)	(0.317)	(0.242)
Health (Very dissatisfied)					
=Dissatisfied	2.447508184**	2.425723927**	4.770655318***	4.57542701***	4.651***
	(0.926)	(0.910)	(2.121)	(2.250)	(2.049)
=No feeling either way	2.936832196***	2.856962467***	7.083878138***	7.35785510***	7.244***

	(0.981)	(0.932)	(3.506)	(3.465)	(3.538)
=Satisfied	3.360373558***	3.311924001***	9.292342811***	10.1302724***	9.865***
	(1.122)	(1.079)	(5.220)	(5.174)	(5.189)
=Very Satisfied	4.347873658***	4.266820745***	11.544419537***	13.7520714***	12.479***
	(1.528)	(1.467)	(6.560)	(7.101)	(6.719)
Abilities (Very dissatisfied)					
=Dissatisfied	1.107911506	1.001402725	0.131920770	0.072640699	0.065
	(1.668)	(1.520)	(0.314)	(0.161)	(0.131)
=No feeling either way	1.307905871	1.194500920	0.344012640	0.17265546	0.174
	(1.910)	(1.754)	(0.639)	(0.275)	(0.271)
=Satisfied	1.729309215	1.601447818	0.585419328	0.320019871	0.289
	(2.519)	(2.347)	(1.033)	(0.498)	(0.429)
=Very Satisfied	2.769593679	2.518810863	1.082913793	0.607969387	0.554
	(4.071)	(3.725)	(1.925)	(0.946)	(0.827)
Bad street access	0.569366603	0.550968085	0.721914225	0.839849658	0.883
	(0.298)	(0.283)	(0.493)	(0.555)	(0.578)
Poor condition	0.319523499***	0.314773428***	0.252927960***	0.263827218***	0.272***
	(0.100)	(0.099)	(0.077)	(0.086)	(0.086)
Damp dwelling	0.539240774***	0.552420155***	0.621393432	0.795240736	0.703
	(0.109)	(0.113)	(0.180)	(0.164)	(0.162)
Difficult to heat	0.458487684***	0.458584660***	0.422640129**	0.421317665**	0.420**
	(0.089)	(0.090)	(0.139)	(0.142)	(0.143)
Having pests	1.381170123	1.426879389	0.928734804	0.845067655	0.854
	(0.386)	(0.399)	(0.361)	(0.342)	(0.355)
Expensive house	0.582306178**	0.590771294**	0.646079031	0.625575417	0.665
	(0.143)	(0.147)	(0.228)	(0.219)	(0.212)
Work distance	1.078305868	1.059964730	1.026516213	1.149958965	1.106
	(0.266)	(0.259)	(0.359)	(0.407)	(0.388)
Far from facilities	0.593158713	0.569423098	0.374052487	0.371563423	0.290
	(0.253)	(0.246)	(0.367)	(0.362)	(0.218)
Unsafe neighbourhood	0.800566683	0.762342803	0.344993067	0.348230894	0.386
	(0.233)	(0.229)	(0.264)	(0.265)	(0.288)
Noise and vibration	0.615082034***	0.604488522***	0.232761350	0.232854660	0.237*
	(0.101)	(0.097)	(0.229)	(0.224)	(0.182)
Air polluted	0.991693656	0.977804044	0.645820222	0.654822117	0.680
	(0.250)	(0.252)	(0.337)	(0.347)	(0.319)
Tenure status (Owner)					
= Renter	0.663684570***	0.659231450***	0.192831085**	0.183503645**	0.185***
	(0.081)	(0.081)	(0.149)	(0.143)	(0.115)
= Family trust	1.161741386	1.161690659	1.744660108	1.79214752**	1.797**
	(0.133)	(0.138)	(0.628)	(0.423)	(0.451)
Free Time (Not enough free time)					

= The right amount of free time	1.268267599**	1.271795973**	1.280079408**	1.1988525	1.194
	(0.127)	(0.127)	(0.155)	(0.147)	(0.132)
= Too much free time	1.229120972	1.241957294	1.338820473	1.199080	1.199
	(0.247)	(0.247)	(0.272)	(0.258)	(0.251)
Socialising (Every day)					
= At least once in The last four weeks	0.878340811	0.867022307	0.706511425	0.6511855	0.683
	(0.152)	(0.148)	(0.190)	(0.174)	(0.180)
= Around once a fortnight	0.829344328	0.838795059	0.962363424	0.975940544	0.976
	(0.116)	(0.117)	(0.366)	(0.358)	(0.340)
= Around 1-6 times a week	0.910256050	0.916176858	1.181398332	1.181642959	1.215
	(0.090)	(0.091)	(0.280)	(0.280)	(0.215)
Local political involvement	0.953736219	0.958025341	0.934692153	0.955205656	0.957
	(0.106)	(0.106)	(0.144)	(0.154)	(0.147)
Recycle==A little	2.157780826	2.047105962	0.454501524	0.387117315	0.334
	(1.335)	(1.211)	(0.746)	(0.627)	(0.475)
Recycle== Some	0.784821911	0.769877047	0.607682764	0.5277852	0.547
	(0.285)	(0.282)	(0.378)	(0.244)	(0.256)
Recycle== Most	0.591061066***	0.603884500***	0.484861286**	0.47510582***	0.475***
	(0.109)	(0.111)	(0.142)	(0.131)	(0.130)
Water use (Never)					
= a little of The time	1.107664995	1.102930555	1.845822547	1.831317942	1.922**
	(0.155)	(0.154)	(0.794)	(0.786)	(0.625)
= Some of The time	1.091404314	1.121023188	1.614880819	1.528100948	1.642
	(0.155)	(0.159)	(0.666)	(0.624)	(0.521)
= Most of The time	0.970748990	0.983651728	1.220999748	1.2052749	1.258
	(0.105)	(0.107)	(0.408)	(0.324)	(0.322)
Energy use (Never)					
= a little of The time	0.654445764**	0.635207109**	0.495161590*	0.483145254**	0.486*
	(0.119)	(0.114)	(0.197)	(0.176)	(0.188)
= Most of The time	0.839120934	0.845126686	0.750440329	0.753210038	0.752
	(0.089)	(0.091)	(0.145)	(0.147)	(0.147)
Council services (Very satisfied)					
= Very Dissatisfied	0.925023150	0.888191192	0.786698072	0. 51024736	0.508
	(0.360)	(0.346)	(1.343)	(0.744)	(0.745)
= Dissatisfied	0.515787504***	0.503614027***	0.611652923	0.660762163	0.543
	(0.090)	(0.090)	(0.433)	(0.477)	(0.350)
= No feeling either way	0.408027765***	0.400285625***	0.581530533	0.624419777	0.566
	(0.069)	(0.069)	(0.271)	(0.292)	(0.243)
= Satisfied	0.447747374***	0.431819530***	0.446460538**	0.4215445**	0.427***
	(0.056)	(0.056)	(0.143)	(0.103)	(0.105)
Facilities access (All of them)					
= Never been to	0.025519605	0.024726423	0.041410846	0.081246187	0.057

	(0.101)	(0.098)	(0.179)	(0.347)	(0.228)
= a few of them	0.226045943	0.236040628	0.662189638	0.65250	0.652
	(0.338)	(0.365)	(1.256)	(1.229)	(1.143)
= None of them	0.858504003	0.824839297	0.403477548	0.415216225	0.372
	(0.430)	(0.426)	(0.287)	(0.296)	(0.222)
Green space access (Few of them)					
= Most of them	0.437441149	0.430856981	0.088032745**	0.087343782*	0.090**
	(0.362)	(0.339)	(0.107)	(0.108)	(0.100)
= All of them	0.999235450	1.002370741	0.804974332	0.775112627	0.833
	(0.285)	(0.292)	(0.299)	(0.295)	(0.302)
Green space state (Very satisfied)					
= Not been to	0.274195988*	0.273465089*	0.644104474	0.654141837	0.492
	(0.202)	(0.198)	(0.648)	(0.656)	(0.501)
= Very Dissatisfied	1.540258308	1.481153660	0.537476569	0.354912782	0.475
	(2.347)	(2.282)	(1.023)	(0.664)	(0.892)
= Dissatisfied	0.882273830	0.859179017	0.802681603	0.813129388	0.765
	(0.273)	(0.264)	(0.283)	(0.290)	(0.253)
= No feeling either way	1.037735358	1.025417088	0.759296344	0.726718321	0.728
	(0.242)	(0.242)	(0.368)	(0.351)	(0.325)
Coastline access (Few of them)					
= Most of them	2.092311863	2.065094543	2.171281989	2.17671892	2.105
	(1.133)	(1.106)	(1.491)	(1.364)	(1.295)
= All of them	1.149109907	1.122285442	0.638860083	0.589552394	0.562
	(0.324)	(0.319)	(0.438)	(0.406)	(0.358)
Coastline state (Very satisfied)					
= Not been to	2.060697400	2.049666683	2.091504767	1.865094217	1.948
	(1.552)	(1.501)	(1.724)	(1.551)	(1.610)
= Very Dissatisfied	1.536216897	1.645792847	2.855426444	3.131174180	3.370
	(1.112)	(1.213)	(2.933)	(3.176)	(3.404)
= Dissatisfied	0.999746727	1.006901009	1.470758780	1.541446475	1.504
	(0.199)	(0.199)	(0.577)	(0.606)	(0.505)
= Satisfied	0.889336775	0.874726397	1.071206841	1.092962331	1.080
	(0.158)	(0.158)	(0.388)	(0.391)	(0.372)
Household size (>4 people)					
= Single person	0.382583093***	0.370635072***	0.814270758	0.4962511**	0.497**
	(0.128)	(0.124)	(0.166)	(0.181)	(0.164)
= 2 people	0.531414224**	0.524396556**	0.875366585	0.6148554*	0.627*
	(0.134)	(0.136)	(0.155)	(0.158)	(0.162)
= 3 people	0.580089479***	0.582084411**	0.760961179	0.7245563	0.658*
	(0.120)	(0.122)	(0.133)	(0.139)	(0.140)
= 4 people	0.720688767	0.728483815	0.809535716	0.7714420	0.784
	(0.152)	(0.152)	(0.154)	(0.168)	(0.165)
Constant	8.381121785	4.3150232357	0.425805695	0.000211	0.0001
	(4.5150)	(2.28043)	(1.6405)	(0.003)	(0.004)

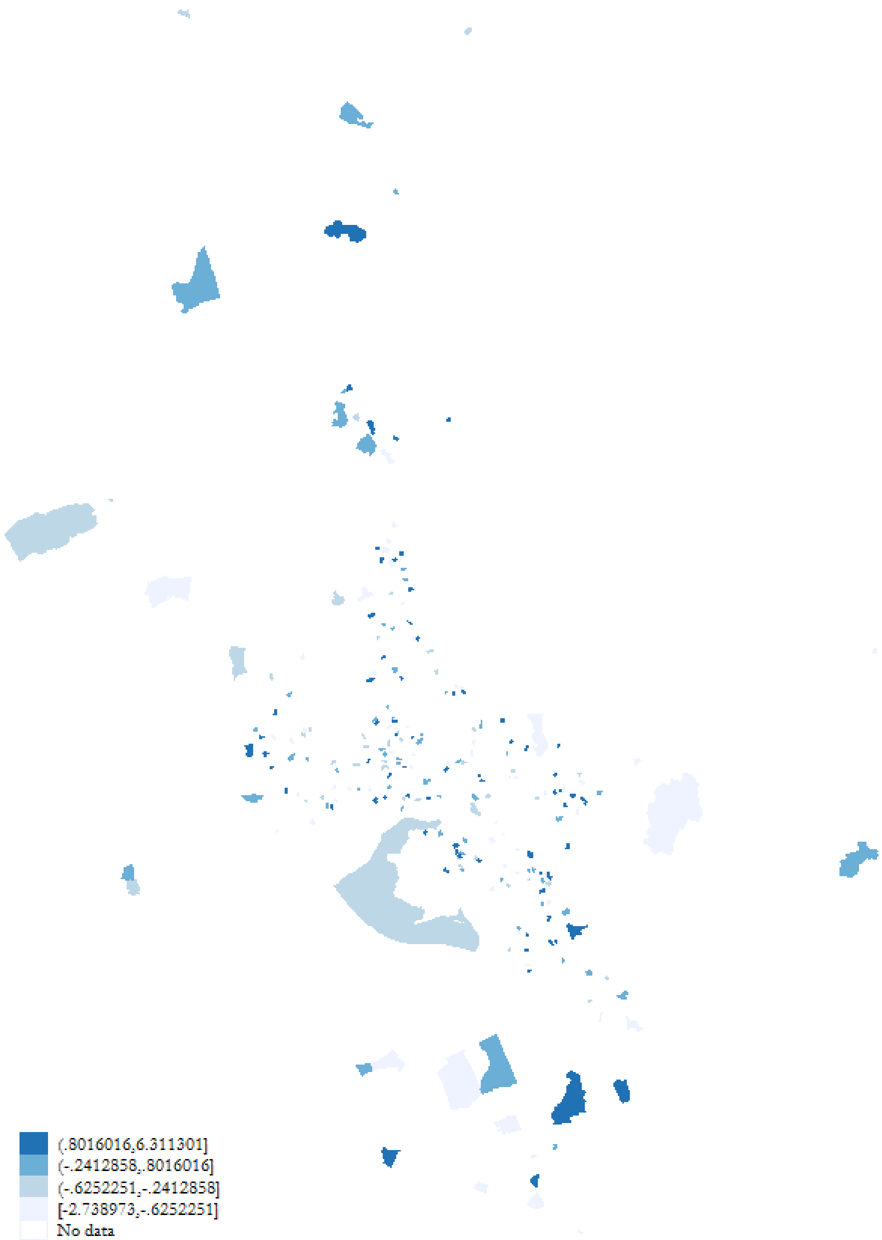
Observations	4607	4607	4607	4607	4607
Population size	904053	904053	904053	904053	904053
R^2	0.192	0.193	0.189	0.196	0.197
AIC	5039.068	5039.068	5075.1	5032.068	5030.573
Joint significance:					
Unexpected relative amenity measures†				<0.1	<0.1
Unexpected deprivation††				Insig.	Insig.
Unexpected crowding measures‡				<0.001	<0.001
Unexpected density measures‡‡				<0.01	<0.01
Expected Deprivation measures~			Insig.	<0.1	<0.1
Expected crowding measuresǀ			Insig.	Insig.	Insig.
Expected density measuresǁ			Insig.	Insig.	Insig.
Absolute deprivation~	Insig.	Insig.			
Absolute crowdingǀ	<0.001	<0.001			
Absolute densityǁ	<0.01	<0.01			

Reported coefficients are exponentiated ones. Standard errors are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
† The significance level for the relative deprivation term is reported on the relative deprivation test row. Note that the reported joint significance p_values are for different coefficients in different models[110]. † The significance level for $(\mathrm{RDEP} - \widehat{\mathrm{RDEP}})\#$ Dummy1 and $\mathrm{ATVARDENS} - \mathrm{ATV\widehat{ARD}ENS}$ is reported jointly on the unexpected relative amenity measures Wald test row. †† The significance level for $\mathrm{DEP} - \widehat{\mathrm{DEP}}$ and $\mathrm{DEP}^2 - \widehat{\mathrm{DEP}}^2$ is reported jointly on the unexpected deprivation measures Wald test row. ‡ The significance level for PPBR-$\widehat{\mathrm{PPBR}}$ and its squared term is reported jointly on the unexpected crowding measures Wald test row. ‡‡ The significance level for $\mathrm{DENSA} - \widehat{\mathrm{DENSA}}$ and its squared term is reported jointly on the unexpected density measures Wald test row. . ~The significance level for $\widehat{\mathrm{DEP}}$ (or absolute deprivation level (DEP) for the models without expected terms) and its squared term is reported jointly on the absolute density measures Wald test row. ǀThe significance level for $\widehat{\mathrm{PPBR}}$ (or absolute crowding level (PPBR) for the models without expected terms) and its squared term is reported jointly on the relative crowding measures Wald test row. ǁThe significance level for $\widehat{\mathrm{DENSA}}$ (or absolute density level (DENSA) for the models without expected terms) and its squared term is reported jointly on the relative density measures Wald test row. The reported AICs are for the models that do not take into account jackknife replications.

