Analysing Proximity Bias

Unravelling the perceived liveability patterns of Groningen, Netherlands

Niek van Wattum
Master Real Estate Studies

Abstract

This paper touches upon the first law of geography, by studying how an individual's current place of residence and environment play a role in the perception of liveability elsewhere. The distance between the place of residence and the urbanisation degree of the current place of residence are the key variables of interest in attempting to explain the perceived liveability patterns. For this study the region of Groningen is used as a case study to explore the perceived liveability by individuals all over the Netherlands for the region of Groningen. By performing an ordered logistic regression and an OLS regression, a uniquely comprehensive dataset with documented opinions about liveability of the region of Groningen (N=4,579), is analysed. Results show that a higher urbanisation degree of current place of residence, negatively impacts the perception of liveability elsewhere. Additionally, the more distance between the region of Groningen and the individuals place of residence, the less positive this individual's perception of liveability for the region of Groningen is. These findings may inform policy makers in their marketing strategies to attract more households to move towards the region of Groningen to mitigated predicted population declines in the region of Groningen.

Keywords: Liveability, Distance, Urbanization, Groningen

Colophon

Title	From Proximity to Prosperity: Unravelling the perceived liveability patterns of Groningen, Netherlands
Version	Final version
Author	Niek van Wattum – S3906337
Supervisor	Dr. M. N. Daams
2nd Assessor	prof. Dr. ir. A.J. van der Vlist.
E-mail	n.van.wattum@student.rug.nl
Date	30-06-2023

Table of Content

1 Introduction

"Er gaat niets boven Groningen", or translated: nothing tops Groningen. Is a well-known place branding slogan used for the region of Groningen. The slogan was first launched in 1989 and is still today one of the best known slogans for regional marketing in the Netherlands(Groningen, 2021). The slogan was used as a place branding strategy by the province in order to attract households to live here, tourists that could visit the area and businesses that could settle in the region to improve the local economy. However, it seems that even though the region has this well-known place branding slogan, this does not put a halt on the predicted population decline the region of Groningen is facing. According to the Dutch Central Bureau of Statistics, the majority of the region around the city of Groningen is facing a population decline (Figure 1) between 2021 and 2035 of averagely -6.88%, and the region of Eemsdelta in Groningen being the region with the highest decline in the whole country of -12.6% (CBS, 2022).

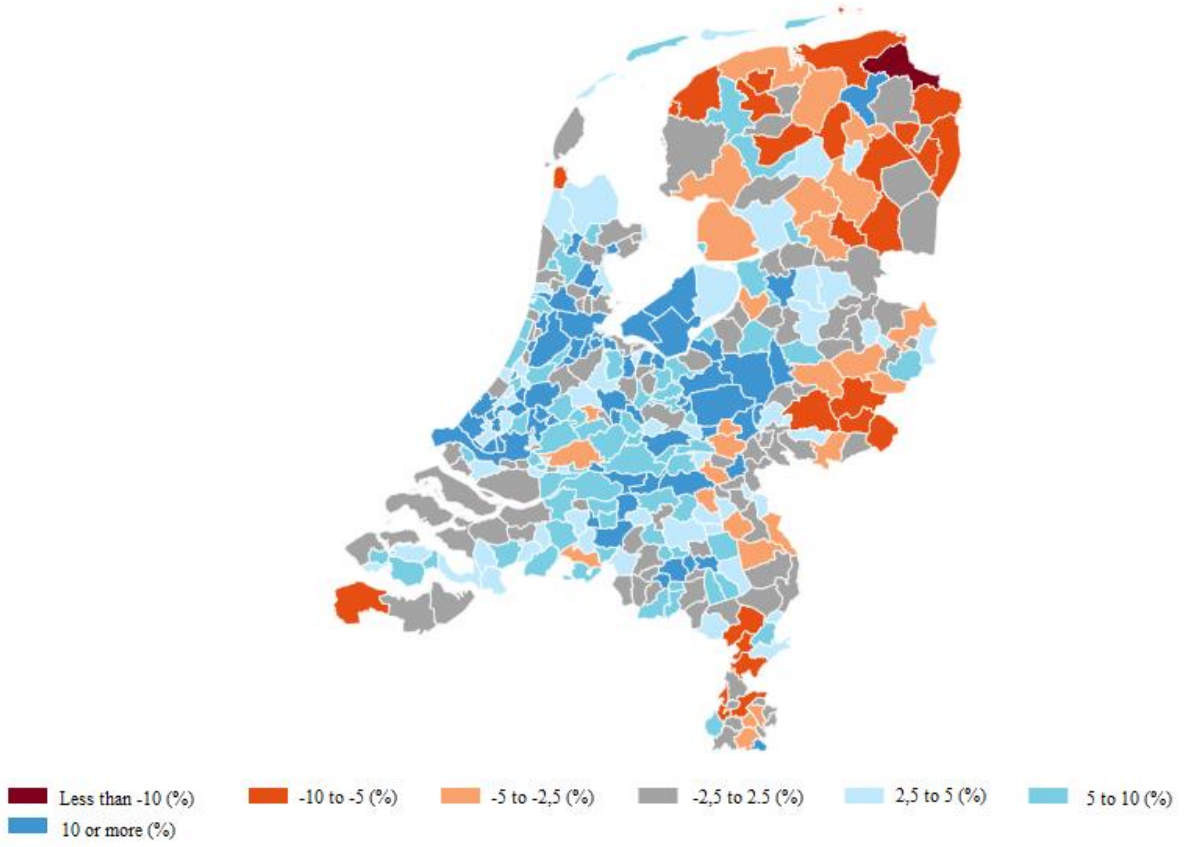


Figure 1: Population growth between 2021 and 2032 (estimation) , source: CBS (2022)

This population decline is driven by way the Dutch population perceives the liveability of the region of Groningen, since the perception of liveability of a region is the strongest determinant for households to relocate (Wheeler, 2013). In order to address the issue of the expected

population decline and potentially reverse it, it is crucial to gain a deeper understanding of the factors that drive the perception of liveability for the region of Groningen. In 2020 a nationwide survey has been distributed by the "Nationaal Programma Groningen", documenting the perception of Dutch individuals about the region of Groningen. In the survey several respondents associated the region of Groningen with words such as: "far away", "boring" , and "left behind" (Groningen, 2020). Furthermore, the local population gave the region of Groningen a score of a 6.5, whereas non-Groningers gave the region a score of a 4.5. If policy makers in the region want to prevent the predicted population decline as shown in figure 1, they have to improve the image of their region. This starts by creating a better understanding of how this perception is formed. The conducted study found that the region of Groningen scores negatively on a place to buy a home and find a job (Groningen, 2020). This perception of buying a house and finding a job puts pressure on the local housing market, because the perception of liveability reflects the image of the area. The image of an area is argued by Cleave and Arku (2020) to be a place's best tool to attract new migrants and boost the local economy. So, if policy makers seek to prevent the presented population decline in figure 1, it is necessary to unravel what drives this perception of liveability in Groningen.

Multiple studies have researched the effectiveness of attracting migrants and how this relates to the notion of the perceived liveability of a place or region. According to Cleave and Arku (2020), place branding is a useful tool for regions and cities to attract highly skilled and educated migrants, and that perceptions about housing and the economy are key drivers in the effectiveness of place branding. In addition, the study of Namazi-Rad et al. (2016), studies the local factors, and how the association with those local factors for a region influences the perceived liveability of a region. Furthermore, besides local physical factors in the place or region that is being assessed on liveability, an individual's current surroundings also play a role. Consisting of physical and social factors that also influences the perception and understanding of the notion "liveability" (Norouzian-Maleki et al., 2018). This would implicate that there are external factors that do not relate to the location itself, that influences the liveability perception of individuals. Existing research and literature already explores notions such as: place branding, liveability, and perceived liveability by residents (Cleave & Arku, 2020; Namazi-Rad et al., 2016; Norouzian-Maleki et al., 2018). However these studies on perceived liveability were conducted on how inhabitants of a specific city/region perceived their own region. Meester and Pellenbarg (2006), conducted a study where the authors asked Dutch entrepreneurs to rank the Netherlands for places to locate their business. The main

conclusions were that the preferences of these entrepreneurs are largely influenced by the centrality of the location and amounts of economic activity (Meester & Pellenbarg, 2006). This study focussed on entrepreneurs and businesses and how they would rate other areas based on business opportunities. Additionally, Donnelly et al. (2020) attempted to map out the identity of places within the UK based on perceptions by students by asking them where they would or would not see themselves study. This study however, does not focus on the country as a whole to be assessed, but rather investigate how the whole country perceives a part or region of the country. In this study we attempt to shed a light on how the perception of liveability is formed not only by the association of local factors of the assessed place, but focussing more on factors relating to the individuals current place of residence. The main focus of this research therefore not only lies in what local factors and associations with those factors drives the perception of liveability, but mostly to see if an individual's current place of residence, shapes their perception about another place. The main research question of this study therefore is: does an individual's current place of residence, shapes their perception of liveability about another place?

By researching the way the perceived liveability of another region is formed, policy makers that represent the region of Groningen get a better understanding of what drives the perception of liveability for their region and are therefore better equipped to deal with the expected population decline by improving the marketing strategies used to attract new households. The main aim of this paper is to study and understand what spatial factors relating to one's current place of residence could have an effect on an individual's perception of liveability elsewhere. In order to gain a better insight on what might influence and drive the opinion of individuals on the topic of liveability several aspects relating to liveability have to be studied and explored. Firstly, this paper will analyse how the spatial factors of the current place of residence might influence the perception of living in another place. These are factors as degree of urbanisation of the current living place, and Euclidian distance between current place of residence and the place of interest: the region of Groningen. Secondly, this paper analyses the association of local factors individuals have with the place of interest and how these associations influence the perceived liveability of a region or place. These associations are if individuals perceive the area as: Safe, Green, Spacious Healthy and if the area has a well-developed infrastructure

Existing literature is analysed and compared to create a better understanding of what factors and associations drive the perception of liveability. Furthermore, data will be analysed to

determine how the degree of urbanisation in one's current place of residence influence the perception of liveability elsewhere and how distance between place of residence and place of interest play a role in forming an opinion on liveability elsewhere. In order to analyse these drivers, a unique comprehensive dataset is used. This dataset has over 5,000 responses containing the opinions of Dutch individuals outside of the region of Groningen, about the region of Groningen.

There are no key ethical issues to be considered relating to the respondents of the survey. Since the data is anonymous and does not contain any personal information. Therefore, it will not be possible to track any individual down based on the filled in survey. Answering the main research question allows for a better understanding on how the perception of liveability is formed, and therefore it provides the policy makers within the region of Groningen with important information about what drives the perception of liveability about their region within the Dutch population.

2 Theoretical framework

2.1 Understanding Liveability

Liveability of a specific area or place is perceived by individuals according to what they think or observe is happing at that particular place or area. How liveable individuals perceive a region is the strongest determinant for migrants to move and settle in that place or region (Wheeler, 2013). This perception is influenced by location specific internal factors and external factors that do not relate to the location itself. Firstly, these external factors are factors like distance towards the place of interest (Meester & Pellenbarg, 2006), and the urbanization degree of the current place of residence (de Haan et al., 2014; Winters & Li, 2017). Secondly, there are local specific factors and associations with those factors that influence the perception of the quality of life within that place. These local factors shape the place its own identity (Wheeler, 2013), which creates an image of the place or region according to the identity. This identity of a place or region will influence individuals perceptions about the place (Cleave & Arku, 2020). Thus, the association of these local internal factors and their presence within the place or region plays a role in assessing the region's liveability. However, one's current place of residence also has an impact on the perception of liveability elsewhere.

2.2 External factors influencing perceived liveability

The perception of regions and places is not only formed by aspects of the region itself, but an individual's own environment plays an important role in assessing other regions (Meester & Pellenbarg, 2006). Distance plays a role (Lee & Pistole, 2012), and urbanization of the current and assessed region matters (Lenzi & Perucca, 2021; Winters & Li, 2017). Geographical distance from a place influence's ones perception about that place, this relates to the first law of geography that mentions that everything is related to everything else, but near things are more related than distant things (Tobler, 1970). The phenomenon of repeated interactions with specific places or individuals has been observed to positively improve an individual's perception and attitudes towards those entities. Thus, improving the liveability perception of that region (Allen et al., 2002). Since individuals who live in closer proximity to places have increased interactions with that place, this is presented in the results from KantarPublic (2022). Where the percentage of individuals who have visited Groningen per province diminishes as distance toward the region of Groningen increases. Furthermore, individuals tend to feel less emotionally attached to places that are physically further away (Lee & Pistole, 2012). In addition, the study of Daams et al. (2019), portrayed monetary value of attractive urban green spaces. The authors established that the added value decreased over distance. Again, proving

that distance plays a role in assessing value's and forming perceptions about places. Distance between the place of residence and the assessed place plays a role in degerming for a household to migrate. However distance towards friends and family also influences our perception of liveability about a place (Mulder & Cooke, 2009). The authors of the paper concluded that family member outside of the household living in close proximity not only prevent some families from moving, but family members living far away also trigger households to move towards these family members (Mulder & Cooke, 2009). The presence of family members relates to the social structure and well-being of individuals within an area (Lloyd et al., 2016), therefore households could be biased towards living in further away places when this relocation will increase the travel distance between family members.

Besides distance towards a place, urbanization degree of the current place of residence also has an effect on satisfaction and perceptions of liveability. Winters and Li (2017) found that a higher degree of urbanization influences one's social well-being. Because higher urbanized areas tend to provide more amenities that improve social well-being. More amenities in an area will therefore lead to higher satisfaction levels and higher liveability standards. The same argument was put forth by Glaeser et al. (2001), that more urbanized areas often offer higher quality and quantity of public services, making these regions more attractive to live in. When other regions and/or places do not offer the same amenities, individuals will rate the liveability of that place or region lower (Winters & Li, 2017). When individuals live further away from these "hotspots" with more amenities, and live in particular smaller cities or less urbanized regions, this negatively impacted their perception of their own place. Implying that when an individual lives in an already highly urbanized area, he/she will think negatively about an area that is less urbanized than the one he/she already lives in (Lenzi & Perucca, 2021). This is mostly fuelled because the perception of liveability, and a perception of what is a normal standard of amenities present in the current place of residence, differs between individuals from a rural landscape and an urban landscape (Winters & Li, 2017). We can conclude that the perception of a place is also influenced by external factors that are related to one's current place of residence rather than only the factors that are associated with the place itself.

2.3 Impact of Associating a region with local specific factors

Local factors that are associated with a region or place play a role in fabricating our perception of liveability of that region or place. It is established that more urbanized areas tend to offer more services and amenities than less urbanized areas (Winters & Li, 2017). Besides amenities

and services, the local economy, the work opportunities, the cultural activities and the local housing market have a positive impact on the perception of liveability in a region when a region is associated with those factors (Cleave & Arku, 2020). Similarly, according to the study of Namazi-Rad et al. (2016) individuals tend to shape their notion of perceived liveability of a certain place according to six different aspects and whether they interrelate the region with these aspects. These six different factors are: (1) home, (2) neighbourhood, (3) work, (4) entertainment, (5) services and (6) connectivity. Thus, when a region is associated with good work opportunities, this will impact the perception of liveability of that region. Besides the urban and economic characteristics of a region and individuals associations with these factors towards a region, there are environmental aspects of regions that also shape the liveability of an area (Antognelli & Vizzari, 2016). The most important environmental aspects that could enhance the liveability in a certain neighbourhood points towards "amount of green spaces" and "presence of trees and natural elements" (Brander & Koetse, 2011). According to the study of Norouzian-Maleki et al. (2015), the perception of liveability is influenced by the association of these "green" attributes when assessing neighbourhoods. Since these green areas are often perceived as more healthy areas to live in. According to (Antognelli & Vizzari, 2016; Daams et al., 2019), the presence of trees and other natural elements in urban areas has been linked to a range of positive effects such as: reduced air pollution, decreased noise levels, improved mental health, and increased social cohesion. Therefore, associating a region with these "green" factors will influence an individual's perception about the place. The presence of trees and other environmental aspects are also interlinked with an increased social cohesion in the area (Brander & Koetse, 2011). Lloyd et al. (2016), discussed in their study that social cohesion factors have been found to have a significant impact on the liveability of communities, and therefore places. When areas or regions are known to have a strong sense of community feeling, this will be portrayed outwards and attract other migrants. By creating a solid foundation in which the community can flourish and create a common feeling of "belonging" within the local community, this can greatly enhance the perception of a region to settle for the long run (de Haan et al., 2014; Lloyd et al., 2016). Furthermore, whether an area is commonly known as "safe", this association will have a big impact on the perception of individuals of that place or location (Antognelli & Vizzari, 2016; Wheeler, 2013). According to Wheeler (2001) crime rate plays an important role in the overall feeling and perception of safety of an region or place. When an area is facing high crime rates this has a negative impact on the liveability of that place, and associating a place with an unsafe feeling will result in less individuals wanting to live in that area and potentially migrate out of that perceived unsafe area (Wheeler, 2001).

Thus, associations we as individuals have with a place or region tend to shape our perception of liveability of that place or region. It is important to stress that all the associations with a place are based on the perception of an individual rather than if that assessed factor is actually present in the assessed region or location. Therefore, these associations are all based on the information the assessor has available. The amount of available information and perception a person has about a place is influenced by the amount of interactions the individual has with the assessed place (Allen et al., 2002), and as mentioned these interactions occur more often when the distance between the person and place is less. Henceforth distance plays a role in information gathering about a place which in turn shapes our perception about the place or region

2.4 Hypothesis

Based on the above-mentioned literature the following hypothesis are formed.
Hypothesis 1: *Distance between the current place of residence and the region of Groningen has a significant, negative impact on perceived liveability about the region of Groningen.*

Distance between areas and places plays a role in generating our perception of that place/region. Repeated interactions between places and individuals will positively improve that individuals perception of those places (Allen et al., 2002). The amount of interactions with a place or region decreases as distance towards the place increases (KantarPublic, 2022). That's why it is expected that there is a negative relationship between the measured distance in km between the region of Groningen and the place of residence, and the perceived liveability of that individual of the region of Groningen. Implying that if the distance in km towards the region of Groningen increases, the perceived liveability of the region of Groningen will decrease when within the statistical model all other factors are kept constant.

Hypothesis 2: *Urbanisation degree of the current place of residence has a significant, negative impact on perceived liveability about the region of Groningen.*

According to Winters and Li (2017), places with higher urbanization degree's offer on average more amenities that contribute to social well-being and liveability. In turn when individuals live in less urbanized areas they tend to criticize their own place of residence more, implying when an individual lives in an area with a higher degree or urbanization, this individual will

have an increased negative perception of liveability towards less urbanized areas (Lenzi & Perucca, 2021). Therefore, a negative relationship between degree of urbanization and the perceived liveability of the region of Groningen is expected to exist. Suggesting that if an individual lives in a more urbanized area, the perceived liveability of the region of Groningen will decrease since these respondents prefer the more urbanized areas with higher degree of amenities (Lenzi & Perucca, 2021).