[110] Joint tests of other variables discussed in text.

Appendix III. The distribution of disturbances

Figure 22. The distribution of error terms over mesh-blocks.



Note: there is no possibility for depicting the distribution of residuals at the individual level, due to the confidentiality issues. Therefore, the aggregates of residuals at mesh-block levels are presented in Figure 22. Due to the small definition of mesh-block boundaries, depicting all boundaries will lead to a meaningless illustration at the city level. Hence, the borders of mesh-blocks and the areas that I do not have any data on are omitted from the illustration. Note that the 'no data' category may refer to areas with large amenities and facilities, such as parks.

10 SPWT: A Stata Command for Creating Spatial Weight Matrices ✩

Abstract

Creating spatial weights for spatial analysis takes an enormous memory space. To overcome the Stata's matrix size limitations, the analysis should be done by using the Mata language. This lets researchers to create large matrices efficiently, in terms of the processing time. The weighting options of **spwt** include the contiguity weights based on the k-nearest neighbour and the walking distance from the areas' centroids methods. -spwt- is a Stata package developed for the purpose of providing a walking distance definition for the neighbourhood boundaries used in the construction of relative measures (of crowding) in the last chapters[111].

10.1 Introduction

As discussed in chapters 8 and 9, a complementary approach to promote the relative measurement is to use the Network Spatial analysis to account for walking distance from mesh-blocks' centroids (for a review of the geographic scales, e.g. mesh-block, see 3.2 and for a review of spatial weight matrix construction see Appendix i.4 of chapter 8).

The regional economic analysis of this thesis needs to combine the complex statistical analysis with geographic analyses. Researchers often give up a combination between the statistical and geographical packages due to the compatibility issues. Hence, most studies are empowered by complicated statistical analyses but are limited to simplified geographical analyses, or the other way around, due to the limitations imposed by analysing packages. For instance, throughout this thesis, Stata is the main statistical package used for statistical inference. At different stages, analyses need to use different forms of the spatial weighting matrices defined at different geographical scales. The construction of the spatial weight matrices, by using a combination of Stata with a geographic software[112], is very time taking, if not infeasible. Due to the lack of spatial tools in Stata, in order to overcome the trade-off between the statistical and geographical available tools, I have written a Stata program. I have called this -spwt-, which may be changed when the package will be published in the future.

While, during the last few years, the use of Stata has boosted remarkably, regional researchers still need to use a combination of packages, such as ArcGIS, Geoda mixed with Stata. The introduction of **spmap** and other 'sp' packages was a significant improvement (Pisati, 2008). However, many of the packages do not

✩ Torshizian appreciates the supports received on the Statalist forum.

[111] The package and the relevant files, including an example for trying this code, are attached to this thesis (see the CD-ROM).

[112] ArcGIS is the geographic package usually used for spatial analyses.

support the use of large datasets. The introduction of Mata in the version 9 of Stata may be a solution for the matrix size limitations that researchers face in Stata.

There are a number of commands for creating spatial weights, e.g. **spmat** (Drukker, Peng, Prucha, & Raciborski, 2013). However, these commands have a limited matrix size, i.e. they cannot handle matrices greater than the maximum matrix size in Stata, which is 11,000 variables.

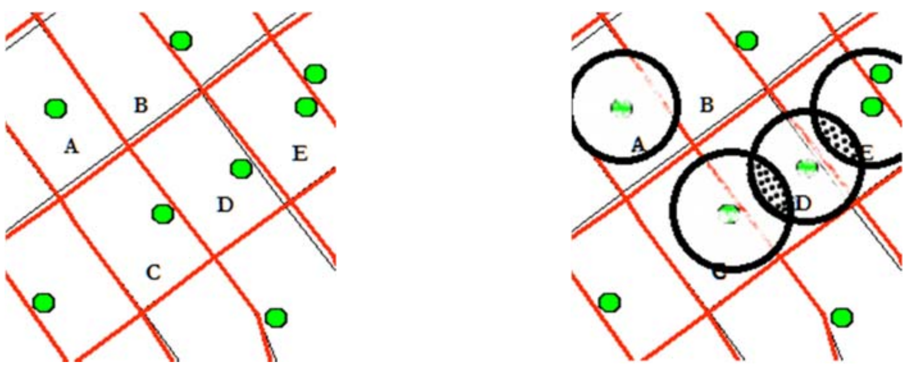
Spatial weight matrices can be constructed based on distance, boundary or a combination of them. spmat supports the distance based methods, such as the k-nearest neighbour. A complementary approach to promote the simple measures is to account for walking distance from areas' centroids (for applications see chapters 8 and 9). In the next section, the 'walking distance' approach will be discussed in details.

10.2 SPWT's weighting methods

10.2.1 The walking distance approach

This method derives polygons of, for example, 15 minutes walking distance around the centroids of areas. Figure 23 depicts areas A, B, C, D and E located next to each other. Lines depict the boundaries and the solid circles are the centroids of each area. On the right hand side of the figure, the circles around the centroids depict a 15-minute walking distance from the centre of each area.

Figure 23. An example of the walking distance method.



Source: Torshizian (2017).

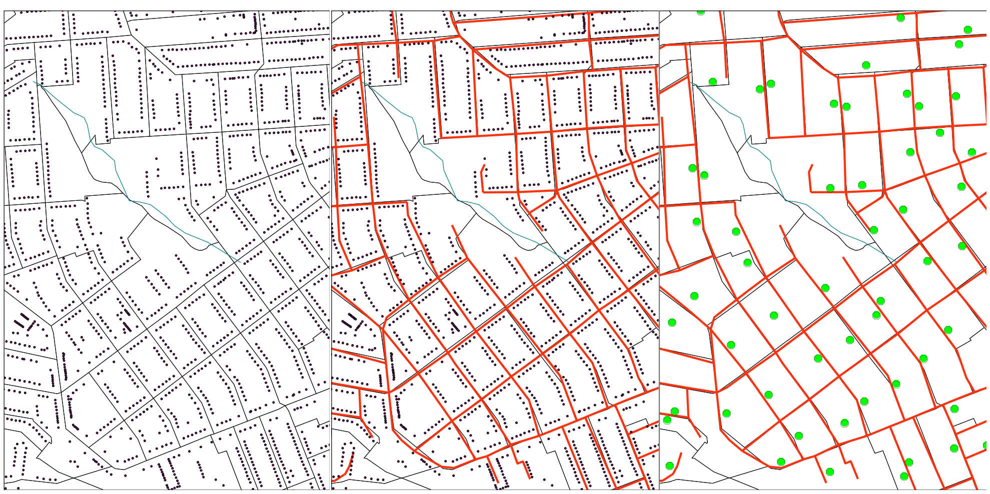
This derives some areas intersecting one another. For example, area D has an overlap with areas D and E. The overlap of an area's polygon layer divided by its total intersecting area with other areas can be considered as the areas' standardised weights.

In continue, to fix ideas a realistic example would be considered. A number of areas located in an area unit of Auckland are depicted on the left hand side of the Figure 24, in which lines represent the boundaries of areas. In order to know about the places that an individual may start a walk, I should know about the places of dwellings. On the right hand side of Figure 24, depicted dots represent the entrance of dwellings.

Figure 24. a map of mesh-blocks and doors.

After knowing about the start points, the walking paths should be recognised. People can either walk through tracks, depicted on the very left hand side of the Figure 25, or through roads, illustrated on the centre of the Figure 25. Before deriving the walking distance polygons, the start points should be clarified. For the sake of simplicity, a dwelling which is located the most centrally in each area, is considered as the start walking point, which is illustrated on the right hand side of the Figure 25.

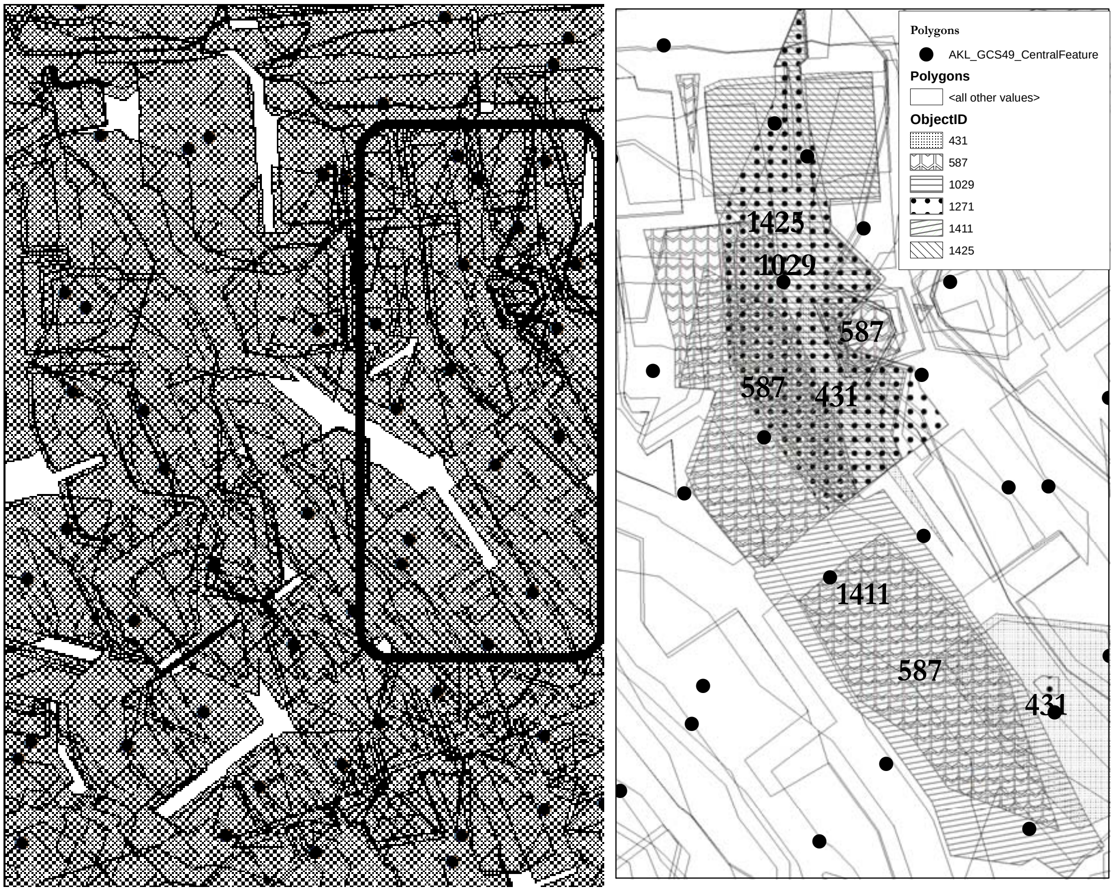
Figure 25. A map of walking paths and centroids.



Source: Torshizian (2017).

Derived polygons for a 5 minute walk from the centroid of each mesh block is illustrated on the left hand side of the Figure 26, in which black dots represent areas' centroids and shaded polygons indicate the walking polygons. The circled area is illustrated in details on the right hand side of the Figure 26, which ables us to explain the spatial matrix derivation.

Figure 26. Walking distance polygons.



Source: Torshizian (2017).

The walking distance polygon of the area number 1411 is overlapped polygons 431, 587, 1029, 1411 and 1425[113]. To derive the geographic weights for the area number 1411 in the spatial matrix, I need to calculate the ratio of the intersecting area between 1411 and the other areas to the total overlapping area. The total overlapping area is equal to 241374 square meters, of which 13227.841 square meters are overlapped area number 431. Thus, as it is depicted in the following matrix (Table 43), the row number 1411 column number 431 of the spatial matrix is equal to 0.054 ($=\frac{13227.841}{241374}$).

Table 43. Spatial walking distance matrix.

	431	587	1029	1411	1425
1411	0.054	0.277	0.053	0	0.049
1029	0.067	0.07	0	0.052	0.233

[113] There are a number of other polygons overlapping with area unit 1411's polygon. As the purpose of this example is to fix idea, I have simplified the illustration by excluding other polygons.

10.3 help: spwt[114]

10.3.1 Syntax

spwt [using filename], id(idvar) xcoord(varname) ycoord(varname) kneighbour(num) distance ratio standardise centroid(name) far(num) clip(filename)

Options

using filename imports id(idvar), which is the polygon identifer, and the xcoord varname x and ycoord varname y coordinates from filename, which presumably is produced by using spmap (or si mmands).

id(idvar) is always required. It specifies the name of the polygons' identifiers.

xcoord(varname) is always required. It specifies the name of the variable containing the x-coordinate of each location object of analysis.

ycoord(varname) is always required. It specifies the name of the variable containing the y-coordinate of each location object of analysis.

kneighbour(num) is always required. It specifies the number of the nearest neighbours that should be presented in the final spatial matrix.

distance generates the distance matrix.

ratio generates the weights as a ratio of the total distance from other areas.

standardise creates row standardised weight matrices, i.e. the weights would be expressed as a ratio of the number of neighbours or, in the case of the walking distance method, the total intersecting area.

centroid: in the case of the walking distance approach, the name of the coordinates of centroids of the geographical areas should be entered by using this option.

far(num): enters the distance that should be considered in the walking distance approach.

clip(varname): keeps the area IDs included in the filename dataset.

[114] This is the help file included in the written package.

normalize: normalises the weights, i.e. the weight would be expressed as a ratio of the number of neighbours or, in the case of the walking distance method, the total intersecting area.

10.3.2 Description

There are a number of commands for creating spatial weights, e.g. spmat, spatwmat. However, these commands have a limited matrix size, i.e. they cannot handle matrices greater than the Stata's maximum matrix size (see -help matsize-). In order to solve this problem, spwt is written in a combination of Stata and Mata. This integration also reduces the processing time.

The user, also, may need to specify the areas that should be included in the analysis. This can be done after importing the shapefiles into Stata by using shp2dta. The outcome of this command is a basemap, which depicts the spatial data in a Stata dataset format.

10.3.3 Outcome

spwt saves the spatial weights in matrix W.