Hypothesis 3: *Local factors that are associated with the region of Groningen by the respondents such as: Safety, Greenery, Economic infrastructure and Social benefits have a significant, positive impact on perceived liveability about the region of Groningen.*

Local specific factors that are associated with a region/place play a role in fabricating our perception of liveability of that region/place. Factors such as Safety, Greenery, Economic infrastructure and social benefits are examples of these meaningful factors that contribute to our perception of liveability (Namazi-Rad et al., 2016). Hence it is expected that there will be a positive relationship between the above mentioned local specific factors relating to the region of Groningen and the perceived liveability with the region of Groningen. Suggesting that if an individual associates Groningen with above mentioned factors, the perceived liveability of the region of Groningen will increase.

3 Methodology

3.1 Empirical model

To answer the main research question: whether an individual's current place of residence, shapes their perception of liveability about another place, a statistical analysis is performed. Within this paper two types of regressions are run, an ordered logistic regression as the main model and an ordinary least squares regression as an additional model. The OLS is run to give a clearer definition of the sizes of the relationships between the independent and dependent variables. The ordered logistic regression is formulated as:

$$\text{Logit}(\text{P}(\text{Y} \leq \text{j})) = \alpha_j + \beta_1 \text{Urbanisation} + \beta_2 \text{Distance} + \sum_{i=3}^{11}(\beta_\text{i} X_\text{i}),\ \text{j} = 1, \dots, \text{c} - 1 \qquad (1)$$

Where $\text{Logit}\big(\text{P}(Y \leq \text{j})\big)$ represents the natural logarithm of the odds of an outcome of the transformed dependent variable being: "Liveability score" categorized in the following five categories: Very Unattractive, Unattractive, Neutral, Attractive and Very Attractive respectively, being in the j-th category or lower (more attractive) given the different predictor variables. It is important to note that the liveability score is a result of an equal weighing of three separate given scores by the respondents, this will be elaborated further after the models are presented. The α_j represents the threshold cut point of the 5 different categories. The two main spatial variables of interest with their corresponding coefficients β_1 and β_2 are: "Urbanisation", measured in addresses per km2 of current place of residence and "Distance(km)" measured as distance in km between current place of residence and the region of Groningen respectively. The operator $\sum_{i=3}^{11}(\beta_\text{i} X_\text{i})$ represents nine independent variables with their corresponding coefficients. These independent variables consist of six binary independent variables containing a Yes or No answer to if the respondents thought the following words were related with the region of Groningen: Safe, Green, Spacious, Healthy, Well connected, and the last binary variable asking if the respondents ever visited Groningen. The remaining three independent variables consist of three Socio Economic control variables with their corresponding coefficients. Being the respondents: sex, age category and highest finished education level.

In order to fit the test, the dependent variable: "liveability score" needs to be transformed. The liveability score is transformed into to a five-point scale for this analysis, ranging from very unattractive to very attractive. The variable has been split up as following: very unattractive

with a score of below three, unattractive ranging from a three or higher until lower than a five. Neutral being a five or higher until lower than a six, attractive ranging from a six or higher until up to a seven, and very attractive ranging from a seven or above. As proposed in DeMaris (1995), happiness or attractiveness in this case, is felt by the respondent and could be implemented as a continuous variable. However, when implementing events such as feelings of safety in Groningen the interpretation of an ordered logistic regression suits the data. Because the results could be implemented as of how the association of safe towards the region of Groningen results in propensity of an event (feelings of attractiveness) to occur. The same method and cut points were used in the study of (Abrudan et al., 2020), where the authors used a ordered logistic regression to measure customer satisfaction on a five-point scale.

Besides the ordered logistic regression also an OLS regression is performed for clearer interpretation of the coefficients of the independent variables. The OLS regression is formulated as:

$$Y = \beta_0 + \beta_1 \text{Urbanisation} + \beta_2 \text{Distance} + \sum_{i=3}^{11} (\beta_{\text{i}} X_{\text{i}}) + \varepsilon_i \qquad (2)$$

Where the left-hand side variable is now the Liveability score, the difference between the ordered logistic regression is mainly that the dependent variables within the OLS regression is not transformed into a categorical variable. The predictor variables in the second equation are similar as those in the first equation. The model assumptions for the OLS are tested with a regression check in Stata and checks for the following assumptions: heteroskedasticity, multicollinearity, normally distributes residuals, specification problem, the appropriate functional form and for influential observations. The following assumptions are violated: heteroskedasticity and the residuals were not normally distributed. The problem of heteroskedasticity is solved by implementing the robust function within the regression. The normality of the residuals is tested by the Shapiro-Wilk W normality test with a null hypothesis being that there is no difference between the distribution of the residuals and a normal distribution. This problem of non-normality of the residuals is explained by some of the independent variables being of a categorical nature and the distance variable is not log transformed. However, the combined interpretation of the ordered logistic regression and the OLS regression will ensure no mis-interpretation of the results Since within the ordered logistic regression there is no assumption of normality within the predictor variables. The model does

however assume that the dependent variable is ordered, one of the predictor variables is either continuous, categorical or ordinal, and there is no multicollinearity between the predictor variables. These assumptions have all been met, therefore the ordered logistic regression is a good fit after transforming the dependent variable into an ordered, five-point scale, variable.

3.2 Describing the Data

This paper uses an unique dataset from the Imago Monitor Groningen (KantarPublic, 2022), which was set up in collaboration with the University of Groningen and the Hanze University of Applied Science. In 2020 surveys where distributes to individuals distributed across the Netherlands. Asking people about their opinion about the province of Groningen on different topics. Mentioned in the introduction of this paper the province faces an expected population decline between 2021 and 2035 (CBS, 2022). The Imago Monitor Groningen askes the public questions regarding the province overall, for this study in particular only a selective amount of data is retrieved from the total dataset, this because the topic of interest for this study is focussing on liveability and thus, what the population thinks about living and or working in the province of Groningen. Furthermore, for this study a spatial element is added to the dataset. GIS software is used to calculate the Euclidian distance between the place the respondents' answered questions about, in this case Groningen, and their place of residence. Additionally, a degree of urbanisation as given in CBS (2020) of the postal code of the respondent is added to the data to determine whether degree of urbanisation of the place of residence influences one's perception of liveability of another place.

Spatial element

This study focusses on how urbanization of current place of residence and distance towards the place of interest impacts the perception of liveability in another place. In the original dataset, the respondents were asked to fill in their postal code. Any missing values or faulty values were dropped from the dataset. The extracted data per postal code from the Dutch Central Bureau of Statistics contained an excel file with all existing postal codes with demographic information such as age of local residents and degree of urbanization of the postal code[1] (CBS, 2020). The extracted data from the BAG register on urbanisation degree were absolute numbers, within the descriptive statistics section the variable will also be presented according to the urbanisation scale as given by the Dutch Central Bureau of Statistics (CBS, 2020). The categories for the

[1] Note: spatial scale of postal code is chosen since within the dataset used for the analysis the respondents only spatial reference is the four-digit postal code

urbanisation degree are according a five-point scale. 1 being very highly urban with 2,500 or more addresses/km2, 2 being highly urban with 1,500 to 2,500 addresses/km2, 3 being moderately urban with 1,000 to 1,500 addresses/km2, 4 being little urban with 500 to 1,000 addresses/km2 and 5 being not urban with less than 500 addresses/km2.

Besides the degree of urbanisation, a second spatial element is added. Mentioned in the introduction of this paper the Dutch population thinks Groningen far away (Groningen, 2020) from the rest of the Netherlands, and analysed literature showed that physical distance to places influences one's perception about that place. To create data on distance all postal codes in the dataset and the one from Groningen are analysed in the software program ArcGIS Pro. ArcGIS pro is a software program that allows spatial data to be created/analysed/explored. To create a variable that contained information about distance the centroid of each postal code was taken as point data. This resulted in 4,070 points in ArcGIS that contained the spatial data linked to the postal code. The software program is used to calculate Euclidean distance between the postal code points spread out across the Netherlands, and the postal code pinpoint of the region of Groningen. The pinpoint in the region of Groningen is the centroid of the polygon representing the municipality of Groningen[2]. The dataset containing the survey information is merged with the dataset containing the spatial data per postal code.

Data cleaning

For this study it would have been interesting to be able to control for other relevant factors such as work status and cultural background, because as explored in the literature these factors shape the notion of liveability (Gentin, 2011; Norouzian-Maleki et al., 2015). However, due to a too large number of missing values, or respondents refusing to answer the questions, these variables had to be dropped. A total of 51% of the respondents chose to not answer the questions regarding work status and income level, dropping all missing values in this case would lead to a too small sample size. After all the data was sorted, and missing values dropped where necessary, a total of 4,579 observations remained in the dataset. The representativeness of the remaining data is checked and presented later in the descriptive statistics subchapter by comparing it to the population averages.