10.3.4 Relevant packages

spmat, spatwmat, shp2dta, spcat, spmap, spatgsa, spatcorr, spatlsa, spatdiag, spatreg

10.4 Examples

NOTE: The examples presented here are based on the mesh-block shapefile, which is available in the directory where the spwt ancillary datasets are located. For illustration purposes, only one of the areas in Auckland is presented here, which is Area unit number '514000'. The complete dataset is available on request.

10.4.1 Example 1: the k-neighbour method

First, the Shapefile should be imported to Stata by using the **shp2dta** command:

```
. shp2dta using AU514000, data("AU514000") coord("AU514000_coord")
```

Then the basemap should be opened in Stata:

```
. use AU514000_coord, clear
```

To create the 5-nearest neighbours weight matrix, the following command would be used:

```
. spwt, id(_ID) x(_X) y(_Y) k(5)
```

Instead of opening the basemap in Stata, it is possible to introduce the basemap file to the spwt command as follows:

```
. spwt using AU514000_coord, id(_ID) x(_X) y(_Y) k(5)
```

10.4.2 Example 2: dropping a number of areas

In this example, only the areas which are included in another dataset (MB_merging) will be kept in the final spatial weights. Since, the using dataset does not include the area identifiers, first I need to merge it with the main dataset.

```
. shp2dta using AU514000, data("AU514000") coord("AU514000_coord")

. use AU514000, clear

. spmap using "AU514000_coord", id(_ID)
```

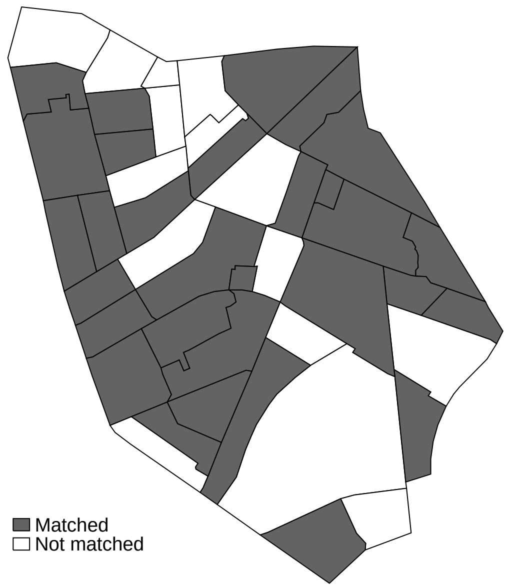
The areas are illustrated in Figure 27.

```
. merge 1:1 MB06 using MB_merging

. spmap _merge using "AU514000_coord", id(_ID)
```

The gray areas, which are the matched areas, should be kept and the rest should be dropped from the analysis. 15 areas out of the total of 45 will be omitted from the analysis.

Figure 27. An example of a merged dataset.



Source: Torshizian (2017).

```
. keep if _merge == 3

. save "AU514000_clipped", replace
```

Then the clipped dataset (AU514000_clipped) should be introduced to **spwt**.

```
. spwt using "AU514000_coord", id(_ID) x(_X) y(_Y) k(5) clip("AU514000_clipped")
```

Example 3: the walking distance method

First, the centroids of areas need to be derived by using the **shp2dta** command:

```
.shp2dta using AU514000, data("AU514000") coord("AU514000_coord") genid(ID) gencentroids(centre) replace
```

Then, the weight matrix should be produced based on a 1000 meters walking distance from the centroids:

```
.spwt using "AU514000", id(ID) cent(centre) far(1000)
```

Appendix I. Written package for Stata – code[115]

```
*! Version 1.0.1 - September 2014

*! -spwt- Generates k-nearest neighbours spatial weight matrices

*! Author: Eilya Torshizian

*! Department of Economics

*! University of Auckland

*! eilya@torshizian.com

capture program drop spwt

program define spwt

version 11.0

syntax [using/], ID(string) [Generate(string) Xcoord(string) Ycoord(string) Kneighbour(string) DIStance
CENtroid(string) FAR(string) STandardize Ratio Clip(string) COUNT HISTogram]

****************

* Check syntax *

****************

if "`using'"=="" & ("`xcoord'"=="" | "`ycoord'"=="") & "`centroid'"=="" {

        di as err "Please specify both x and y coordinates using options "

        di as err "{bf:{ul:x}coord({it:varname})} and " _c

        di as err "{bf:{ul:y}coord({it:varname})}"

        exit

}

************************************

*Import the coordinations from using*

************************************

if "`using'"!="" {

        preserve

        qui use `"`using'"', clear

        }

if `"`using'"'=="" & "`centroid'"=="" {

capture qui assert `xcoord'!=.

if _rc!=0 {
```

[115] The programming code of this command and the relevant files, including an example for trying this code, are attached to this thesis (see the CD-ROM).

```
		di as err "Variable `xcoord' has missing values"
		exit
		}
capture qui assert `ycoord'!=.
if _rc!=0 {
		di as err "Variable `ycoord' has missing values"
		exit
		}
local N=_N
}
if "`clip'"!=""{
		merge m:1 `id' using "`clip'"
		keep if _merge==3
		drop _merge
}


if "`standardize'"!="" {
		local normalize = 1
		}
		else {
		local normalize = 0
		}
if "`ratio'"!="" {
		local ratio = 1
		}
		else {
		local ratio = 0
		}
if "`distance'"!="" {
		local distance = 1
		}
		else {
```

```
        local distance = 0
        }
if "`centroid'"!="" {
        local centre = 1
        local xcoord = "x_`centroid'"
        local ycoord = "y_`centroid'"
        local far = `far'
        }
        else {
        local centre = 0
        local far = 0
        egen IDgen = group(`id')
        qui drop if `xcoord'==.
        }
******************************
*THE DISTANCE FUNCTION IN MATA*
******************************
qui putmata `xcoord' `ycoord' `k', replace
mata: W = DIST(`xcoord',`ycoord',`kneighbour',`normalize', `ratio', `distance', `far')
if ("`count'"!="") {
mata: C = count(W)
dis "The number of neighbours is saved in matrix C."
dis "The average number of neighbours is equal to " n
}
if ("`histogram'"!="") {
svmat C, name(COUNT)
dis "The number of neighbours is saved in the COUNT variable."
hist COUNT, xtitle("Number of neighbours")
}
end
mata
function DIST(X,Y,k, norm, wratio, distance, FAR){
if (FAR!=0){
```

```
        Z = 0
        A = 0
        W=J(rows(X),rows(X),0)
        for(i=1;i<=rows(X);i++) {
                for(j=i+1; j<= rows(X); j++) {
                        Z = sqrt((X[i]-X[j])^2+(Y[i]-Y[j])^2)
                        if (Z < FAR){
                                A = 2*((FAR^2)*acos(Z/(2*FAR))-(Z/4)*sqrt(4*(FAR^2)-(Z^2)))
                                W[i,j] = A
                                }
                        }
                }
W = uppertriangle(W) + uppertriangle(W)'
if (norm==1) {
for(i=1;i<=rows(X);i++) {
        R=rowsum(W[i,.])
        if (R!=0){
                W[i,.] = W[i,.]/R
                }
 }
}
}
else{

F = J(rows(X), rows(X), .)
        for(i=1; i<= rows(X); i++) {
                for(j=1; j<= rows(X); j++) {
                F[i,j] = sqrt((X[i]-X[j])^2+(Y[i]-Y[j])^2)
                }
        }

i = 1
r=0
```

```
r1=0
st_view(ID,.,"IDgen")
m = max(ID)
D = J(1,m,.)
while (i<rows(ID)) {
        for (j=1;j<=rows(ID);j++) {
        if (ID[j] == i) {
                D[1,i]=j-r-1
                r=r+1
                }
        else{
                i=i+1
                r=0
                r1++
                }
        }
}
D = D, rows(ID)

l=1
i = 1
Z = J(cols(D), cols(D), .)
while (i<=cols(D)){
        for(j=1; i+j+1<= cols(D); j++){
                Z[i,j+i] = min(F[D[i]+1..D[i+1]-1,D[j+i+1]..D[i+j+1]-1])
                }
        i++
}
Z = uppertriangle(Z) + uppertriangle(Z)'
W=J(cols(D),cols(D),0)
if (distance == 1) {
W = Z
        if (ratio==1){
```

```
                    for(i=1; i<= cols(D); i++){
                              W[i,.] = W[i,.]/rowsum(W[i,.])
                              }
          }
}
else{
for(i=1; i<= cols(D); i++){
          minindex(Z[i,],k,IN=.,junk=.)
  W[i,IN]=J(1,rows(IN),1)
          if (norm==1){
                    W[i,.] = W[i,.]/rowsum(W[i,.])
                    }
}
}
W=W[|1,1\cols(D)-1,cols(D)-1|]
}
st_matrix("W",W)
printf("The weight matrix W saved in the matrix directory. Type -matrix list W- to see the matrix.")
return(W)
}


function count(M){
          CS = colshape(M:!=0,1)
          IND = selectindex(CS)
          C = cols(M)
          R = ceil(IND/C)
          COUNT = J(C,1,0)
          j = 1
          for (i=1; i<=C; i++){
                    z = 0
                    while (R[j]==i & j!=rows(R)){
                              j++
                              z = z+1
```

```
        }
        COUNT[i] = z
    }
    n = colsum(COUNT)/C
    st_matrix("C",COUNT)
    st_numscalar("n",n)
    return(COUNT)
}
End
```

11 Conclusion

This thesis assessed the impact of household crowding, population density and the proximity of rich and poor in the city on people's satisfaction with their residence (which I term 'residential satisfaction'). In order to contribute to more sustainable development, there is a growing interest in increasing the intensification of urban areas by using policy instruments. However, the outcome of intensification may be a higher crowding level in addition to the increase in population density.

In addition to the substantial question of interest (the impact of the factors of space on the welfare of residents), this thesis is motivated by a range of theoretical, empirical and technical questions.

Despite the recent attention of economists to alternative measures of welfare, urban economists have only rarely considered outcomes other than prices and rents. Urban intensification can raise rents without raising (and possibly lowering) residential satisfaction. Therefore the economists' conventional measure of value, price, may not be adequate to judge the welfare consequences of intensification. The relative merits of judging policy outcomes on the basis of market price and individual measures of satisfaction are unclear. Rent (land price) is a collectively (market) generated objective measure of value accorded a site, while residential satisfaction is a subjective measure of the qualitative value of that site to the individual. The problem is that when applied to any given residence the two measures of 'value' may be weakly correlated. They may even be negatively correlated – reflecting the strong preference of New Zealanders for low density living. Using the satisfaction of residents with their living environment as its outcome of interest, the current thesis provided a more holistic view of the features of urban development.

Governments are trying to make economic development more sustainable. When it comes to the growth of cities, the primary policy tool for increasing sustainability is residential intensification. However, the primary outcome of intensification is greater population density, higher rents (which leads to higher levels of crowding within dwellings) and, in many cases, the increased proximity of widely disparate income groups. It is possible therefore that in an attempt to become more sustainable, increasing intensification will lead to a reduction in the overall wellbeing of urban residents.

A two stage least squares approach is used for addressing the location choice issue. The thesis used a rich data set derived from three rounds of the New Zealand General Social Survey, 2008, 2012, 2014 plus 2006 Census data covering the area administered by the Auckland Council. The relationships are examined at a fine geographic level accounting for a wide range of socioeconomic factors and by using appropriate spatial regression techniques.

This thesis first examines how household crowding and population density affect residential satisfaction after accounting for people's location choice and their social comparisons. Then the thesis assesses if relatively poor (rich) areas experience a positive (negative) amenity (NIMBY) effect, as opposed to a negative envy effect, when located next to less (more) deprived areas.

Household crowding can be measured using a number of indexes. Some, such as the Canadian National Occupation Standard index (CNOS), account for cultural issues and households' compositions, while others, such as the people per bedroom measure (PPBR), do not. In the first substantive essay (Chapter 6), I compare these measures with one another and with people's perceptions about the crowding of their dwelling in terms of the variables' explanatory power for the satisfaction of residents. Our results indicate that amongst objective measures, CNOS is a better predictor than PPBR for residential satisfaction, although PPBR remains a useful predictor of residential satisfaction in the absence of other measures. After taking people's location choice into account, a higher level of population density in the resident's own suburb lowers residential satisfaction. At the same time, residents are more sensitive to the perception of crowding [in the dwelling] than with objective measures of crowding.

I analysed the role of households' residential positions, particularly the level of household crowding and the level of area density, in determining the satisfaction of residents. The results indicate that a higher level of household crowding than could be expected based on an individual's background (reflecting that individual's social norms), which is estimated based on their background, decreases residential satisfaction. The study also showed that, after taking people's location choice into account, a higher level of density in the resident's own suburb lowers residential satisfaction.

The two goals of the third study, presented in Chapter 8, are to find an optimal reference group and an optimal set of subjective measures for social comparisons. The results indicated that people respond to reference groups, which in the Auckland case are found to be those who live within a 15-minute walk of the resident. After taking into account the endogeneity of residential choice, I found that people take the density of their neighbourhood compared to other suburbs into account when evaluating their living environment. By contrast, relative crowding positions do not significantly affect residential satisfaction.

The (chapter 9 of the) thesis discussed the idea that residents of a deprived area located next to an affluent area may experience: (i) higher levels of residential satisfaction as a result of the amenities provided by the rich area, or (ii) lower levels of residential satisfaction as a result of envy. The results indicate that people are negatively affected by locating in an area with a deprivation level that is higher than that they are expected to experience. I also find that residents of poor (rich) areas are positively (negatively) affected by the affluence of neighbouring areas, which confirms the presence of a positive amenity effect for the residents of poor neighbourhoods and a negative 'not-in-my-backyard' (NIMBY) effect for the residents of rich neighbourhoods.

In summary, this thesis demonstrates the worth of using residential satisfaction in urban economic studies. The results indicate that people apply expectations based on their social norms when evaluating their living environment. After taking these social norms into account, people value living in less crowded houses, in less dense suburbs and in proximity to affluent areas.

11.1 Introduction

This thesis aimed to study the impact of household crowding, area density and deprivation on the satisfaction of residents with their living environment. The first chapter presented an overview of the thesis, including motivations for this research. A part of the literature, the descriptive statistics, the outcome of interest, and the methodology of the main topics of this thesis are in common. Therefore, the common parts of the literature were presented in chapter 2, the common descriptive statistics presented in chapter 3, discussion about the outcome of interest were presented in chapter 1, and the common methodology issues were discussed in chapter 5.

Three cross-sections of the General Social Survey, 2008, 2010 and 2012, are merged with Census 2006 for the purpose of this study. This provides the thesis with 5716 observations after the drop of outliers, as explained in section 3.3. This dataset is used for the analysis of chapter 1, but due to some changes needed for spatial analyses in the next chapters, the sample population for chapters 7, 8 and 9 is equal to 4607 observations. According to the resampling weights used for the estimations of this thesis, the sample population is representative of 904,053 Auckland residents, as discussed in 3.1.3.

11.2 Summary and discussion

I start this section by presenting a short summary of the topics discussed in Chapters 2 to 5. Then, in the five subsequent subsections, I provide a summary of each of the four main essays, presented in Chapters 6 to 9, and the additional technical essay related to spatial effects estimation presented in Chapter 10.