[2] The coordinates of this point are: 53.218984, 6.622509

3.3 Descriptives

Table 1
Comparison sample - population

Variable	Frequency	Percentage dataset (%)	Percentage NL (%)
Province			
Drenthe	180	3.93	2.82
Flevoland	160	3.49	2.50
Friesland	226	4.94	3.70
Gelderland	509	11.12	11.98
Limburg	360	7.86	6.34
Noord-Brabant	714	15.59	14.74
Noord-Holland	691	15.09	16.58
Over Ijssel	326	7.12	6.65
Utrecht	323	7.05	7.79
Zeeland	137	2.99	2.20
Zuid-Holland	953	20.81	21.36
Sex			
Male	2410	52,63	49,72
Female	2169	47,37	50,28
Age			
16 – 29 Years	390	8,52	17,70
30 – 39 Years	508	11,09	12,84
40 – 49 Years	686	14,98	11,92
50 – 59 Years	982	21,45	14,26
60 – 69 Years	1307	28,54	12,41
70 – 79 Years	636	13,89	9,54
80 Years or more	70	1,53	4,91
Education			
Primary school	102	2,23	
VMBO	686	14,98	
HAVO/VWO	562	12,27	
MBO	1706	37,26	
HBO	1216	26,56	
University	307	6,70	
Total N = 4579			

Note: Percentage NL represents the population percentage as of 2023.
Source: CBS (2023)

Table 1 above shows the distribution of the province the respondents live in, and the percentage of population in the Netherlands as of 2023(CBS, 2023). The distribution of the dataset has a similar distribution pattern as that of the Netherlands. Ensuring equal and proper representation of each province in the dataset to prevent a bias, or overrepresentation of a particular province.

Table 1 also presents the Socioeconomic controls used in the data analysis which represent several characteristics of the respondents. The table presents an equal representation of both sexes (male and female). Important to note that the age of the respondents is relatively high compared to the average in the Netherlands and there is an over representation for the age groups above 40 on average. However, according to Kadaster (2021), the average age of the home buyers has increased with 15% to an average of 43 years of age. Since this study focusses on liveability and tries to inform policy makers to face the population decline, there has to be a good understanding how the group of homebuyers shapes their perception of liveability. The education variable could not be compared one-on-one with the averages in the Netherlands. Because the dataset uses a different distribution of highest finished education level than the Netherlands is currently using. The data however shows that over 75 % of the respondents has finished either an HAVO/VWO, MBO (secondary vocational education) or an HBO (higher vocational education) degree. Furthermore, only 2% only finished primary school as their highest educational degree whereas the percentage of individuals having primary school as their highest finished education is on average 7.2% in the Netherlands (CBS, 2018). Showing that the respondents from the dataset on average have a higher educational level as the average in the Netherlands.

Table 2
Descriptive statistics

Variable	N	Minimum	Maximum	Mean	Std. Deviation
Liveability score	4,579	0	10	5.14	2.22
Place to live	4,579	0	10	5.14	2.50
Place to find a job	4,579	0	10	4.80	2.33
Place to grow old	4,579	0	10	5.49	2.53
Place to study	4,579	0	10	6.45	2.38
Distance (Km)	4,579	9.89	306.48	166.58	58.79
Addresses per Km2	4,579	18	11,324	1,771.81	1,404.77

Liveability score categorized	Frequency	Percentage (%)
Very Unattractive	885	19.33
Unattractive	684	14.94
Neutral	914	19.96
Attractive	942	20.57
Very Attractive	1,154	25.2
Total	4,579	100

Urbanisation degree categories (Addresses/km2)	Frequency	Percentage (%)
1 = Very Highly Urban (2,500 or more)	893	19.50
2 = Highly Urban (1,500 to 2,500)	1,437	31.38
3= Moderately Urban (1,000 to 1,500)	883	19.28
4= Little Urban (500 to 1,000)	776	16.95
5= Not Urban (Less than 500	590	12.88
Total	4,579	100

		Yes		No	
	N	Frequency	Percentage (%)	Frequency	Percentage (%)
Safe	4,579	3,330	72.72	1,249	27.28
Green	4,579	3,851	84.10	728	15.90
Spacious	4,579	4,121	90.00	458	10.00
Healthy	4,579	3,564	77.83	1,015	22.17
Well Connected	4,579	2,367	51.69	2,212	48.31
Visited Groningen	4,579	3,480	76.00	1,099	24.00

***Note**: Liveability score consists of an equal weighing of the score of: Place to live, place to grow old and place to find a job*

Table 2 presents the descriptives of the dependent variable Liveability score and the scores given by the respondents to calculate this liveability score. These elements influence the notion of liveability for individuals as discusses in the literature, related to Groningen. The table shows that overall, the respondents think negatively on living in Groningen, to find a job in Groningen and think negatively about the region of Groningen as a place to grow old. However, the region of Groningen scores a 6.45 as a place to study. Appendix 1 presents the distribution of the given scores that relate to the perceived liveability/ These variables have an influence on the perceived liveability of a region (Antognelli & Vizzari, 2016; Cleave & Arku, 2020; Namazi-Rad et al., 2016). Due to their interlinkage with the liveability of a region a check is performed to combine the scores. First the data is "eyeballed" with the use of the histograms. It is clear that the first three variables show a similar pattern, Place to study is different than the other 3 by it being more right skewed. The correlation matrix is presented appendix 2 presents a strong correlation between the variables; Place to live, Place to find a job and place to grow old. Based on discussed and shown relations between the variables, the three separate variables place to live, place to find a job and place to grow old are combined by equal weighing of the three separate scores into one score that represents the perceived liveability score. This score is the average of the respondent's grade on; Place to live, Place to find a job and Place to grow old. Place to study is excluded from this combination since the correlation is not strong enough to include this variable in the combined score. Comparing the descriptives of the combined score with the separate scores presented the table it shows similarity and the standard deviation is lower. Meaning better consistency, producibility and quality. Because the standard deviation serves as a measure that quantifies the extent of variation or dispersion (the spread) within a dataset. When the standard deviation is smaller, it indicates that the values are more tightly clustered, implying greater precision. Conversely, a large standard deviation suggests that the values are more diverse, leading to reduced precision.

Table 2 also presents the spatial independent variables used in the regression models. Where the minimum distance in kilometres is less than 10 and the maximum over 300. The observation that is furthest away from the region of Groningen comes from the province of Zeeland, and the closest by observation is from the province of Drenthe. The individual that came from the place with the lowest density was also from Zeeland, and the individual that came from the place with the highest density came from the province of Noord-Holland. Table 2 also presents the distribution of the categorized dependent variable: Liveability score, used for the ordered

logistic regression. Within the data 34.27%[3] has a negative perception (unattractive or very unattractive) of liveability towards the region of Groningen. Whereas 45.77%[4] of the respondents within the dataset showed a positive perception (attractive or very attractive) of liveability towards the region of Groningen. Furthermore the table presents the urbanisation degree per category determined by the Dutch Central Bureau of Statistics (CBS, 2020). It shows that the Netherlands is a highly urbanized country, comparing the percentages in the dataset to the population average was not possible since this data was not available. However, over 70% of the respondents lives in an Urban are of at least a 1,000 addresses per km2, and the largest chunk of the respondents even lives in a Highly Urban area. According to the National Environmental Planning office in the Netherlands, almost 75% of the population lives in an urbanized area (PBL, 2015), therefore the degree of urbanization within the dataset is representative towards the population.

Table 2 furthermore presents the descriptives for the variables that represent the respondent's association with Groningen for: Safe, Green, Spacious, Healthy and Well Connected. Additionally, the table also shows how many individuals from the dataset have ever visited the region of Groningen before. The descriptives show that majority associates the region of Groningen with: Safe, Green, Spacious and Healthy. The respondents show mixed feelings whether the region of Groningen is well connected with an almost fifty – fifty distribution between yes and no. Furthermore, 76% of individuals analysed within the data has even visited Groningen before.

[3] Calculation: 19.33+14.94 = 34.27
[4] Calculation: 20.57+ 25.2 = 45.77

4 Results

4.1 Ordered logistic regression

Table 3 presents the ordered logistic regression. All four models show an overall significance, we observe that the log-likelihood increases between the models. The log-likelihood is an indicator of measure of goodness of fit. The score measures how probable it is that our model accurately describes the data (Stata, 2021; Statology, 2021). The log-likelihood should not be interpreted as an absolute number but rather be compared between models as it ranges between negative infinite and positive infinite. When the score increases between models it shows that the model fits the data better. We observe in model 1 a lower log-likelihood compared to model 4, indicating that the overall goodness of fit of the model has improved. Suggesting that the perceived liveability is not only explained by the urbanisation degree of the current place of residence and distance between place of residence and Groningen, but also partially if the respondents associate the region of Groningen as: Safe, Green, Spacious, Healthy and well Connected.

The urbanisation degree of the current place of residence shows a significant odds ratio of .651. Indicating that when urbanisation degree increases by 10,000 addresses per km2, the chance of the respondent giving a higher score on liveability decreases with 34.9%[5]. These results indicate that when an individual lives in a higher urbanized area the change of giving a higher liveability score decreases. Reflecting to the main research question in this paper, it seems that factors that relate to the individuals place of residence, particularly urbanisation degree in this case, matters in this particular case when assessing liveability in another region. Furthermore, the variable Distance (km) reports a significant odds ratio of .899. Indicating that when distance in km increases by 50 kilometres, the change of the respondent giving a higher score on liveability decreases with 10.1%[6]. These findings indicate that when distance increases, the change of giving a higher liveability score decreases. Relating this to the main research question where factors of the current place of residence, location in this case, matter when assessing liveability elsewhere. Since the location of the place of residence determines the distance between the place of interest and the current place of residence.

[5] Calculation((1-0.651)*100=34,9
[6] Calculation (1-0.101)*100=10.1

Table 3 also reports the variables asking about the association of Safe, Green, Spacious, Healthy and Well connected with the region of Groningen. Comparing a no to a yes. Safe, Green, Healthy and Well Connected were significant, whereas spacious is not. The presented relationships within the ordered logistic regression show a positive direction. Indicating that when an individual associates the region of Groningen with these factors the odds of them being in a higher liveability score category increases. The question if the respondent has visited Groningen also showed a positive relationship, where if an individual did ever visited Groningen before, they had an increased change of giving a higher liveability score. The sex variable was not significant, and the only age category showing a significant result is the age category of 80 years or older.

Table 3
Ordered Logistic Regression Results

	(1)	(2)	(3)	(4)
Addresses km2(10,000)	.606**(.113)	.725*(.136)	.663 **(.127)	.651**(.125)
Distance(50km)		.840***(.019)	.899***(.022)	.899***(.022)
Safe (ref. No)			1.751***(.120)	1.771***(.122)
Green (ref. No)			1.108* (.095)	1.114* (.095)
Spacious (ref. No)			1.162 (.124)	1.159 (.125)
Healthy (ref. No)			2.270***(.173)	2.256***(.172)
Well Connected (ref. No)			2.327***(.129)	2.324***(.130)
Visited_Gro (ref. No)			1.579***(.102)	1.558***(.103)
Sex (ref. male)				.966 (.052)
Age (ref. 19 – 29 years)				
30 – 39 years				1.061 (.130)
40 – 49 years				.987 (.114)
50- 59 years				.981 (.107)
60 – 69 years				1.070 (.113)
70 – 79 years				1.181 (.141)
80 > years				2.078** (.484)
Education level (ref. primary school)				
VMBO				.692* (.135)
HAVO/VWO				.871 (.173)
MBO				.792 (.149)
HBO				.764 (.146)
University				.736 (.155)
Observations	4,579	4,579	4,579	4,579
Model Chi2	0.0071	0.0000	0.0000	0.0000
Log likelihood	-7304.3	-7274.9	-6869.3	-6857.3

Note: Coefficients represent odd ratios. Dependent variable is the liveability score categorized, Addresses per km (10,000) is the variable addresses per km2 / 10,000, Additionally Distance (50km) is the variable of distance per km / 50
*, **, *** Significance at 10%, 5%, and 1%, respectively

4.2 OLS regression

Appendix A3 presents table 4 where the OLS regression results are presented. We observe the first key variable of interest: Addresses per km2, which measures the degree of urbanisation in the place of residence of the respondent. This variable presents a negative coefficient of -.408, at a 90% confidence interval. Meaning that when the degree of urbanization increases by 10,000 addresses, the given liveability score decreases with .408 when keeping all other variables constant. Implicating that the individuals from our study that live in the most urbanized area of this study, being Amsterdam, with 11,324 addresses per km2 will give 0.46 [7] lower score on the perceived liveability of Groningen, when keeping all other variables constant. Therefore, the degree of urbanization within the current place of residence plays a role when assessing liveability in another region. Showing similar results as in the ordered logistic regression. The following key variable is the variable Distance (km), measuring the amount of increase in distance per fifty kilometres between the region of interest (Groningen) and the current place of residence of the respondent. This key independent variable presents a negative coefficient of -.112, at a 99% confidence interval indicating that when distance to the place of interest increases by fifty kilometre the liveability score decreases with -.120. To put it in the same perspective as earlier, the same person living in the highest urbanized area also lives 150.5 kilometres from Groningen. This would result in a lower given liveability score of 0,34 [8]. Implying that distance towards a place plays a role in assessing its liveability. Also showing similar results as presented in the ordered logistic regression

After these key independent variables, the variables asking about the association of location specific factors about the region of Groningen are presented, comparing a no to a yes, the corresponding coefficients relate to when a respondent thinks the words match with the place of interest. Safe, Green, Spacious, Healthy and Well Connected all have a positive significant impact on the liveability score, these results are similar as those presented within the ordered logistic regression. The socio-economic control variables within the OLS regression also show similar results as those presented in the ordered logistic regression.

[7] Calculation: (11,324/10,000) * .408 = 0.46
[8] Calculation: (150.5/50) * .112 = 0.34

4.3 Robustness checks

Concerns could arise when assessing the results presented in table 3 and 4. Especially taking into the account spatial differences, being urbanisation patterns, across the Netherlands. The Randstad in the Netherlands is an example of a highly urbanized area that might overshadow trends and patterns elsewhere in the Netherlands. Coefficients of the essential core variables should be relatively unaffected by the removing or adding of observations (Lu & White, 2014). By removing all observations within the province of Zuid Holland the robustness of the findings is checked. Furthermore, if there are no large differences observed it can be concluded that the results presented in the main model are grounded and robust enough to withstand changes and alterations with the dataset. The robustness check results are presented in table 5. The resulting estimates are similar to those that were presented in the unrestricted model presented in table 3 and 4. The robustness check demonstrates that the key variables remain unaffected by the inclusion or exclusion of additional observations in both regressions.

Table 5
Robustness Check

	Ordered Logistic Regression (A)	OLS Regression (B)
Addresses km2 (10,000)	.478**(.115)	-.752**(.292)
Distance km (50)	.896***(.023)	-.115***(.028)
Safe (ref. No)	1.725***(.134)	.585***(.090)
Green (ref. No)	1.073 (.1035)	.154 (.111)
Spacious (ref. No)	1.080 (.130)	.165 (.141)
Healthy (ref. No)	2.248*** (.192)	1.020***(.100)
Well Connected (ref. No)	2.351*** (.149)	.957***(.070)
Visited_Gro (ref. No)	1.583*** (.121)	.570***(.089)
Socioeconomic controls	Yes	Yes
Constant		3.449
Observations	3,626	3626
Model Chi2	0.0000	
Log likelihood	-5,419.2	
R-Squared		.197

Note: Dependent variables are the liveability score categorized in model A for the ordered logistic regression and the liveability score in model B for the OLS regression. Coefficients in model A represent odds ratios Addresses per km (10,000) is the variable addresses per km2 / 10,000, Additionally Distance km (50km) is the variable of distance per km / 50. Socioeconomic controls consist of respondent Sex, Age and Education Level. Within this robustness check all observations from "Zuid Holland" are removed.
*, **, *** Significance at 10%, 5%, and 1%, respectively

4.4 Decomposition of the liveability score

The reader will recall that the dependent variable in the regression is a combination of three separate questions asked to the respondents in the dataset. Being: place to live, place to find a job and place to grow old linked to the region of Groningen. The literature showed that the perceived liveability of a region or place, is amongst other things influenced by its local housing market and job opportunities (Cleave & Arku, 2020). Which is why the liveability score, which is analysed in the main model, consist of the combination of the three separate factors. However, as presented in the beginning of the paper the region of Groningen scores negative on a place to buy a home, place to find a job and place to grow old. With place to find a job being the most negative amongst the three. Therefore, it is interesting for this study to check if the established relationships in the main model are similar when the liveability score is separated in its original parts. To check whether the three original parts respond differently to the influence of the independent variables. Furthermore, this in-depth analysis can provide useful information for policy makers of the region of Groningen if the goal is to target specific groups of individuals that might not look for a place to grow old yet, but rather a place to find a job. The main results of each full model are presented below in table 6.

Table 6
Decomposition liveability score categorized

	Place to live (A)	Place to find a job (B)	Place to grow old (C)
Addresses km2 (10,000)	.622** (.121)	.604** (.118)	.871 (.169)
Distance km (50)	.933**(.023)	.843***(.020)	.953** (.023)
Safe (ref. No)	1.752***(.121)	1.391***(.095)	1.783***(.124)
Green (ref. No)	1.093 (.093)	.946 (.081)	1.375**(.116)
Spacious (ref. No)	1.180 (.125)	.995 (.107)	1.404**(.148)
Healthy (ref. No)	2.134***(.162)	1.959***(.148)	2.390***(.184)
Well Connected (ref. No)	2.002***(.113)	2.714***(.153)	1.833***(.103)
Visited_Gro (ref. No)	1.676***(.111)	1.253***(.083)	1.577***(.105)
Socioeconomic controls	Yes	Yes	Yes
Model Chi2	0.000	0.000	0.000
Log-likehood	-6,649.3	-6,956.8	-6,703.3

Note: Coefficients represent odds ratios, the dependent variables are categorized similarly as the liveability score. The dependent variables changes per model as A) Place to Live, B) Place to find a job and C) Place to grow old respectively. Number of observations for each panel (A, B, C) is 4,579. Addresses per km (10,000) is the variable addresses per km2 / 10,000, Additionally Distance km (50km) is the variable of distance per km / 50. Socioeconomic controls consist of respondent Sex, Age and Education Level. *, **, *** Significance at 10%, 5%, and 1%, respectively

Important to stress out is that the dependent variables presented in table 6 are categorized similarly as the dependent variable liveability score which is analysed in table 4. The separation of the liveability score allows for an analysis per segment of the perceived liveability factor. Being place to live, place to find a job and place to grow old. Comparing the results between the liveability score and the respondents asking about place to live presented in model A, we do not observe any significant changes between the two models regarding the key independent variables. Checking model B with the place to find a job as dependent variable we observe a change in odds ratio for whether the respondents think the region of Groningen is well connected. Signalling that connectivity is more important for individuals when they asses regions based on a place to find a job compared to assessing a place for growing old. Model C however does show different results than the preceding models. This is the only model where degree of urbanisation is not significant. Implying that when individuals assess whether a place is suitable to grow old, degree of urbanisation of their current place of residence does not play a significant role. However, local factors such as whether the region is associated with notions as: green, or spacious is significant. These factors were insignificant when individuals assessed the region as a place to live or to find a job. Suggesting that the association with these local factors matters less when assessing location for job opportunities compared to assessing a location to grow old. Furthermore, the coefficients of "green" and "spacious" experienced the biggest increase when comparing the separate variables with the combined liveability score presented in the unrestricted model. Appendix 4 contains table 7, which is the same decomposition of the liveability score as in table 6 but with an OLS regression. This test shows similar results as presented in table 6.