As mentioned in Chapter 2, Auckland, which has been described as one of the world's most livable cities, has experienced a significant increase in its housing prices during the last few years. One reason for the high house prices is that they reflect an increase in demand that is not compensated by a matching increase in supply. A higher supply can be achieved by an increase in the supply of land, through an expansion of the city's urban area, or further intensification in the current urban area. Considering that Auckland Council has aimed to limit further expansion of the city in the next 30 years, we need to improve our understanding of the impact of intensification on residents' satisfaction – a topic that is informed by this dissertation.

Residential satisfaction, which is the outcome of interest in the current study, is derived from the answers to New Zealand General Social Survey on "How do you feel about where you are currently living'. Originally, five categories were defined for the answers to this question. In this study, based on their statistical properties, I aggregate the five categories into two groups: satisfied and dissatisfied.

Two main variables of interest in this thesis are household crowding and density, both of which can be measured using different definitions. Perceived crowding is introduced as a subjective measure of household crowding. Density can be measured using the area density and the residential density measures, both of which were tested in the current study.

Neighbourhood effects are the externalities involved in living in specific types of neighbourhoods. As discussed in Section 2.4, considering the complexities involved in identifying neighbourhood effects, a careful understanding of these effects must account for a range of issues and a spatial analysis must provide a correct definition of neighbourhood boundaries. Measuring the causal impact of residential segregation on residents' satisfaction is difficult because of the other factors that affect individuals' location choice, and the estimation methodology needs to be of an appropriate form to overcome the potential endogeneity issues.

In order to find the correct definition of neighbourhood boundaries, I have considered three conventional definitions of boundaries at the mesh-block level, area-unit level and territorial Authority level. Mesh-blocks are very small neighbourhoods that consist of a few houses located next to each other, with an average population of 110 people. Area units are aggregates of mesh-blocks and are the suburbs that people usually account themselves as a member of. Territorial authorities are larger areas determining communities of interest that share a common local government.

Table 2 presents descriptions of a wide range of variables that have been used in the current study. Figures 9 and 10 illustrate the percentages of various housing problems across each residential satisfaction category and the proportion of various problems with neighbourhoods amongst different residential satisfaction levels, respectively. In each figure, it can be seen that the prevalence of problems (with the house and the neighbourhood respectively) decline along with the degree of residential satisfaction. This indicates that our interpretation of the residential satisfaction variable as being determined both by house-specific and neighbourhood-specific factors is likely to hold. Thus, I consider it appropriate to interpret the residential satisfaction variable as reflecting factors at both the house and the neighbourhood levels. I will discuss this further in the next section.

The most important residents' problems with their dwellings are heating the dwelling and the having a small house. Noise or vibration are reported as the most important neighbourhood problem, followed by safety issues in the neighbourhood. All problems with dwelling and with neighbourhood are included in the regression analyses undertaken in this thesis.

As noted in Chapter 4, using prices as the outcome of interest is very common amongst urban economic studies. One of the shortcomings of the price measure is that it is not informative about people's desired choice that has not been provided in the market; that is, prices cannot be used to price individual unrevealed preferences.

The use of subjective measures in the economics literature became more common after Easterlin's (1974) study on the economics of happiness. Many economists discussed that subjective measures are similar to measures of the consumer's utility because they account for perceptions as well as observed economic characteristics, and also because they account for the desired choices of consumers that have not been provided in the market.

As per our discussion in Chapter 5, the ordered logit method has been used for the regression analyses of this theses. However, I needed to ensure that the ordered logit's parallel assumption is not violated. I tested this assumption by using the Brant test. The result suggests that I can rely on the parallel assumption of the ordered logit method.

Aggregation of the five-category residential satisfaction variable into a binary variable simplifies the presentation and interpretation of results. To adopt this aggregation, I needed to ensure that no information was lost as a result of the re-grouping. Hence, I compared the estimation results derived from three models: a model using the ordered logit method, and two binary logistic regressions with different definitions for the binary outcome of interest. The results suggest that the ordered logit's results are similar to those derived from estimation using the binary variable in statistical and economic terms. Hence, residential satisfaction in the current study is aggregated into two categories such that the very satisfied category is considered as one group (0) and the other four categories are grouped as the other category (1).

11.2.1 Essay 1 – Raw, Adjusted and Perceived Crowding Measures

The first essay of this thesis (Chapter 6) introduces the best measure of household crowding (in isolation of other factors) in terms of its predictive power for residential satisfaction. As discussed in Section 2.2, residential satisfaction is a combination of objective and subjective characteristics of the living environment, where the latter includes cultural perspectives of crowding. Besides perceived crowding (PC) and the objective measures, including people per room (PPR), people per bedroom (PPBR) and adjusted people per bedroom (APPBR), I used the Canadian National Occupation Standard (CNOS) as an adjusted crowding measure that accounts for some characteristics of individuals, such as partnership status and the number of children. Also, a combination of these measures is constructed by taking a principal component approach. The short list of the crowding measures and the factors they account for is as follows:

Table 44. A list of crowding measures and their characteristics.

Index	Based on	Couple status	Age that boys and girls can share	Age that same sex children can share
PPR	Rooms	Not used	-	-
PPBR	Bedrooms	Not used	-	-
APPBR	Bedrooms	Used	-	-
CNOS	Bedrooms	Used	Under 5	Under 18

The study hypothesises that a higher residential satisfaction is affected both by a perception of less crowding and by a lower value of the raw (objective) crowding measures. The impact of the variables of crowding on residential satisfaction is estimated based on a quadratic function by fitting a logistic model for our binary response outcome: residential satisfaction. Controls for personal characteristics are included in all regressions.

The results show that the perceived crowding measure has higher predictive power than the objective crowding measures, which indicates that perceptions are an important manifestation of people's evaluations of their household crowding. However, all crowding measures are significantly related to each other; that is, each gives a similar qualitative impact on residential satisfaction with all measures of greater crowding impacting negatively on residential satisfaction. Amongst objective measures, CNOS is a better predictor than PPBR for residential satisfaction . Nevertheless, PPBR remains a useful predictor of residential satisfaction in the absence of the other measures and, given its wide availability, is used as the measure for household crowding in the subsequent chapters. The formula for PPBR is as follows:

$$\text{PPBR} = \frac{\text{Number of people in the dwelling}}{\text{Number of bedrooms}}$$

11.2.2 Essay 2 – Residential Satisfaction, Household Crowding and Density

The second essay analyses the role of households' absolute residential positions, particularly the level of household crowding and the level of area density, in determining the satisfaction of residents. In their residential environment evaluations, people may think of their neighbourhood boundaries at a very local level, such as an area just 'over the road', at the area known as their neighbourhood or suburb or at a larger area level. For the purpose of this study, the measures of both crowding and area density need to be chosen appropriately. The chosen measure of household crowding, as discussed in the last chapter, is PPBR, although I add perceived crowding (PC) as a robustness check in my estimation.

The location choice of residents is a very important issue that is addressed in this chapter. It is very likely that segregation/desegregation in neighbourhoods is affected by ethnicities and occupations. Not accounting for location choice may lead to biased estimates. As Table 45 shows, for the variable measuring the number of generations in New Zealand, for example, an excess concentration of at least 25 per cent for people with different numbers of generations in New Zealand happens in 19 per cent of neighbourhoods. Therefore, as explained in Section 7.6, this variable will be used as an instrument for capturing individuals' location choice. Two other variables that are used as instruments for capturing the population location patterns are the number of people who raised the individual who were born in New Zealand, and the ease of expressing one's own identity in New Zealand (for details, see Section 7.2.1).

Table 45. Factors of population distribution pattern, an example.

			Excess concentration relative to mean			
Variables	Categories	Mean	10%	15%	20%	25%

The number of generations in New Zealand	First generation	0.36	0.3	0.17	0.15	0.08
	Second generation	0.1	0.14	0.08	0.05	0.03
	Third generation or more	0.52	0.3	0.2	0.12	0.07
	Total		0.77	0.48	0.33	0.19

* The non-stated categories of the variables are not included in the table.

I estimate an individual's expected level of household crowding based on all characteristics of people and their residential environment, as well as their factors of location choice. In other words, people choose their living environment such that it suits them the best. The best residential environment is one that provides people with the amenities and facilities that they need. The need is defined based on an individual's background. For example, people whose parents are New Zealanders may have their needs defined differently from those whose parents are foreigners.

Therefore, this study uses a two-stage least squares approach to estimate the level of household crowding and the level of density that an individual may expect based on his or her background. I am interested in the impact of the difference between the actual (absolute) residential position measures and the expected values on residential satisfaction; in other words, the difference between actual and expected levels is an indicator of the level of crowding and density that an individual does not expect (or wish) to experience.

I then construct a range of density measures based on both residential and population definitions. The measures are constructed at different neighbourhood boundaries: at a very local level, such as an area just 'over the road' (mesh-block); at the area known as their suburb (area unit); or at a larger area level (territorial local authority). My aim is to find the neighbourhood boundary that provides the best predictive power for residential satisfaction. Based on this density measure, the study hypothesises that a density level above an individual's expected density level lowers residential satisfaction.

Table 24 summarises the results derived from the equations shown in Table 23. The absolute crowding measure (PPBR) has a very significant negative effect. The impact of the expected household crowding on RS is positive; that is, $\widehat{\text{PPBR}}$ raises RS. This shows that when people can satisfy their crowding expectations, they are happy with their residential environment. For example, a person whose background favours children sharing a bedroom may wish to live in a dwelling with two bedrooms for his or her four-person household. This positive significant effect confirms the role of norms in individuals' satisfaction with their residential environment. However, an unexpected level of additional crowding lowers RS, which suggests that once the level of crowding is more than the level that an individual may expect to have, that person is more likely to be dissatisfied with their living environment.

Table 46. A summary of the results of the 7th chapter.

Category	Result	Significance
Household crowding (PPBR)	A higher household crowding level lowers RS.	Yes
Household crowding derived from backgrounds ($\widehat{\text{PPBR}}$)	A higher household crowding stemming from 'own group' attraction raises RS.	Yes
Unexpected household crowding ($\text{PPBR} - \widehat{\text{PPBR}}$)	A higher unexpected level of crowding lowers RS.	Yes
Absolute area density (DENSA)	A higher absolute density lowers RS.	Yes
Absolute area density derived from backgrounds ($\widehat{\text{DENSA}}$)	A higher area density level stemming from 'own group' attraction does not affect RS.	No
Unexpected area density ($\text{DENSA} - \widehat{\text{DENSA}}$)	A higher level of unexpected area density lowers RS.	Yes
Residential density	A higher level of residential density does not affect RS.	No
Perceived crowding (PC)	A higher PC level affects RS negatively.	Yes

The comparisons between different measures of our other variable of interest, density, shows that area density (based on the area unit) is the best measure for residential satisfaction in terms of its predictive power. After taking a range of individual and area characteristics into account, a higher area density level (DENSA), measured at the area unit (or suburb) level, has a significant negative impact on RS. This impact, before taking the endogeneity into account, confirms our hypothesis that people (in Auckland) do not like to live in a dense area. When I account for the endogeneity of location choice in the area density measure, I find that a higher exogenous area density level (i.e. density greater than that expected or desired) lowers RS. Consistent with this study's hypothesis that people choose to live in an area with their preferred density, they endogenise their location choice. However, if because of other constraints the area is denser than they expect, they will not be happy with it.

People's perception of crowding (PC) is associated with lower RS. Including PC in the equation does not affect the significance level of the estimated negative effects of unexpected household crowding or density on residential satisfaction.

11.2.3 Essay 3 – Residential Satisfaction, Relative Crowding and Relative Density

Concentrating on objective measures of an individual household's situation may be insufficient when dealing with the determinants of satisfaction with residential environment. There is a long history of relative measurement in the economics literature, especially in welfare studies, in which 'keeping up with the Joneses' may play an important role in people's evaluations of their different life circumstances. Therefore, this essay aims to introduce an optimal reference group and an optimal set of subjective measures for social comparisons.

In this chapter, neighbours, as a group who live shoulder-to-shoulder with the person, are considered as the relevant reference group for some situations. I analyse the role of relative residential positions, in relation to neighbours, in determining the satisfaction of residents. Reference groups are considered as neighbours living over different geographic scales, neighbours who are the symbols of success over different geographic scales, and neighbours who may be visited within a 15-minute walk.

For the construction of the relative measures, different boundaries for neighbourhoods need to be taken into account. In order to take into account the geographic proximity of mesh-blocks, four geospatial datasets are used: mesh-blocks dataset, address points dataset, roads and tracks datasets. After adopting the appropriate projection system and editing the datasets, a matrix of spatial proximity is derived by using a network analysis toolbox (for more details, see Section 8.3.2).

In order to define spatial weight matrices by taking a walking distance definition for neighbourhood boundaries, I programmed a new tool for spatial analysis for this study's statistical software (Stata). The developed package is presented in Chapter 10.

The results of this study indicate that the most relevant reference group consists of people who may be visited in a 15-minute walk. Results for the effects of expected and unexpected (plus overall) household crowding and area density are the same as in the previous chapter. The key additional result is that living in a relatively dense area (overall, expected and unexpected) raises residential satisfaction. I conjecture that this reflects an amenity impact with greater civic amenities being provided in denser areas. Given that relative density is a zero-sum game across the city, the negative effect of greater (overall and unexpected) density has an overall negative impact on residential satisfaction within the city; but also that, given this effect, residential satisfaction can be improved by locating in an area with greater amenities associated with high relative density.

Table 47. A summary of the results of the 8th chapter.

Category	Result	Significance

Household crowding derived from backgrounds ($\widehat{\text{PPBR}}$)	A higher household crowding stemming from 'own group' does not affect RS.	No
Exogenous household crowding ($\text{PPBR} - \widehat{\text{PPBR}}$)	A higher unexpected level of crowding lowers RS.	Yes
Absolute area density derived from backgrounds ($\widehat{\text{DENSA}}$)	A higher area density level stemming from 'own group' attraction does not affect RS.	No
Exogenous area density ($\text{DENSA} - \widehat{\text{DENSA}}$)	A higher level of unexpected area density lowers RS.	Yes
Expected relative crowding ($\widehat{\text{CVARS}}$)	A higher level of expected relative crowding does not affect RS.	No
Unexpected relative crowding ($\text{CVAR} - \widehat{\text{CVARS}}$)	A higher level of unexpected relative crowding does not affect RS.	No
Expected relative density ($\widehat{\text{ATVARDENS}}$)	A higher level of expected relative density raises RS.	Yes
Unexpected relative density ($\text{ATVARDENS} - \widehat{\text{ATVARDENS}}$)	A higher level of unexpected relative density raises RS.	Yes
Perceived crowding (PC)	A higher PC level affects RS negatively.	Yes

11.2.4 Essay 4 – Negative Envy or Positive Amenity Effects

This essay starts with motivations for inferring the role of poorer residents locating nearby rich areas. Based on studies of the poverty trap, the residents of poor areas are likely to adapt to their own living conditions and, therefore, may not be dissatisfied with living in a poorer area. However, they may benefit from locating near the amenities offered in a richer area. To analyse such effects, a relative deprivation measure needs to be constructed.