Table 8
Province segregation

	Ordered Logistic Regression (A)	OLS regression (B)
Addresses km2 (10,000)	.841(.176)	-.135 (.248)
Dummy_Drenthe	1.797***(.281)	.636***(.163)
Dummy_Flevoland	1.397**(.220)	.416***(.160)
Dummy_Friesland	1.629***(233)	.473***(152)
Dummy_Gelderland	.893(.091)	-.118 (123)
Dummy_Limburg	1.005(.116)	.039 (.129)
Dummy_Noord-Brabant	.950(.087)	-.028 (.103)
Dummy_Noord-Holland	.787***(.071)	-.279***(.102)
Dummy_Overijssel	1.500***(.180)	.426***(.126)
Dummy_Utrecht	1.016(.118)	.092 (.124)
Dummy_Zeeland	1.072(.178)	.009(.192)
Safe (ref. No)	1.800***(.125)	.633***(.080)
Green (ref. No)	1.133 (.098)	.201** (.098)
Spacious (ref. No)	1.199* (.130)	.289** (.125)
Healthy (ref. No)	2.242*** (.171)	.998***(.089)
Well Connected (ref. No)	2.298*** (.129)	.919***(.061)
Visited_Gro (ref. No)	1.564*** (.104)	.540***(.076)
Socioeconomic controls	Yes	Yes
Constant		2.737
Observations	4,579	4,579
R-Squared		0.204
Model Chi2	0.0000	
Log likelihood	-6,836.2	

Note: Dependent variable is the liveability score categorized in model A for the ordered logistic regression and the liveability score in model B for the OLS regression. Coefficients of model A represent odds ratios. Socioeconomic controls consist of respondent Sex, Age and Education Level. Within this robustness check all observations from "Zuid Holland" are removed. Addresses per km (10,000) is the variable addresses per km2 / 10,000, Additionally Distance (50km) is the variable of distance per km / 50. The dummy variable representing the province of Zuid-Holland has been omitted by the model and used as the reference category. *, **, *** Significance at 10%, 5%, and 1%, respectively

Table 8 above presents an analysis per Province, to see beyond the relationship of just linear distance. The provinces of Drenthe, Flevoland, Friesland and Overijssel show significant results and a positive relationship towards the perceived liveability in the region of Groningen. These are also the four Provinces in closest proximity to the province of Groningen. This relationship with proximity and assessing liveability presented in table 8 is in line with the findings of Allen et al. (2002), where the author concludes that increased distance leads to

increased negative perceptions since the amount of interactions diminishes when distance increases. Furthermore, the province of Noord-Holland has a significant negative relationship with the liveability score of Groningen. Suggesting that individuals from this province on average give a lower liveability score towards the region of Groningen. Hence is why another in depth analysis is run and is presented in table 9 in appendix 5 and will be used in the discussion of the results.

4.5 Mapping liveability score

On the following page the first map is presented in figure 2. This map is based on predicted values from the second statistical equation, the OLS regression. The map is made to visualize the spatially mentioned factors of influence. These being degree of urbanisation of current place of residence and distance from current place of residence towards the region of Groningen. The map makes it easier to interpret the effect of distance and urbanisation on the liveability score within a spatial scale. The legend on the left in figure 3 shows the predicted liveability score ranges from lower than 1 and above a 7. It is immediately clear that as the distance towards the region of Groningen increases, the score of liveability goes down. Furthermore, we see a concentration of the lower scores around urbanized areas. This is especially clear around the city of Leeuwarden, where the distance variable does not come into play as much as further away in the country. We can observe the effect of the higher urbanized area of Leeuwarden on the liveability score estimated by our model.

On page 33 the reader finds figure 3, where the residuals of the model are mapped to see where in the Netherlands the model performs well, this has similarly been done in the study of (Scholte et al., 2018). The distribution of the residuals can be found in appendix 6. Within the analysed dataset there were explicitly no cases from the province of Groningen itself since this study analysed the perception of individuals outside of the province. This needs to be taken into account when interpreting the map. The model's residuals are zero or close to zero around the area of the Randstad, suggesting that the model in this paper performs relatively good in this particular area of the Netherlands. It is interesting to see that the model underestimates the liveability score of the region of Groningen within the outskirts of the province of Zeeland. This could possibly have something to do with these areas of Zeeland also being perceived as remote and far away, and that therefore the individuals living within these areas are less

negatively biased towards the region of Groningen on liveability score and actually have a more positive perception towards the region of Groningen than the model suggests. Another interesting note is that in the east corner of the province of Friesland the model seems to overestimates the liveability score of the region of Groningen. Suggesting that the individuals within this region of Friesland are more negatively biased towards the region of Groningen than the model predicts. This could be explained by the socio-cultural differences across the Netherlands since the Netherlands is a relatively stratified country in this sense.

Figure 2: Expected liveability score based on statistical model 2 (own source)

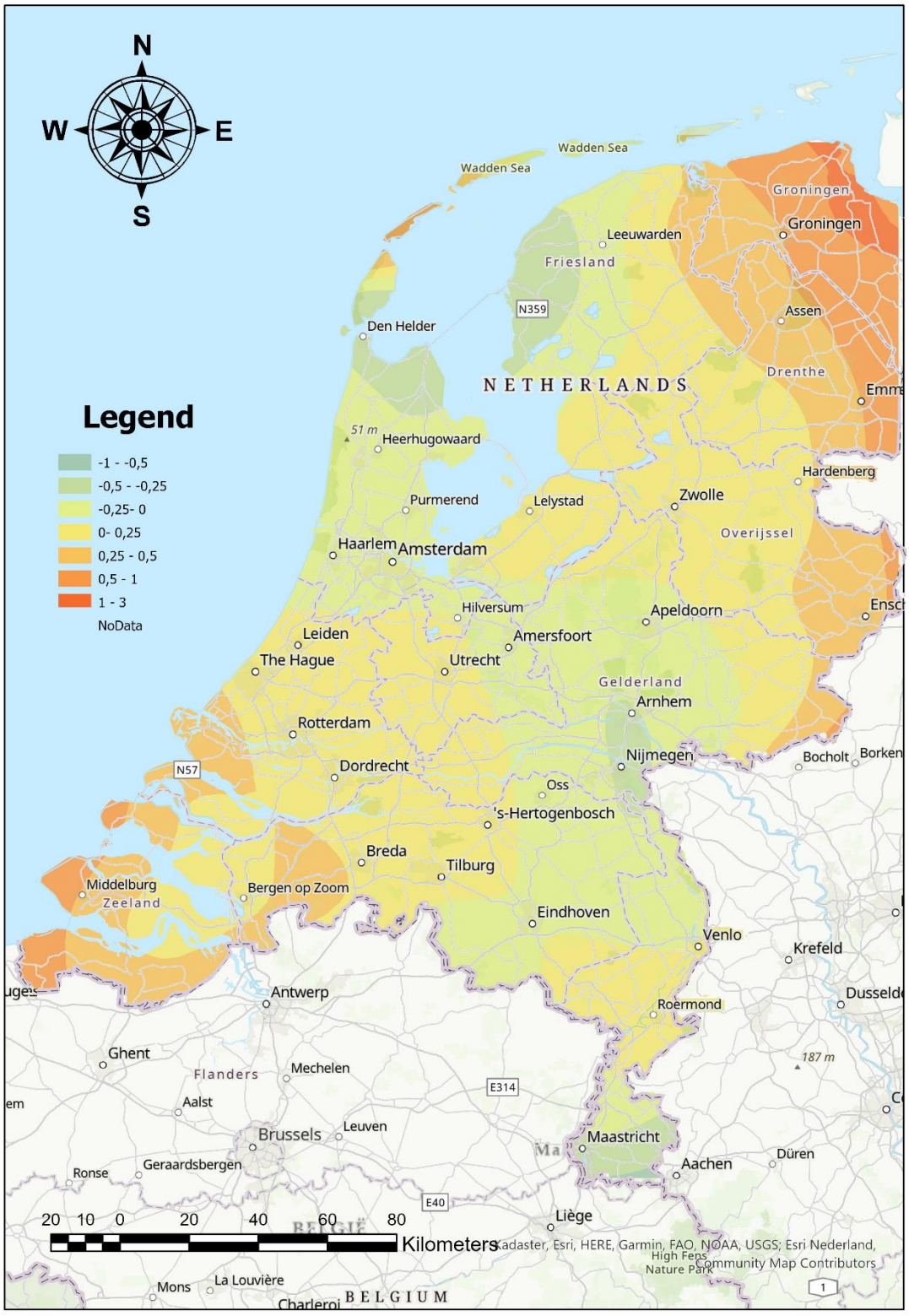


Figure 3: Mapped residuals based on statistical model 2 (own source). A difference of zero indicates a perfect performance of the model, any positive values suggest that the model underestimates intensity and negative values suggest that the model overestimates the intensity.

5 Discussion

Perceived liveability is influenced by local factors and one's associations with a place (Antognelli & Vizzari, 2016; Brander & Koetse, 2011; Cleave & Arku, 2020; Namazi-Rad et al., 2016; Norouzian-Maleki et al., 2015; Wheeler, 2001; Winters & Li, 2017). However, often overlooked are external factors related to an individual's current place of residence. Whereas literature showed degree of urbanisation and distance do play a role in shaping perceptions (Allen et al., 2002; Lee & Pistole, 2012; Lenzi & Perucca, 2021). This study examines how urbanisation degree and proximity to one's place of residence shape the perception of liveability in other areas, using the case of the region of Groningen. Quantifying and analysing subjective matters like liveability opinions is challenging (DeMaris, 1995). The study uses an ordered logistic regression and an OLS regression on a unique comprehensive dataset for grounded conclusions. However, the results are limited to the Dutch context, and future research could explore other case studies. The study has limitations, including a potential omitted variable bias due to uncontrolled work status and social background. Furthermore, amongst the respondents some interpretation error could have occurred when asking to assess the region of Groningen. Since the capital city and the province as a whole share the same name. Qualitative methods could have been beneficial for understanding liveability perception better, and incorporating more cultural control variables might offer additional insights.

5.1 Interpreting the spatial scale

This study has established the relationship between the perception of liveability elsewhere, and an individual’s current place of residence. Specifically, the distance between one’s current place of residence, and the place or region that individual is assessing on liveability. Both models shows that if distance (km) between place of residence and in this case the region of Groningen increases, the liveability score for Groningen goes down. The paper of Allen et al. (2002), presents that if more interactions occur between a person and a place, this will improve one’s perception and opinion’s on liveability of that place, these findings are in line with results presented when individuals did visited Groningen, since individuals who have visited Groningen before on average rate the region higher on liveability score. We also know from the paper of Allen et al. (2002) and results presented in KantarPublic (2022) that if the distance increases less interactions will occur. Therefore, one could conclude that when distance between place of residence and the place that is being assessed on liveability increases, the less

interactions these individuals have with that place or region. Thus, when distance towards that place increases less interactions will occur and therefore the perception of liveability about that place will decrease. This relationship between distance and perception of liveability is also established in the paper of Lee and Pistole (2012). Both regressions, the ordered logistic regression and the OLS regression show negative relationships between the liveability score and distance (km). Thus, confirming that when distance towards a place increases, the perception of liveability of that place will decrease.

Another spatial element studied in this research is the degree of urbanisation of the current place of residence. Both models show a significant relationship between the dependent variable: liveability score, and the key independent variable urbanization degree (measured in addresses per km2). The regressions concludes that when degree of urbanisation goes up, liveability score decreases. These findings are in line with corresponding literature (Lenzi & Perucca, 2021; Winters & Li, 2017), where the authors found that degree of urbanization matters when it comes to social wellbeing, and especially that one's notion towards less urbanized places is more negative when the individual itself lives in a more urbanized area already. Thus, when individuals current place of residence has a high degree of urbanization, they will on average give a lower liveability score when assessing the region of Groningen. In light of Winters and Li (2017), and findings in this paper, it is therefore expected that individuals who live in the Randstad will often have a more negative perception of liveability about the region of Groningen. Since the Randstad is a higher urbanized area than the region of Groningen. Suggesting that the degree of urbanization has a negative relationship with liveability towards less urban areas when assessing the liveability of in this study the region of Groningen (Lenzi & Perucca, 2021; Winters & Li, 2017).

Both regressions checked the relationship between local factors such as Safe, Green, Spacious, Healthy and Well Connected. The regressions shows that if a person thinks an area is either: Safe, Green, Spacious, Healthy and Well Connected their perceived liveability score of that place will go up. Implying that both regressions show a positive relationship between previously mentioned associations with a place and the liveability score of that place. Confirming the established relationship already presented in existing literature: (Antognelli & Vizzari, 2016; Brander & Koetse, 2011; Cleave & Arku, 2020; Namazi-Rad et al., 2016; Norouzian-Maleki et al., 2015; Wheeler, 2001; Winters & Li, 2017).

5.2 Disentangling the Liveability score

Earlier in the paper the liveability score which consists on the score of places to live, place to find a job and place to grow old has been taken apart into its three original parts. The results showed that the perception of liveability contains different segments that can be influenced by different factors. Where degree of urbanisation mattered when assessing the region of Groningen based on place to live or a place to find a job, it did not matter when assessing the region of Groningen based on it being a suitable place to grow old. However, the results showed increased significance of degree of urbanisation and distance when respondents were asked to assess the region of Groningen based on a place to find a job. Again, stressing that the different segments of liveability are influenced by different factors such as: degree of urbanisation of current place of residence, distance to the place of interest, and local factors relating to the situation within the place of interest. Particularly in this case when individuals will assess a region for its suitability to find a job. These are findings that policy makers of the region of Groningen can take into account when attempting to attract economic migrants looking for job opportunities.

5.3 Implications for policy makers

This paper has analysed the influence of urbanisation degree in one's current place of residence and distance between place of residence and the region of Groningen on the perceived liveability. After performing the regressions and the in-depth analysis on liveability, we can conclude that the degree of urbanisation in current place of residence and distance towards the region of Groningen from current place of residence play a role in the perception of liveability of the region of Groningen.

Furthermore, by analysing the different perceived liveability patterns per province and what drives the perception within the different provinces more adequate and precise strategies can be advises for policy makers of the region of Groningen. Table 9 in appendix 5 shows what associations have the most impact when respondents from the region of Noord-Holland were asked to assess liveability in the region of Groningen. This in-depth analysis was performed since the province of Noord-Holland showed to be the only province to have a negative and significant perception of liveability towards the region of Groningen. Therefore, it is interesting for policy makers what the respondents of the province of Noord-Holland find important when assessing liveability elsewhere. Table 9 shows that associations of: Safe, Spacious, Healthy

and Well Connected had a positive significant relationship with the liveability score. Especially the association with Healthy and Well Connected seemed to have the largest impact on the perception of liveability for the region of Groningen. Suggesting that if policy makers want to trigger more households from Noord-Holland to move to the region of Groningen it is advised that the perception of Healthy and Well connected is improved for the region. Since these are the biggest drivers for liveability for individuals from the region of Noord-Holland.

By creating a better understanding of what drives the perception of liveability, better marketing strategies can be developed by policy makers which are responsible for regional growth in the province of Groningen. By knowing what factors drive what perceptions of liveability, it being a place to find a job or a place to grow old. The policy makers can target specific groups better and therefore creating better place branding strategies. Knowing that place branding is a useful tool for regions and cities to attract highly skilled and educated migrants, that could bring potential economic benefits to a region. (Cleave & Arku, 2020), the findings in this paper can greatly contribute to existing and forthcoming marketing strategies for the region of Groningen. This paper has presented that it is not only essential for a region to ensure local factors that influence liveability greatly matter, but these perceptions that will determine if people move or visit other areas are also influenced by external factors which can't be influenced by the local policy makers. It is therefore of great importance to understand these external factors such as degree of urbanisation in order to adapt marketing and place branding strategies accordingly. Besides just policy makers of Groningen the findings in this paper are meaningful to other policy makers when setting up their place branding strategies, to be aware of the target group you want to approach and what is important for specific groups located in higher urbanized or more distant areas.

6 Conclusion

This paper investigates how geographical characteristics of an individual's place of residence influence the perception of liveability of another region. These geographical characteristics are degree of urbanisation, which was measured in number of addresses per square kilometre and distance between the region of interest and the respondents place of residence. This study used a unique dataset from the Imago Monitor Groningen (KantarPublic, 2022), which is a survey distributed in collaboration with the University of Groningen and Hanze University of Applied Science. This dataset contained information about individuals' perception of Groningen,

including liveability. Spatial data from the Dutch Central Bureau of Statistics was added to this existing dataset. The results of the multiple ran regressions presented the relationship between the dependent and independent variables. After the interpretation of the results and the discussion the three mentioned hypothesis cannot be rejected. Since firstly, the distance between current place of residence and the region of Groningen proved to have a significant impact on the perceived liveability of Groningen. Secondly, this paper presents a significant relationship between the degree of urbanisation of current place of residence and perceived liveability of the region of Groningen. And thirdly, the association of local factors such as Safety, Greenery, Economic infrastructure and social benefits towards the region of Groningen also present to have a significant impact on the perceived liveability of the region of Groningen. Therefore, we can conclude that individuals from more distant and urbanized areas to the place of interest give on average lower scores for liveability elsewhere, than individuals who lived in less urbanized and closer by areas to the place of interest. This thesis has successfully investigated that an individual's current place of residence, shapes their perception of liveability about another place. It is therefore imperative that local policy makers take these findings into account when attempting to attract households and migrants towards their region.

Acknowledgments

I would like to thank my supervisor dr. M.N. Daams for providing me with guidance and feedback during the process of writing my master thesis. Also, I would like to thank prof. dr. S. Koster for providing me with the uniquely comprehensive dataset to perform this study.

References

Abrudan, I.-N., Pop, C.-M., & Lazăr, P.-S. (2020). Using a general ordered logit model to explain the influence of hotel facilities, general and sustainability-related, on customer ratings. *Sustainability*, *12*(21), 9302.

Allen, M., Bourhis, J., Burrell, N., & Mabry, E. (2002). Comparing student satisfaction with distance education to traditional classrooms in higher education: A meta-analysis. *The American Journal of Distance Education*, *16*(2), 83-97.

Antognelli, S., & Vizzari, M. (2016). Ecosystem and urban services for landscape liveability: A model for quantification of stakeholders' perceived importance. *Land use policy*, *50*, 277-292.

Brander, L. M., & Koetse, M. J. (2011). The value of urban open space: Meta-analyses of contingent valuation and hedonic pricing results. *Journal of environmental management*, *92*(10), 2763-2773.

CBS. (2018). *Cijfers - Onderwijs*. https://longreads.cbs.nl/trends18/maatschappij/cijfers/onderwijs/#:~:text=In%20Nederland%20heeft%20bijna%2040,heen%20in%20omvang%20vrij%20stabiel.