As discussed in Essay 3, the most appropriate relative crowding measure is the one constructed based on a comparison between an individual's position and the position of people living in a 15-minute walking distance from that individual's location. The relative deprivation measure is constructed in the same way, based on a comparison between the deprivation level of an individual and the deprivation level of the areas that the individual may visit in a 15-minute walk around his or her location.

The present study uses a similar approach to the previous chapters for estimating the impact of the variables of interest on the satisfaction of residents. The variables of interest in this study are absolute and relative deprivation measures, as well as all the variables of interest discussed in the previous essays, including household crowding, area density and perceived crowding.

After taking account of endogenous location choice, both with respect to the level of deprivation of the individual's own neighbourhood and the level of deprivation of his or her adjacent neighbourhoods, I find that an increase in expected level of own deprivation lowers RS. However, locating in an area with a higher deprivation level than the deprivation level that an individual is expected to experience does not affect that person's residential satisfaction. Locating next to an affluent area raises the satisfaction of those people in the poor areas. Therefore, residents of a poor area enjoy a positive amenity effect when they choose to locate adjacent to a rich area. However, the residents of the rich area do not enjoy living near poor areas, which can be interpreted as a negative NIMBY effect.

All crowding and density results reflect those of previous chapters and are therefore shown to be robust to the inclusion of deprivation variables.

Table 48. A summary of the results of the 9th chapter.

Category	Result	Significance
Predicted deprivation level derived from backgrounds ($\widehat{\text{DEP}}$)	A higher level of predicted deprivation lowers RS.	Yes
Unexpected deprivation ($\text{DEP} - \widehat{\text{DEP}}$)	A higher level of unexpected deprivation does not affect RS.	No
Expected relative deprivation ($\widehat{\text{RDEP}}$)	If an individual lives in a poor area: a higher relative deprivation level raises his or her RS.	Yes
	If an individual live in a rich area: a higher relative deprivation level lowers his or her RS	Yes
Unexpected relative deprivation ($\text{RDEP} - \widehat{\text{RDEP}}$)	A higher level of unexpected relative deprivation for the residents of poor area raises RS.	Yes
	A higher level of unexpected relative deprivation for the residents of rich area lowers RS.	Yes
Absolute crowding derived from backgrounds ($\widehat{\text{PPBR}}$)	Higher absolute crowding stemming from 'own group' attraction behaviour does not affect RS.	No
Exogenous household crowding ($\text{PPBR} - \widehat{\text{PPBR}}$)	A higher level of unexpected crowding lowers RS.	Yes
Expected area density level ($\widehat{\text{DENSA}}$)	Higher area density level stemming from individual's backgrounds does not affect RS.	No
Exogenous area density ($\text{DENSA} - \widehat{\text{DENSA}}$)	A higher level of unexpected area density lowers RS.	Yes

Expected relative crowding ($\widehat{\mathrm{CVARS}}$)	A higher relative expected household crowding does not affect RS.	No
Unexpected relative crowding ($\mathrm{CVAR} - \widehat{\mathrm{CVARS}}$)	A higher level of unexpected relative crowding does not affect RS.	No
Expected relative density ($\mathrm{ATV}\widehat{\mathrm{ARD}}\mathrm{ENS}$)	Living in an area with a relatively higher expected density does not affect RS.	No
Unexpected relative density ($\mathrm{ATVARDENS} - \mathrm{ATV}\widehat{\mathrm{ARD}}\mathrm{ENS}$)	A higher level of unexpected relative density raises RS.	Yes
Perceived crowding (PC)	Higher PC level affects RS negatively.	Yes

11.2.5 Essay 5 – A Stata Command for Creating Spatial Weight Matrices

Researchers often face compatibility issues when combining statistical and geographical packages. Also, if there is a need for constructing different spatial outputs based on a varying parameter, using multiple software packages, such as a mix of ArcGIS and Stata, is very time-consuming. For example, in Chapter 8, reference groups are considered as neighbours living over different geographic scales, one of which is neighbours who may be visited in a 15-minute walk. In order to find the best definition for the dynamic neighbourhood boundaries based on the walking distance approach, different walking distances need to be tested. For instance, this parameter can be a two-minute walking distance or a 20-minute walking distance. Due to limitations imposed by various packages, most statistical studies are limited to simplified geographical analyses.

The memory space needed for creating spatial weights for spatial analysis is huge. The existing commands for creating spatial weights in Stata, such as spmat and spatwmat, have a limited matrix size; that is, they cannot handle matrices greater than the Stata's maximum matrix size. To overcome Stata's matrix size limitations, the analysis should be conducted using the Mata language. This enables researchers to create large matrices efficiently, in terms of the processing time.

Due to the lack of spatial tools in Stata, in order to overcome the trade-off between available statistical and geographical tools, I have written a Stata programme that I called -spwt-.The weighting options of **spwt** include the contiguity weights based on the k-nearest neighbour and the walking distance from the areas' centroids methods. This programme is outlined in Chapter 8.

11.3 Limitations of the current study

An important contribution of the current study is its attempt to distinguish between expected and unexpected levels of the variables of interest. The expected levels are estimated based on an individual's backgrounds. In other words, I have assumed that an individual's previous life experience affects his or her choice of dwelling and his or her expected levels of crowding, density, deprivation and relative position. This can be justified by the retrospective nature of our outcome of interest (as discussed in Chapter 4). An interesting question concerns how people substitute between their current consumption of space and their future consumption of it. For example, an individual may not live in a large house, hoping that he or she will live in a large house in the future. In this situation, the current level of consumption of space may be desirable as the individual is planning on living in a bigger house in the future. This point has not been fully discussed in the current thesis considering the cross-sectional nature of the dataset.

Expectations are also subject to change over time. An individual may be happy with a simple shelter when his or her city of residence is not prosperous. However, the same individual may dream of a luxurious house if his or her city of residence is affluent. In Auckland, for example, current cohorts of first-homebuyers may be less satisfied than previous cohorts with the homes available in their affordable price range.[116] This change in tastes and expectations can be an interesting question for a future study.

Another possibility for overstating satisfaction with residential environment is having a second home. An individual who lives in a crowded neighbourhood may be happy with his or her living condition as he or she may have a second home for vacations. This study's dataset does not provide us with any information about second homes. The individual answers the questions in relation to his or her current usual residence. Future studies may attempt to address this issue.

A number of reasons were provided in Section 2.2 to support the reliability of the residential satisfaction measure used in this study as a measure of respondents' satisfaction with both their dwelling and their neighbourhood. As noted in Section 3.3, a higher level of problems with both housing attributes and neighbourhood features is associated with a lower residential satisfaction. This confirms our interpretation of the residential satisfaction variable as being determined both by house-specific and neighbourhood-specific factors. Also, the study's finding that neighbourhood features significantly affect the outcome of interest suggests that people clearly think about their neighbourhood features, as well as their housing characteristics, when answering the question about how they feel about where they are currently living.[117] However, future studies may use another survey to check on the results of the current study in terms of its inclusion of both house-specific and neighbourhood characteristics for the outcome of interest (residential satisfaction).

Intertemporal substitution may affect the results of the current thesis. Individuals may not reflect their dissatisfaction with a higher crowding level as they plan to provide a better dwelling in the future. In other words, people may substitute between their consumption of space over their life cycles. This thesis accounts for the endogeneity of the variables of interest with respect to a number of instrumental variables, as well as age and other socioeconomic factors. Assuming that people with similar characteristics are planning for

---

[116] We do not have enough information to analyse this issue because of the lack of a time-series data on first home buyers' mortgages.

[117] In the presence of the measurement error in dependent variable, the results are not biased but there will be inflated standard errors for the variables of interest.

a similar situation, the current study partly accounts for people's current understanding of their situation relative to their future plans. Also, as discussed in Chapter 6, people reflect their crowding level in their perception of their current dwelling's crowding level. In other words, if our estimation of crowding's impact on RS would be biased, PC would capture the unbiased impact. The study also accounts for the age, income level and homeownership status of individuals, which may capture a part of the substitution between the current consumption of the factors of space and future consumption. Considering the limitations of this study's cross-sectional dataset, I am not able to precisely account for people's substitution of space consumption over their life cycle; this could be an interesting topic for a future panel-data study.

The first essay of this thesis found no significant difference between the estimated results using Auckland's sample population and the estimations derived from using New Zealand's sample population. However, due to computational limits, the subsequent essays could not verify the results using New Zealand's sample population. Future studies may use improved technology to verify the results of this thesis using a larger sample population.

The current thesis uses a two stage least squares approach for capturing the impact of social norms on people's choice of location. The list of instrument used for this analysis is limited to the variables available in the NZGSS dataset. While the validity of the instruments is validated using statistical tests, a future study may use another set of instruments, derived from a richer dataset, to check on this thesis's results.

The current study controls for the impact of ethnicity in its estimations of the impacts of the variables of interest on residential satisfaction. The estimated coefficients, however, may be significantly different for different ethnic groups. An assessment of the differences in estimated coefficients for different ethnic groups remains an interesting question for future research.

11.4 Suggestions to the data provider

This thesis sheds light on the role of households' expectations and relativities in their evaluation of their living environment. In addition to interpreting a new method for taking expectations into account, the thesis integrates spatial statistics, using geographic information systems, with individual-level datasets provided by SNZ. Due to the confidentiality of the individual-level data, the author needed to use facilities provided in SNZ's datalabs for accessing the full file datasets, i.e. the CURFs provided by SNZ do not contain all needed variables for the current study.

In datalabs, the access to geographic processing software packages, such as ArcGIS, is not provided. The research highlights the need for providing researchers with basic spatial analysis information, such as commonly used proximity matrices. The other issue that the researcher faced in his access to datasets was a huge inconsistency amongst different waves of a dataset. Fixing this issue takes long hours from a researcher while it could be processed and provided in the first place by the data provider.

Another suggestion for the data provider is to clarify the possibility of merges between datasets provided. Understanding the possibility of merging the NZGSS with Census was a very time-consuming process that could only be resolved after a number of meetings with Statistics New Zealand. This issue could be addressed by providing the researcher with the dimensions of the datasets through which a linkage between them is possible.

The NZGSS dataset needs to account for a wider range of the factors related to a specific survey question. For example, the survey provides information on household crowding, but it does not include any information on the importance of 'privacy' to the respondents.

11.5 Policy implications

Residential living conditions and residential satisfaction may be influenced by regulatory settings, including those related to planning restrictions. For instance, in order to contribute to a more sustainable urban development, there is growing interest in increasing the intensification of urban areas by using policy instruments (Howley et al., 2009).

Expansion of the city at the city fringe requires costly infrastructure upgrades. Therefore, one goal of Auckland Council is to encourage greater intensification in parts of Auckland. Many of the building regulations in the Auckland Plan aim for higher density in areas with relatively good transport connectivity. Auckland Council's chief economist discussed the need for intensification (Auckland Economic Quarterly, 2015). This is in line with a large proportion of the regulations introduced by the Proposed Auckland Unitary Plan (PAUP) that aim at intensification (Auckland Council, 2013; Grimes & Mitchell, 2015). The outcome of intensification may be a higher household crowding level – that is, an increase in the number of people per room[118] – in addition to an (intended) increase in population density (number of people per unit of area). However, prior to this study I was not aware of how crowding and density affect people's residential satisfaction.

Auckland Council's chief economist recognised the need for a better understanding of the society's residential preferences, saying that 'The most fundamental challenge is how to balance what communities judge is best for them with what is best for Auckland and the country. There is a need to relax regulatory controls on height and density, while retaining provisions for quality urban design. That will allow more dwellings per unit of land, more choices, and better housing affordability' (Parker, 2015).

The present thesis verifies the tradeoff between housing intensification and residential preferences. For a precise analysis of the impact of factors of intensification on residents' satisfaction, we need to know the level of intensification that an individual is likely to experience based on his or her characteristics. For example, while living in a dense area may not significantly affect the residential satisfaction of one person

[118] Currently, the population pattern in Auckland is also changing as the number of one- or two-person households is increasing (Statistics New Zealand, 2012a); this factor needs to be taken into account in the intensification policies.

who is likely to live in a dense area, it may affect the satisfaction of another person who is less keen on living in a dense area.

A policy question to be answered is whether residents enjoy locating in dense areas. To answer this question, I should consider both the satisfaction of people located in a denser area and the satisfaction of people located nearby the dense area. Living in a dense area may be satisfactory due to better accessibility to a range of facilities, such as libraries and restaurants. However, it may not be desirable due to higher congestion. This thesis concludes that living in a dense area, on average, lowers the satisfaction of residents. However, to know the impact of a higher density, we also need to know how living next to a dense area affects residents' satisfaction. They may enjoy accessing the facilities provided in their area, or may be negatively affected by congestion. The results in Chapter 9 indicate that:

- A higher level of unexpected density reduces RS while a higher level of unexpected relative density raises RS. The latter is a zero-sum game so the negative unexpected density effect predominates.

Based on the results, the negative impact of higher density levels is relatively smaller in less dense areas. On the other hand, the positive impact of living in dense areas is larger for the residents of the less dense areas. The positive impact of locating next to denser areas outweighs the negative impact of higher density levels for the areas with a density level less than the average level. Hence, density regulations may aim to provide higher density provisions nearby areas with lower levels of density. Also, considering that locating in a dense area has a higher positive marginal impact on the residents of the relatively denser areas, a continuous development of dense areas is more desirable to the whole population than a leapfrogging development.

A second policy question is: What is the impact of a higher household crowding?

Amongst policy makers, it is generally accepted that people have imaginary privacy circles around themselves. An individual may be annoyed if someone else stands too close to him or her. In the policy environment, the definition of 'too close' depends on an individual's background; that is, where and how he or she has been brought up. It is commonly believed that, for an estimation of the radius of their privacy circles, people look at the conventional distance between like others. Therefore, the third question, which needs to be answered alongside the second question, is: How does living in a relatively crowded house affect people's satisfaction?

The results of this study suggest that:

- A higher unexpected household crowding lowers residential satisfaction. Crowding is only considered undesirable when it is more than the level that she expects to have.

The last question is about locating in a poor (rich) area nearby affluent (poor) areas. The appropriate location for poor areas has always been a controversial issue. An example is John Key's statement that pepper-potting policies are 'economic vandalism' (Key, 2006). The question is: How does locating nearby an affluent (poor) area affect the residents of a poor (rich) area?

The results indicate that locating in a poor area near an affluent area increases the satisfaction of the residents of the poor area. On the other hand, this same policy lowers the residential satisfaction of the residents of rich areas. Therefore,

- Locating poor areas nearby rich neighbourhoods does not have a homogenous effect on the residential satisfaction across the city. If a policy aims to be in favour of the residents of poor neighbourhoods, it is advisable to locate the deprived neighbourhoods next to affluent areas.

Strategic advisers often look for a basis for their proposed plans. For example, many plans are based on the assumption that neighbourhood segregation reflects job locations. This study emphasises the importance of the exogeneity of the factors affecting segregation. Based on this study, people's backgrounds affect their understandings of their living environment significantly. This is observed in terms of their choice of residence location and, therefore, in their understanding of the desirable living environment. Consequently, policy makers and planners must take these person-specific preferences about crowding and density into account when considering the effect of their planning choices.

References

Abas, M. A., Vanderpyl, J., Robinson, E., Le Prou, T., & Crampton, P. 2006. Socio-economic deprivation and duration of hospital stay in severe mental disorder. ***The British Journal of Psychiatry***, 188(6): 581–582.

Amérigo, M. A., & Aragonés, J. I. 1997. A theoretical and methodological approach to the study of residential satisfaction. ***Journal of Environmental Psychology***, 17(1): 47–57.

Anderson, S. T., & West, S. E. 2006. Open space, residential property values, and spatial context. ***Regional Science and Urban Economics***, 36(6): 773–789.

Anselin, L. 1988. ***Spatial econometrics: methods and models***. Dordrecht: Luwer academic publishers.

Ansell, B. 2013. The Political Economy of Ownership: Housing Markets and the Welfare State. ***January. Available from: http :// www. polisci. umn. edu/~ansell/papers/Ansell Political Economy of Ownership Jan***, 202013.

Atkinson, R., & Kintrea, K. 2001. Disentangling area effects: evidence from deprived and non-deprived neighbourhoods. ***Urban Studies***, 38(12): 2277–2298.

Auckland Council, 2013. The Proposed Auckland Unitary Plan: Auckland Council.

Bache, I., & Reardon, L. 2013. An Idea Whose Time has Come? Explaining the Rise of Well-Being in British Politics. ***Political Studies***, 61(4): 898–914.

Baker, M. G., Goodyear, R., Telfar Barnard, L., & Howden-Chapman, P. 2006. ***The distribution of household crowding in New Zealand: An analysis based on 1991 to 2006 Census data***: Wellington.

Barbosa, O., Tratalos, J. A., Armsworth, P. R., Davies, R. G., Fuller, R. A., Johnson, P., & Gaston, K. J. 2007. Who benefits from access to green space? A case study from Sheffield, UK. ***Landscape and Urban Planning***, 83(2–3): 187–195.

Barnett, R., & Lauer, G. 2003. Urban deprivation and public hospital admissions in Christchurch, New Zealand, 1990–1997. ***Health & Social Care in the Community***, 11(4): 299–313.

Barro, R. J., & Gordon, D. B. 1981. A positive theory of monetary policy in a natural-rate model: National Bureau of Economic Research Cambridge, Mass., USA.

Baum, S., Arthurson, K., & Rickson, K. 2010. Happy people in mixed-up places: the association between the degree and type of local socioeconomic mix and expressions of neighbourhood satisfaction. ***Urban Studies***, 47(3): 467–485.

Bjørnskov, C. 2003. The happy few: Cross–country evidence on social capital and life satisfaction. ***Kyklos***, 56(1): 3–16.

Blume, L., & Durlauf, S. 2001. The interactions-based approach to socioeconomic behavior. In S. N. Durlauf, & H. P. Young (Eds.), ***Social Dynamics***: 15–44.

Boehm, T. P., & Schlottmann, A. M. 1999. Does home ownership by parents have an economic impact on their children? ***Journal of Housing Economics***, 8(3): 217–232.

Bolster, A., Burgess, S., Johnston, R., Jones, K., Propper, C., & Sarker, R. 2007. Neighbourhoods, households and income dynamics: a semi-parametric investigation of neighbourhood effects. ***Journal of Economic Geography***, 7(1): 1–38.

Bonnes, M., Bonaiuto, M., & Ercolani, A. P. 1991. Crowding and Residential Satisfaction in the Urban Environment A Contextual Approach. ***Environment and Behavior***, 23(5): 531–552.

Boyko, C. T., & Cooper, R. 2011. Clarifying and re-conceptualising density. ***Progress in Planning***, 76(1): 1–61.

Brock, W. A., & Durlauf, S. N. 2002. A multinomial-choice model of neighborhood effects. ***American Economic Review***: 298–303.

Brock, W. A., & Durlauf, S. N. 2004. ***Identification of binary choice models with social interactions***.

Brodsky, A. E., O'Campo, P. J., & Aronson, R. E. 1999. PSOC in community context: Multi-level correlates of a measure of psychological sense of community in low-income, urban neighborhoods. ***Journal of Community Psychology***, 27(6): 659–679.

Brueckner, J. K. 1997. Consumption and investment motives and the portfolio choices of homeowners. ***The Journal of Real Estate Finance and Economics***, 15(2): 159–180.

Burrows, S., Auger, N., Gamache, P., St-Laurent, D., & Hamel, D. 2011. Influence of social and material individual and area deprivation on suicide mortality among 2.7 million Canadians: A prospective study. ***BMC Public Health***, 11(1): 577.

Carlino, G. A., Chatterjee, S., & Hunt, R. M. 2007. Urban density and the rate of invention. ***Journal of Urban Economics***, 61(3): 389–419.

Carmon, N. 1999. Three generations of urban renewal policies: analysis and policy implications. ***Geoforum***, 30(2): 145–158.

Chan, Y.-K. 1999. Density, crowding, and factors intervening in their relationship: Evidence from a hyper-dense metropolis. ***Social Indicators Research***, 48(1): 103–124.

Chapman, D. W., & Lombard, J. R. 2006. Determinants of neighborhood satisfaction in fee-based gated and nongated communities. ***Urban Affairs Review***, 41(6): 769–799.

Cheshire, P. 2007. Segregated neighbourhoods and mixed communities. ***York: Joseph Rowntree Foundation***.

Cheshire, P., & Sheppard, S. 1998. Estimating the demand for housing, land, and neighbourhood characteristics. ***Oxford Bulletin of Economics and Statistics***, 60(3): 357–382.

Cheshire, P., & Sheppard, S. 2004. Capitalising the Value of Free Schools: The Impact of Supply Characteristics and Uncertainty*. ***The Economic Journal***, 114(499): F397–F424.

Cheshire, P. C., Nathan, M., & Overman, H. G. 2014. ***Urban economics and urban policy: Challenging conventional policy wisdom***: Edward Elgar Publishing.

Churchman, A. 1999. Disentangling the concept of density. ***Journal of Planning Literature***, 13(4): 389–411.

Clark, A. E. 2012. Happiness, habits and high rank: Comparisons in economic and social life: SOEP papers on multidisciplinary panel data research.

Clark, A. E., Westergård-Nielsen, N., & Kristensen, N. 2009. Economic satisfaction and income rank in small neighbourhoods. ***Journal of the European Economic Association***, 7(2–3): 519–527.

Clark, J., & Kim, B. 2012. The effect of neighbourhood diversity on volunteering: Evidence from New Zealand. ***BE Journal of Economic Analysis and Policy***, 12(1).

Clark, W. A., & Morrison, P. S. 2012. Socio-spatial mobility and residential sorting: evidence from a large-scale survey. ***Urban Studies***, 49(15): 3253–3270.

Coleman, A. M. G., & Scobie, G. M. 2009. A simple model of housing rental and ownership with policy simulations: Motu Economic and Public Policy Research.

Cook, C., & Bruin, M. 1994. Determinants of housing quality: A comparison of white, African-American, and Hispanic single-parent women. ***Journal of Family and Economic Issues***, 15(4): 329–347.

Court, A. T. 1939. Hedonic price indexes with automotive examples. ***The dynamics of Automobile Demand***.

Cox, K. R. 1982. Housing tenure and neighborhood activism. ***Urban Affairs Review***, 18(1): 107–129.

Daneshvary, N., Daneshvary, R., & Schwer, R. K. 1998. Solid-waste recycling behavior and support for curbside textile recycling. ***Environment and Behavior***, 30(2): 144–161.

De Graaff, T., & Van Leuvensteijn, M. 2011. A European Cross-Country Comparison of the Impact of Homeownership and Transaction Costs on Job Tenure. ***Regional Studies*** (ahead-of-print): 1–19.

Deiner, E., Lucas, R., & Oishi, S. 2002. Subjective well-being: The science of happiness and life satisfaction. ***Handbook of Positive Psychology***: 63–73.

Dekker, K., de Vos, S., Musterd, S., & van Kempen, R. 2011. Residential Satisfaction in Housing Estates in European Cities: A Multi-level Research Approach. ***Housing Studies***, 26(04): 479–499.

Diaz-Serrano, L. 2009. Disentangling the housing satisfaction puzzle: Does homeownership really matter? ***Journal of Economic Psychology***, 30(5): 745–755.

Dıaz, A., & Luengo-Prado, M. J. 2010. Homeownership: Economic Benefits. The International Encyclopedia of Housing and Home.

Diener, E. 2009. ***Well-being for public policy***: Oxford University Press, USA.

Diener, E., Sandvik, E., Seidlitz, L., & Diener, M. 1993. The relationship between income and subjective well-being: Relative or absolute? ***Social Indicators Research***, 28(3): 195–223.

Dietz, R. D., & Haurin, D. R. 2003. The social and private micro-level consequences of homeownership. ***Journal of Urban Economics***, 54(3): 401–450.

DiPasquale, D., & Glaeser, E. L. 1999. Incentives and social capital: are homeowners better citizens? ***Journal of Urban Economics***, 45(2): 354–384.

Dolan, P., Layard, R., & Metcalfe, R. 2011. Measuring subjective well-being for public policy.

Dowling, J. M., & Yap, C.-F. 2007. ***Modern developments in behavioral economics: social science perspectives on choice and decision making***: World Scientific.

Drukker, D. M., Peng, H., Prucha, I. R., & Raciborski, R. 2013. Creating and managing spatial-weighting matrices with the spmat command. ***Stata Journal***, 13(2): 242–286.

Duesenberry, J. S. 1949. Income, saving, and the theory of consumer behavior.

Duranton, G., & Overman, H. G. 2005. Testing for localization using micro-geographic data. ***The Review of Economic Studies***, 72(4): 1077–1106.

Durlauf, S. N. 2004. Chapter 50 Neighborhood effects. In J. V. Henderson, & T. Jacques-François (Eds.), ***Handbook of Regional and Urban Economics***, Vol. Volume 4: 2173–2242: Elsevier.

Easterlin, R. A. 1974. Does economic growth improve the human lot? Some empirical evidence. ***Nations and Households in Economic Growth***, 89: 89–125.

Easterlin, R. A. 1995. Will raising the incomes of all increase the happiness of all? ***Journal of Economic Behavior & Organization***, 27(1): 35–47.

Eichholtz, P., & Lindenthal, T. 2014. Demographics, human capital, and the demand for housing. ***Journal of Housing Economics***, 26: 19–32.

Ellaway, A., Macintyre, S., & Bonnefoy, X. 2005. Graffiti, greenery, and obesity in adults: secondary analysis of European cross sectional survey. ***BMJ: British Medical Journal***, 331(7517): 611.

Engelhardt, G. V. 1996a. Consumption, down payments, and liquidity constraints. ***Journal of Money, Credit and Banking***, 28(2): 255–271.

Engelhardt, G. V. 1996b. House prices and home owner saving behavior. ***Regional Science and Urban Economics***, 26(3): 313–336.

Engelhardt, G. V., Eriksen, M. D., Gale, W. G., & Mills, G. B. 2010. What are the social benefits of homeownership? Experimental evidence for low-income households. ***Journal of Urban Economics***, 67(3): 249–258.

Erdős, P., & Rényi, A. 1960. On the evolution of random graphs. ***Magyar Tud. Akad. Mat. Kutató Int. Közl***, 5: 17–61.

Ewing, R., & Rong, F. 2008. The impact of urban form on US residential energy use. ***Housing Policy Debate***, 19(1): 1–30.

Fabling, R., Grimes, A., & Stevens, P. 2012. The relatives are fine: use of qualitative firm data in economic analysis. ***Applied Economics Letters***, 19(7): 615–618.

Ferrer-i-Carbonell, A., & Frijters, P. 2004. How Important is Methodology for the estimates of the determinants of Happiness? ***The Economic Journal***, 114(497): 641–659.

Fisher, J. D., & Gervais, M. 2011. Why has home ownership fallen among the young? ***International Economic Review***, 52(3): 883–912.

Fleming, M. M. 2004. Techniques for estimating spatially dependent discrete choice models, ***Advances in Spatial Econometrics***, 145–168: Springer.

Fleurbaey, M. 2009. Beyond GDP: The quest for a measure of social welfare. ***Journal of Economic Literature***: 1029–1075.

Flint, J., & Rowlands, R. 2003. Commodification, normalisation and intervention: cultural, social and symbolic capital in housing consumption and governance. ***Journal of Housing and the Built Environment***, 18(3): 213–232.

Frank, L. D., Kershaw, S. E., Chapman, J. E., Campbell, M., & Swinkels, H. M. 2014. The unmet demand for walkability: Disparities between preferences and actual choices for residential environments in Toronto and Vancouver. ***Canadian Journal of Public Health***, 106(1): eS12–eS20.

Frank, R. H. 1985. ***Choosing the right pond: Human behavior and the quest for status***: Oxford University Press.

Frey, B. S., & Stutzer, A. 2002. What can economists learn from happiness research? ***Journal of Economic Literature***, 40(2): 402–435.

Friedman, M. 1953. The methodology of positive economics. in Essays in Positive Economics. University of Chicago Press, 1966: 3-16, 30-43.

Galbraith, J. K., & Galbraith, J. K. 1984. The affluent society. Fortieth Anniversary Edition. Houghton Mifflin Company: New York, 1998.

Galster, G., & Hedman, L. 2013. Measuring Neighbourhood Effects Non-experimentally: How Much Do Alternative Methods Matter? ***Housing Studies***, 28(3): 473–498.

Gerkman, L. 2011. Empirical spatial econometric modelling of small scale neighbourhood. ***Journal of Geographical Systems***: 1–16.

Gibbons, S., Silva, O., & Weinhardt, F. 2013. Everybody Needs Good Neighbours? Evidence from Students' Outcomes in England. ***The Economic Journal***.

Gilderbloom, J. I., & Markham, J. P. 1995. The impact of homeownership on political beliefs. ***Social Forces***, 73(4): 1589–1607.

Glaeser, E. L. 2010. Agglomeration economics. University of Chicago Press.

Glaeser, E. L., & Gottlieb, J. D. 2009. The wealth of cities: Agglomeration economies and spatial equilibrium in the United States: National Bureau of Economic Research.

Glaeser, E. L., Gottlieb, J. D., & Ziv, O. 2014. Unhappy cities: National Bureau of Economic Research.

Glaeser, E. L., & Kahn, M. E. 2001. Decentralized employment and the transformation of the American city (No. w8117). National Bureau of Economic Research.

Glaeser, E. L., & Sacerdote, B. 1996. Why is there more crime in cities? **Journal of political economy**, 107(S6), S225-S258.

Glaeser, E. L., Sacerdote, B., & Scheinkman, J. A. 1996. Crime and social interactions. ***Quarterly Journal of Economics***, 111(2): 507–548.

Goodyear, R., & Fabian, A. 2012. Household crowding in New Zealand compared with selected countries: Wellington: Statistics New Zealand ISBN.

Goodyear, R., Fabian, A., & Hay, J. 2011. Finding the crowding index that works best for New Zealand. ***Statistics New Zealand Working Paper*** (11-04).

Gourinchas, P. O., & Parker, J. A. 2002. Consumption over the life cycle. ***Econometrica***, 70(1): 47–89.

Green, R. K., & Hendershott, P. 2001. Home-ownership and the duration of unemployment: a test of the Oswald hypothesis. ***NBER Working paper***.

Greene, W. H., & Hensher, D. A. 2010. Modeling ordered choices: a primer: Cambridge University Press.

Grimes, A., & Aitken, A. 2006. What's the beef with house prices? Economic shocks and local housing markets. ***Economic Shocks and Local Housing Markets*** (October 2004).

Grimes, A., Kerr, S., Aitken, A., & Sourell, R. 2006. The housing fulcrum: Balancing economic and social factors in housing research and policy. ***Kōtuitui: New Zealand Journal of Social Sciences Online***, 1(1): 65–79.

Grimes, A., & Liang, Y. 2009. Spatial Determinants of Land Prices: Does Auckland's Metropolitan Urban Limit Have an Effect? ***Applied Spatial Analysis and Policy***, 2(1): 23–45.

Grimes, A., & Mitchell, I. 2015. Impacts of Planning Rules, Regulations, Uncertainty and Delay on Residential Property Development. ***Motu Economic and Public Policy Research***, Motu Working Paper 15–02.

Grimes, A., Oxley, L., & Tarrant, N. 2012. Does Money Buy Me Love? Testing Alternative Measures of National Wellbeing.

Grimes, A., & Reinhardt, M. G. 2015. Relative Income and Subjective Wellbeing: Intra-national and International Comparisons by Settlement and Country Type. ***Available at SSRN***.

Grimes, A., & Young, C. 2013. Spatial effects of urban rail upgrades. ***Journal of Transport Geography***, 30: 1–6.

Guest, A. M., & Wierzbicki, S. K. 1999. Social Ties at the Neighborhood Level Two Decades of GSS Evidence. ***Urban Affairs Review***, 35(1): 92–111.

Halpern, D. 2005. Social Capital. Malden: USA: Polity Press.

Hamermesh, D. S. 2004. Subjective outcomes in economics. In NBER Working paper 10361. Cambridge, MA: National Bureau of Economic Research.

Hancock, R. 1998. Housing wealth, income and financial wealth of older people in Britain. ***Ageing and Society***, 18(01): 5–33.

Hansen, L. P. 1982. Large sample properties of generalized method of moments estimators. ***Econometrica: Journal of the Econometric Society***: 1029–1054.

He, S., & Wu, F. 2007. Socio-spatial impacts of property-led redevelopment on China's urban neighbourhoods. ***Cities***, 24(3): 194–208.

Helderman, A. C., Mulder, C. H., & Ham, M. 2004. The changing effect of home ownership on residential mobility in the Netherlands, 1980–98. ***Housing Studies***, 19(4): 601–616.

Helliwell, J., Layard, R., & Sachs, J. 2012. World happiness report.

Helliwell, J. F. 2003. How's life? Combining individual and national variables to explain subjective well-being. ***Economic Modelling***, 20(2): 331–360.

Hilber, C. A. 2010. New housing supply and the dilution of social capital. ***Journal of Urban Economics***, 67(3): 419–437.

Howley, P., Scott, M., & Redmond, D. 2009. Sustainability versus liveability: an investigation of neighbourhood satisfaction. ***Journal of Environmental Planning and Management***, 52(6): 847–864.

Immergluck, D. 1998. Progress confined: Increases in Black home buying and the persistence of residential segregation. ***Journal of Urban Affairs***, 20(4): 443–457.

Ingersoll, R. 1992. The disappearing suburb. ***Design Book Review***, 26.

Ioannides, Y. M. 2002. Residential neighborhood effects. ***Regional Science and Urban Economics***, 32(2): 145–165.

Johnston, R., M. Pousen, and J. Forrest. 2009. Evaluating changing residential segregation in Auckland, New Zealand, using spatial statistics. Bristol: The Centre for Market and Public Organisation, Bristol University.

Kearns, A., & Parkes, A. 2003. Living in and leaving poor neighbourhood conditions in England. ***Housing Studies***, 18(6): 827–851.

Kenny, D. A. 2012. Measuring model fit. Retrieved from http://www.davidakenny.net/cm/fit.htm.

Key, J. 2006. Scoop Talks To Key About Tax Cuts. In K. List (Ed.), ***Audio 'N' Transcript***.

Kingdon, G. G., & Knight, J. 2007. Community, comparisons and subjective well-being in a divided society. ***Journal of Economic Behavior & Organization***, 64(1): 69–90.

Kingston, P. W., Thompson, J. L., & Eichar, D. M. 1984. The politics of homeownership. ***American Politics Research***, 12(2): 131–150.

Knight, J., & Gunatilaka, R. 2010. Great Expectations? The Subjective Well-being of Rural–Urban Migrants in China. ***World Development***, 38(1): 113–124.

Knowles, E. S., & Nathan, K. T. 1997. Acquiescent responding in self-reports: Cognitive style or social concern? ***Journal of Research in Personality***, 31(2): 293–301.

Knox, P., & Pinch, S. 2009. ***Urban social geography: an introduction*** (6th ed.): Pearson Education Limited.

Kőszegi, B., & Rabin, M. 2007. Mistakes in choice-based welfare analysis. ***The American Economic Review***, 97(2): 477–481.

Kott, P. S. 2001. The delete-a-group jackknife. ***Journal of Official Statistics***, 17(4): 521–526.

Krupka, D. J. 2007. Are big cities more segregated? Neighbourhood scale and the measurement of segregation. ***Urban Studies***, 44(1): 187–197.

Larwood, L., & Whittaker, W. 1977. Managerial myopia: Self-serving biases in organizational planning. ***Journal of Applied Psychology***, 62(2): 194.

Layard, R. 2003. Has social science a clue?: What is happiness? Are we getting happier. ***Lionel Robbins Memorial Lecture Series***: 03–05.

Leach, H. 2014. ***The New Zealand kitchen in the 20th century***: Otago Univeristy Press.

Li, H., Calder, C. A., & Cressie, N. 2007. Beyond Moran's I: testing for spatial dependence based on the spatial autoregressive model. ***Geographical Analysis***, 39(4): 357–375.

Lipset, S. M. 1960. ***Revolution and counterrevolution: Change and persistence in social structures***: Transaction Publishers.

Lu, M. 1999. Determinants of Residential Satisfaction: Ordered Logit vs. Regression Models. ***Growth and Change***, 30(2): 264–287.

Lucas, R. E., Diener, E., & Suh, E. 1996. Discriminant validity of well-being measures. ***Journal of Personality and Social Psychology***, 71(3): 616.

Luigino, B., & Porta, P. 2007. Handbook on the Economics of Happiness: Cheltenham: Edward Elgar.

Luttmer, E. F. P. 2005. Neighbors as negatives: relative earnings and well-being. ***The Quarterly Journal of Economics***: 963–1002.

Maani, S. A. 2004. Why have Maori relative income levels deteriorated over time? ***Economic Record***, 80(248): 101–124.

Maani, S. A., Vaithianathan, R., & Wolfe, B. L. 2006. Inequality and health: Is housing crowding the link?: Motu Economic and Public Policy Research.

Manski, C. F. 1987. Semiparametric analysis of random effects linear models from binary panel data. ***Econometrica: Journal of the Econometric Society***: 357–362.

Manski, C. F. 1990. Nonparametric bounds on treatment effects. ***The American Economic Review***, 80(2): 319–323.

Manski, C. F. 1993. Identification of endogenous social effects: The reflection problem. ***The Review of Economic Studies***, 60(3): 531–542.

Manski, C. F. 2000. Economic Analysis of Social Interactions. ***National Bureau of Economic Research Working Paper Series***, No. 7580.

Manski, C. F., & Pepper, J. V. 2013. Deterrence and the Death Penalty: Partial Identification Analysis Using Repeated Cross Sections. ***Journal of Quantitative Criminology***, 29(1): 123–141.

Marans, R. W., & Stimson, R. 2011. ***An overview of quality of urban life***: Springer.

Mare, D., & Coleman, A. 2011. Estimating the determinants of population location in Auckland.

Maré, D. C., Coleman, A., & Pinkerton, R. 2011. Patterns of population location in Auckland.

Mayer, C. 1993. Taxes, income distribution, and the real estate cycle: Why all houses do not appreciate at the same rate. ***New England Economic Review*** (May): 39–50.

Mayer, C. J., & Engelhardt, G. V. 1996. Gifts, down payments, and housing affordability. ***Journal of Housing Research***, 7: 59–78.

McCarthy, G. W., Van Zandt, S., & Rohe, W. M. 2001. The economic benefits and costs of homeownership: A critical assessment of the research: Research Institute for Housing America.

Margaret McClure, 'Auckland region - Government, education and health', Te Ara - the Encyclopedia of New Zealand, http://www.TeAra.govt.nz/en/map/30852/auckland-local-government-1989-2010 (accessed 15 Augutst 2016).

Meen, G., Nygaard, C., & Meen, J. 2013. The Causes of Long-Term Neighbourhood Change, ***Understanding Neighbourhood Dynamics***: 43–62: Springer.

Menard, S. 2002. Applied logistic regression analysis (No. 106). Sage.

Mercer. 2015. Full 2015 rankings, ***Mercer survey***.

MfE. 2010. Resource Management Act 1991 (RMA), ***Ministry for the Environment***.

Moffitt, R. A. 2001. Policy interventions, low-level equilibria, and social interactions. ***Social Dynamics***: 45–82.

Moon, G., Pearce, J., & Barnett, R. 2012. Smoking, Ethnic Residential Segregation, and Ethnic Diversity: A Spatio-temporal Analysis. ***Annals of the Association of American Geographers***, 102(5): 912–921.

Morrison, P. 2014a. Reviews: Investigating quality of urban life: theory, methods, and empirical research. ***Environment and Planning B: Planning and Design***, 41(2): 379–380.

Morrison, P. 2015a. Who cares about income inequality? ***Policy Quarterly***, 11(1): 56–62.

Morrison, P. S. 1994. Housing occupancy and the changing size of households and dwellings in New Zealand 1951–1991. ***New Zealand Population Review***, 20(1–2): 69–100.

Morrison, P. S. 2011. Local expressions of subjective well-being: The New Zealand experience. ***Regional Studies***, 45(8): 1039–1058.

Morrison, P. S. 2014b. The measurement of regional growth and wellbeing, ***Handbook of regional science***: 277–289: Springer.

Morrison, P. S. 2015b. The inequaliy debate, the neglected role of residential sorting. ***The Policy Quarterly***, 11 (2): 72–79.

Morrison, P. S., & Nissen, K. 2010. Moving in and out of areas of deprivation: evidence from the New Zealand census. ***New Zealand Population Review***, 36: 55.

Morrison, P. S., & Weijers, D. 2012. Well-being in Wellington. ***Policy Quarterly***, 8(4): 51–55.

Moshrefi, M., Kim, J., & Torshizian, E. 2015. The Significance of the impact of MUL on the house prices in the CBD. mimeo.

Mundlak, Y. 1978. On the pooling of time series and cross section data. ***Econometrica: Journal of the Econometric Society***: 69–85.

Nettleton, S., & Burrows, R. 1998. Mortgage debt, insecure home ownership and health: an exploratory analysis. ***Sociology of Health & Illness***, 20(5): 731–753.

Nieuwenhuis, J., Völker, B., & Flap, H. 2013. A Bad Neighbour is as Great a Plague as a Good One is a Great Blessing: On Negative Relationships between Neighbours. ***Urban Studies***.

OECD. 2015. OECD Economic Surveys, New Zealand.

Oswald, A. J. 1996. A conjecture on the explanation for high unemployment in the industrialized nations: part 1. Warwick Economic Research Papers, No. 475. University of Warwick, Conventry, United Kingdom.

Palmer, E. M. 1985. Graphical evolution: an introduction to the theory of random graphs. John Wiley & Sons, Inc..

Parker, C. 2015. No, Mike, house crisis is not our fault, ***NZ Herald***. Auckland.

Pavot, W., Diener, E., Colvin, C. R., & Sandvik, E. 1991. Further validation of the satisfaction with life scale: Evidence for the cross-method convergence of well-being measures. ***Journal of Personality Assessment***, 57(1): 149–161.

Pearce, J., & Dorling, D. 2006. Increasing geographical inequalities in health in New Zealand, 1980–2001. ***International Journal of Epidemiology***, 35(3): 597–603.

Pearson, A. L., Griffin, E., Davies, A., & Kingham, S. 2012. An ecological study of the relationship between socioeconomic isolation and mental health in the most deprived areas in Auckland, New Zealand. ***Health & Place***.

Pinkerton, R. M., D. C. Mare, and J. Poot. 2011. Immigrant segregation in Auckland, New Zealand. In *Migration, economic change, social challenges, April 6-9, 2011*.

Pingle, M., & Mitchell, M. 2002. What motivates positional concerns for income? ***Journal of Economic Psychology***, 23(1): 127–148.

Pisati, M. 2008. SPMAP: Stata module to visualize spatial data. ***Statistical Software Components***.

Pollakowski, H. O., Stegman, M. A., & Rohe, W. 1991. Rates of Return on Housing of Low-and Moderate-Income Owners. ***Real Estate Economics***, 19(3): 417–425.

Posthumus, H., Bolt, G., & van Kempen, R. 2013. Why do Displaced Residents Move to Socioeconomically Disadvantaged Neighbourhoods? ***Housing Studies***, 28(2): 272–293.

Quercia, R. G., McCarthy, G. W., Ryznar, R. M., & Can Talen, A. 2000. Spatio-temporal measurement of house price appreciation in underserved areas. ***Journal of Housing Research***, 11(1): 1–28.

Robbins, L. 1932. The nature and significance of economic science. ***The Philosophy of Economics: An Anthology***: 73–99.

Robson, B., Lymperopoulou, K., & Rae, A. 2008. People on the move: exploring the functional roles of deprived neighbourhoods. **Environment and planning** A 40.11 (2008): 2693-2714.

Rodgers, W. 1982. Density, crowding, and satisfaction with the residential environment. ***Social Indicators Research***, 10(1): 75–102.

Rosen, S. 1974. Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. ***Journal of Political Economy***, 82(1): 34–55.

Roskruge, M., Grimes, A., McCann, P., & Poot, J. 2012. Social Capital and Regional Social Infrastructure Investment Evidence From New Zealand. ***International Regional Science Review***, 35(1): 3–25.

Roskruge, M., Grimes, A., McCann, P., & Poot, J. 2013. Homeownership, Social Capital and satisfaction with Local Government. ***Urban studies*** 50.12 (2013): 2517-2534.

Salmon, P., & Shand, D. 2009. Royal commission on Auckland governance report. ***Royal Commission, New Zealand***.

Salmond, C., Crampton, P., & Atkinson, J. 2007. NZDep2006 Index of Deprivation. Wellington.

Samuelson, P. A. 1938. A note on the pure theory of consumer's behaviour. ***Economica***, 5(17): 61–71.

Schluter, P., Carter, S., & Kokaua, J. 2007. Indices and perception of crowding in Pacific households domicile within Auckland, New Zealand: findings from the Pacific Islands Families Study. ***New Zealand Medical Journal***, 26.

Sen, A. K. 1977. Rational fools: A critique of the behavioral foundations of economic theory. ***Philosophy & Public Affairs***: 317–344.

Sengupta, N. K., Luyten, N., Greaves, L. M., Osborne, D., Robertson, A., Armstrong, G., & Sibley, C. G. (2013). Sense of community in New Zealand neighbourhoods: A multi-level model predicting social capital. **New Zealand Journal of Psychology**, 42(1), 36-45.

Shaw, M. 1999. ***The widening gap: health inequalities and policy in Britain***: The Policy Press.

Sheiner, L. 1995. Housing prices and the savings of renters. ***Journal of Urban Economics***, 38(1): 94–125.

Sheppard, S. 1999. Chapter 41 Hedonic analysis of housing markets. In C. Paul, & S. M. Edwin (Eds.), ***Handbook of Regional and Urban Economics***, 3: 1595–1635: Elsevier.

Sinai, T., & Souleles, N. S. 2005. Owner-occupied housing as a hedge against rent risk. ***The Quarterly Journal of Economics***, 120(2): 763–789.

Singh, G. K., & Siahpush, M. 2006. Widening socioeconomic inequalities in US life expectancy, 1980–2000. ***International Journal of Epidemiology***, 35(4): 969–979.

Sloan, M. 2013. ***The post-move satisfaction of individuals moving within New Zealand***. Victoria University of Wellington.

Sloan, M. K., & Morrison, P. S. 2015. The post-move satisfaction of internal migrants. ***Migration Studies***, 4(1), 97-115.

Smith, L. B., & Ho, M. H. 1996. The relative price differential between higher and lower priced homes. ***Journal of Housing Economics***, 5(1): 1–17.

Solnick, S. J., & Hemenway, D. 2005. Are positional concerns stronger in some domains than in others? ***The American Economic Review***, 95(2): 147–151.

Song, Y., & Knaap, G.-J. 2003. New urbanism and housing values: a disaggregate assessment. ***Journal of Urban Economics***, 54(2): 218–238.

Statistics New Zealand. 2006. 2006 Census of Population and Dwellings: QuickStats 2. Statistics New Zealand.

Statistics New Zealand. 2009. New Zealand general social survey: 2008. Wellington.

Statistics New Zealand. 2011. New Zealand General Social Survey 2010. ***Wellington: Statistics New Zealand. http://www. stats. govt. nz/browse_for_stats/people_and_communities/Households/nzgss_HOTP2008/Tables. aspx***.

Statistics New Zealand. 2012a. ***Demographic trends: 2011***: Wellington: Statistics New Zealand ISSN.

Statistics New Zealand. 2012b. ***Ethnicity and crowding: A detailed examination of crowding among ethnic groups in New Zealand 1986–2006***: Wellington: Statistics New Zealand ISBN.

Statistics New Zealand. 2013. New Zealand general social survey: 2012. ***Wellington***.

Stokols, D. 1972. A social-psychological model of human crowding phenomena: some implications for environmental research and design. **Journal of the American Institute of Planners** 38:72-83.

Tella, R. D., & MacCulloch, R. 2007. Happiness, contentment and other emotions for central banks (No. w13622). National Bureau of Economic Research.

Tella, R. D., MacCulloch, R. J., & Oswald, A. J. 2003. The macroeconomics of happiness. ***Review of Economics and Statistics***, 85(4): 809–827.

Torshizian, E. 2015a. Are We Limiting Regional Development by Requiring Planning Permissions?, ***North American Regional Science Conference***.

Torshizian, E. 2016a. Affordability Decompositions for Auckland. Presented to NZ Association of Economists conference, Auckland.

Torshizian, E. 2016b. A Positive Expected Future Rent or A Negative Supply Effect? A Study of Auckland's MUL Expansion. Presented to NZ Association of Economists conference, Auckland.

Torshizian, E. 2017. Effects of Crowding, Density and Deprivation on Residential Satisfaction. PhD Dissertation. The University of Auckland.

Troy, A., & Grove, J. M. 2008. Property values, parks, and crime: A hedonic analysis in Baltimore, MD. ***Landscape and Urban Planning***, 87(3): 233–245.

Van Ham, M., & Feijten, P. 2008. Who wants to leave the neighbourhood? The effect of being different from the neighbourhood population on wishes to move. ***Environment and Planning. A***, 40(5): 1151.

Veblen, T. 1899. The theory ofthe leisure class. ***New York: The New American Library***.

Veenhoven, R. 1994. Is happiness a trait? ***Social indicators research***, 32(2): 101–160.

Viton, P. A. 2010. Notes on spatial econometric models. ***City and Regional Planning***, 870(03).

Wang, X., & Maani, S. 2012. Assimilation, ethnic capital and earnings of immigrants in New Zealand: a spatial approach, ***International Conference: Futures for the metropolis: Pathways for development in Auckland and beyond***, Vol. Session 7. Massey University, Auckland.

White, H. 1980. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. ***Econometrica: Journal of the Econometric Society***: 817–838.

White, P., Gunston, J., Salmond, C., Atkinson, J., & Crampton, P. 2008. Atlas of Socioeconomic Deprivation in New Zealand NZDep2006. Wellington.

Williams, K., Jenks, M., & Burton, E. (2000). Achieving sustainable urban form. ***Taylor & Francis***.

Winkler, H. 2011. The effect of homeownership on geographic mobility and labor market outcomes. ***Manuscript, University of California Los Angeles***.

Witten, K., McCreanor, T., & Kearns, R. 2003. The place of neighbourhood in social cohesion: Insights from Massey, West Auckland. ***Urban Policy and Research***, 21(4): 321–338.

Wooldridge, J. 2015. ***Introductory econometrics: A modern approach***: Nelson Education.

Wu, F., & He, S. 2005. Changes in traditional urban areas and impacts of urban redevelopment: a case study of three neighbourhoods in Nanjing, China. ***Tijdschrift voor economische en sociale geografie***, 96(1): 75–95.

Yeabsley, J., & Stephenson, J. 2012. Grow for it – How population policies can promote economic growth. ***NZIER working paper 2012/1***.

Yoshikawa, H., & Ohtaka, F. 1989. An analysis of female labor supply, housing demand and the saving rate in Japan. ***European Economic Review***, 33(5): 997–1023.

Zheng, G. 2013. The effect of Auckland's Metropolitan Urban Limit on land prices. ***New Zealand Productivity Commission. Available online at http://www. productivity. govt. nz/sites/default/files/research-note-mar-13-auckland-mul. pdf***.

Glossary

Abbreviation	Name	Definition (if different from Name)
	Residential density	$= \frac{\text{Number of dwellings in an area}}{\text{unit of area}}$
	Objective crowding measure	Refers to the measures of crowding that do not account for the perception of crowding (PC).
	Developable land	Land available for construction (excluding parks and roads).
A		
AC		$\text{AC}_{ij} = \frac{\text{PPBR of the observation i}}{\text{Average PPBR of observation i's AU}} = \frac{C_i}{\overline{CA_j}}$
ACI	American Crowding Index	
AHC	Comparison between individual's crowding level and the crowding of symbols of success living in the same AU	Compares an individual's PPBR with the average crowding of individuals with an income at the 90th percentile in his or her AU.
AHVARC	Comparison between individual's crowding level and the crowding of symbols of success living in the same AU	Compares an individual's PPBR with the average crowding of individuals with an income at the 90th percentile in his or her AU. A variation method is used for constructing this measure.
AHVARC95	Comparison between individual's crowding level and the crowding of symbols of success living in the same AU.	Compares an individual's PPBR with the average crowding of individuals with an income at the 95th percentile in his or her AU. A variation method is used for constructing this measure.
AIC	Akaike information criterion	
APPBR	Adjusted PPBR	In the denominator of PPBR calculations, couples are considered as one person rather than two.
ATDENS		If people, in their evaluation of their living environment, compare the density level of the area unit that they are located in with the density level at their territorial authority, the constructed relative density variable will be called ATDENS.
ATVARDENS	Relative area density	$\text{ATVARDENS}_i = \frac{\text{DENSA}_i - \text{DENST}_i}{\sigma(\text{DENST})}$
$\text{ATV}\widehat{\text{ARD}}\text{ENS}$	Expected relative density	

$ATVARDENS - \widehat{ATVARDENS}$	Unexpected relative density	
AU	Area-unit	The larger (than MB) neighbourhoods are defined as area units, which approximate city suburbs (as opposed to administrative boundaries). Most people account themselves as a member of their suburb or area unit.
AVARC		$AVARC_i = \frac{C_i - CA_i}{\sigma(CA)}$
B		
BIC	Bayesian Information Criterion	
C		
C_i	Household crowding of individual i	PPBR of the observation i
CA		Average Crowding level in AUs
CBD	Central Business District	
CBD	Central Business District	
CC	Crowding Component	Crowding measure calculated by using a principal component approach. This is a summary of PPBR, PC and CNOS measures.
CM		Average crowding level in MBs
CNOS	Canadian National Occupation Standard index	
CT		Average crowding level in TLAs
$CVAR - \widehat{CVARS}$	Unexpected relative crowding	
CVARS	Relative crowding	
$\widehat{CVARS}$	Expected relative crowding	
D		
DENS	Area density	$= \frac{\text{Population size}}{\text{unit of area}}$
DENSA	AU's area density	$= \frac{\text{Population size in AU}}{\text{unit of area}}$
$\widehat{DENSA}$	Absolute area density derived from backgrounds	

$\mathrm{DENSA} - \widehat{\mathrm{DENSA}}$	Exogenous area density	
$\mathrm{DENSA} - \widehat{\mathrm{DENSA}}$		
DENSD		DENSD is the area density measure constructed based on the developable land area.
DENSM	Mesh-block's area density	$= \frac{\text{Population size in MB}}{\text{unit of area}}$
DENST	TLA's area density	$= \frac{\text{Population size in TLA}}{\text{unit of area}}$
DEP	Deprivation level	Deprivation levels of MBs
$\widehat{\mathrm{DEP}}$	Predicted deprivation level derived from backgrounds	
DUMMY	dummy variable	= 1 (RDEP > 0); 0 (Otherwise).
G		
GDP	Gross domestic product	
GOF	Goodness of fit	Measures the variability in the outcome that has been explained by the model
GSS	General Social Survey	
H		
HIR	Housing improvement regulations	
M		
MADENS		If people, in their evaluation of their living environment, compare the density level of the MB that they are located in with the density level at their area unit (AU), the constructed relative density variable will be called ATDENS.
MAVARDENS		$\mathrm{MAVARDENS_i} = \frac{\mathrm{DENSM_i} - \mathrm{DENSA_i}}{\sigma(\mathrm{DENSA})}$
MB	Mesh-block	Small neighbourhood, which consists of a few houses located next to each other, with an average population of 50 people.
MC		$\mathrm{MC_{ij}} = \frac{\text{PPBR of the observation i}}{\text{Average PPBR of observation i's MB}} = \frac{\mathrm{C_i}}{\mathrm{CM_j}}$
MELAA		Middle Eastern, Latin American, African ethnicities.

METHN	$= \frac{\text{The number of people with the same ethnicity as individual i living in her MB}}{\text{Population of individual i's MB}}$	
METHT		$= \frac{\text{The number of people of a specific ethnicity}}{\text{Population of individual i's MB}}$
MHVARC	Comparison between individual's crowding level and the crowding of symbols of success living in the same MB.	Compares an individual's PPBR with the average crowding of individuals with an income at the 90th percentile in his or her MB. A variation method is used for constructing this measure.
MOCCT		The ratios of different occupations constructed as follows: $= \frac{\text{The number of people of a specific occupation}}{\text{Population of individual i's MB}}$
MTDENS		If people in their evaluation of their living environment, compare the density level of the MB that they are located in with the density level at their territorial authority, the constructed relative density variable will be called ATDENS.
MTVARDENS		$\text{MTVARDENS}_i = \frac{\text{DENSM}_i - \text{DENST}_i}{\sigma(\text{DENST})}$
MVARC		$\text{MVARC}_i = \frac{\text{C}_i - \text{CM}_i}{\sigma(\text{CM})}$
N		
NIMBY	Not In My Back Yard	
NZDep	New Zealand Deprivation Index	
NZGSS	New Zealand General Social Survey	
P		
PC	Perceived crowding	People's subjective evaluations of the number of people living in a small area, generally their dwelling.
PPBR	People per bedroom measure	$= \frac{\text{Number of people in the dwelling}}{\text{Number of bedrooms}}$
$\widehat{\text{PPBR}}$	Expected crowding level	The probability of living in a crowded household predicted based on individuals' backgrounds.
$\text{PPBR} - \widehat{\text{PPBR}}$	Exogenous household crowding	
$\text{PPBR} - \widehat{\text{PPBR}}$	Unexpected crowding level	

PPR	People per room measure	$= \frac{\text{Number of people in the dwelling}}{\text{Number of rooms}}$
R		
R^2	Coefficient of determination	
$r_i - \hat{r}_i$	Unexpected relative measure	The level of the absolute housing position of an individual that is beyond (or below) his or her expected level.
RDEP	Relative deprivation of individuals	$= \text{DEP}_i - \text{ADEP}_i$, where ADEP is the average deprivation level of adjacent neighbourhoods.
$\widehat{\text{RDEP}}$	Expected relative deprivation	
$\text{RDEP} - \widehat{\text{RDEP}}$	Unexpected relative deprivation	
RS	Residential Satisfaction	In NZGSS, respondents answer the residential satisfaction question of 'How do you feel about where you are currently living?' on a scale from '1 – very dissatisfied' to '5 – very satisfied', which is an ordered categorical base.
S		
SDEM	Spatial Durbin Error Model	
SDM	Spatial Durbin Model	
SLX	Spatial Explicit Model	
SNZ	Statistics New Zealand	
SPWT	Spatial weight command	The Stata package for constructing contiguity weights based on the k-nearest neighbour and the walking distance from the areas' centroids methods.
T		
TC		$\text{MC}_{ij} = \frac{\text{PPBR of the observation i}}{\text{Average PPBR of observation i's TLA}} = \frac{C_i}{\overline{\overline{CT_j}}}$
THC	Comparison between individual's crowding level and the crowding of symbols of success living in the same TLA	Compares an individual's PPBR with the average crowding of individuals with an income at the 90^{th} percentile in his or her TLA.
THVARC	Comparison between individual's crowding level and the crowding of symbols of success living in the same TLA	Compares an individual's PPBR with the average crowding of individuals with an income at the 90^{th} percentile in his or her TLA. A variation method is used for constructing this measure.

THVARC95	Comparison between individual's crowding level and the crowding of symbols of success living in the same TLA	Compares an individual's PPBR with the average crowding of individuals with an income at the 95th percentile in his or her TLA. A variation method is used for constructing this measure.
TLA	Territorial Authorities	At a larger (than AU) scale, I consider the boundaries of territorial authorities (TLAs) – i.e., local councils – that are administratively determined community of interest that share a common local government.
TVARC		$TVARC_i = \frac{C_i - CT_i}{\sigma(CT)}$

Index

I

J

K

L

M

N

O

P

Q

R

S

T

U

W