CBS. (2020). *Kerncijfers per postcode*. https://www.cbs.nl/nl-nl/dossier/nederland-regionaal/geografische-data/gegevens-per-postcode#:~:text=Het%20gaat%20het%20om%20gegevens,demografische%20en%20sociaaleconomische%20variabelen%20gekoppeld.

CBS. (2022). *Groei en krimp per gemeente*. Retrieved 22/11/2022 from https://www.cbs.nl/nl-nl/visualisaties/dashboard-bevolking/regionaal/groei-en-krimp

CBS. (2023). *Regionale kerncijfers Nederland*. https://opendata.cbs.nl/statline/#/CBS/nl/dataset/70072NED/table?fromstatweb

Cleave, E., & Arku, G. (2020). Immigrant attraction through place branding? Evidence of city-level effectiveness from Canada's London. *Cities*, *97*, 102502.

Daams, M. N., Sijtsma, F. J., & Veneri, P. (2019). Mixed monetary and non-monetary valuation of attractive urban green space: A case study using Amsterdam house prices. *Ecological Economics*, *166*, 106430.

de Haan, F. J., Ferguson, B. C., Adamowicz, R. C., Johnstone, P., Brown, R. R., & Wong, T. H. (2014). The needs of society: A new understanding of transitions, sustainability and liveability. *Technological Forecasting and Social Change*, *85*, 121-132.

DeMaris, A. (1995). A tutorial in logistic regression. *Journal of Marriage and the Family*, 956-968.

Donnelly, M., Gamsu, S., & Whewall, S. (2020). Mapping the relational construction of people and places. *International Journal of Social Research Methodology*, *23*(1), 91-108.

Gentin, S. (2011). Outdoor recreation and ethnicity in Europe—A review. *Urban Forestry & Urban Greening*, *10*(3), 153-161.

Glaeser, E. L., Kolko, J., & Saiz, A. (2001). Consumer city. *Journal of economic geography*, *1*(1), 27-50.

Groningen. (2021). *'Er gaat niets boven Groningen' meest bekende slogan voor regiomarketing in Nederland*. Retrieved 21/11/2022 from

https://www.marketinggroningen.nl/nl/merk-groningen?ins_rd=d3d3Lm1lcmtncm9uaW5nZW4ubmw%3D

Groningen, N. P. (2020). *het imago van Groningen in 2020*. Retrieved 28/11/2022 from https://www.nationaalprogrammagroningen.nl/nieuws/2021/01/02/niet-groningers-positiever-over-groningen-dan-groningers-denken/

Kadaster. (2021). *Spanning tussen vraag en aanbod woningmarkt groeit*. https://www.kadaster.nl/-/spanning-tussen-vraag-en-aanbod-woningmarkt-groeit#:~:text=Woningen%20staan%20minder%20lang%20in,jaar%20naar%20ongeveer%2040%20jaar.

KantarPublic. (2022). *Imago Monitor Groningen*.

Lee, J.-y., & Pistole, M. C. (2012). Predictors of satisfaction in geographically close and long-distance relationships. *Journal of Counseling Psychology*, *59*(2), 303.

Lenzi, C., & Perucca, G. (2021). Not too close, not too far: Urbanisation and life satisfaction along the urban hierarchy. *Urban Studies*, *58*(13), 2742-2757.

Lloyd, K., Fullagar, S., & Reid, S. (2016). Where is the 'social'in constructions of 'liveability'? Exploring community, social interaction and social cohesion in changing urban environments. *Urban Policy and Research*, *34*(4), 343-355.

Lu, X., & White, H. (2014). Robustness checks and robustness tests in applied economics. *Journal of econometrics*, *178*, 194-206.

Meester, W. J., & Pellenbarg, P. H. (2006). The spatial preference map of Dutch entrepreneurs: subjective rating of locations, 1983, 1993 and 2003. *Tijdschrift voor economische en sociale geografie*, *97*(4), 364-376.

Mulder, C. H., & Cooke, T. J. (2009). Family ties and residential locations. *Population, space and place*, *15*(4), 299-304.

Namazi-Rad, M.-R., Perez, P., Berryman, M., & Wickramasuriya, R. (2016). A semi-empirical determination of perceived liveability. *Bulletin of Sociological Methodology/Bulletin de Méthodologie Sociologique*, *129*(1), 5-24.

Norouzian-Maleki, S., Bell, S., Hosseini, S.-B., & Faizi, M. (2015). Developing and testing a framework for the assessment of neighbourhood liveability in two contrasting countries: Iran and Estonia. *Ecological indicators*, *48*, 263-271.

Norouzian-Maleki, S., Bell, S., Hosseini, S.-B., Faizi, M., & Saleh-Sedghpour, B. (2018). A comparison of neighbourhood liveability as perceived by two groups of residents: Tehran, Iran and Tartu, Estonia. *Urban forestry & urban greening*, *35*, 8-20.

PBL. (2015). Plan Bureau voor de Leefomgeving. https://www.pbl.nl/sites/default/files/downloads/PBL_2015_De_stad_verbeeld_1744.pdf

Scholte, S. S., Daams, M., Farjon, H., Sijtsma, F. J., van Teeffelen, A. J., & Verburg, P. H. (2018). Mapping recreation as an ecosystem service: Considering scale, interregional differences and the influence of physical attributes. *Landscape and Urban Planning*, *175*, 149-160.

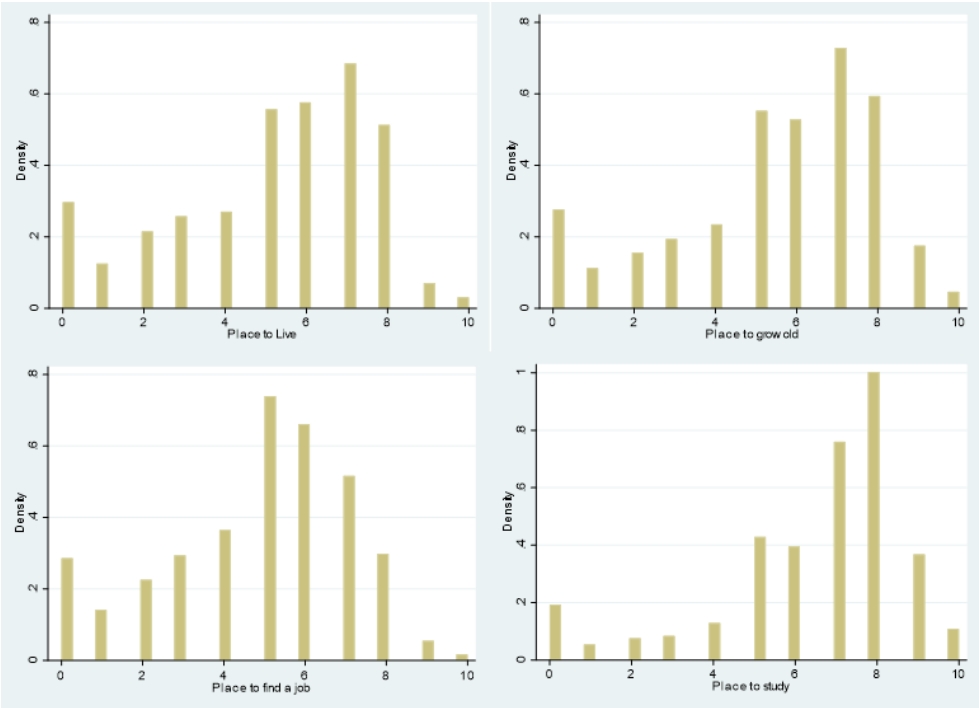
Stata. (2021). *Stata Data Analysis Examples*. Retrieved 02/07/2023 from https://stats.oarc.ucla.edu/stata/dae/truncated-regression/

Statology. (2021). *How to Interpret Log-Likelihood Values*. https://www.statology.org/interpret-log-likelihood/

Tobler, W. R. (1970). A computer movie simulating urban growth in the Detroit region. *Economic geography*, *46*(sup1), 234-240.
Wheeler, S. (2001). Livable communities: Creating safe and livable neighborhoods, towns, and regions in California.
Wheeler, S. (2013). *Planning for sustainability: creating livable, equitable and ecological communities*. Routledge.
Winters, J. V., & Li, Y. (2017). Urbanisation, natural amenities and subjective well-being: Evidence from US counties. *Urban Studies*, *54*(8), 1956-1973.

Appendixes

A1 Histogram Liveability scores



A2 Correlation matrix

Correlation Matrix

	Place to live	Place to find a job	Place to grow old	Place to study
Place to live	1.0000			
Place to find a job	.7929	1.0000		
Place to grow old	.7937	.6921	1.0000	
Place to study	.5811	.6921	.6233	1.0000

A3 OLS regression

Table 4
OLS regression results

	(1)	(2)	(3)	(4)
Addresses km2 (10,000)	-.567**(.249)	-.347 (.253)	-.374*(.140)	-.408*(.229)
Distance (km) (50)		-.201***(.028)	-.112***(.018)	-.112***(.267)
Safe (ref. No)			.626***(.080)	.622***(.080)
Green (ref. No)			.177*(.097)	.183*(.098)
Spacious (ref. No)			.268**(.064)	.255**(.125)
Healthy (ref. No)			1.018***(.056)	1.009***(.089)
Well Connected (ref. No)			.929***(.061)	.938***(.062)
Visited_Gro (ref. No)			.566***(.076)	.540***(.078)
Sex (ref. male)				-.075 (.060)
Age (ref. 19 – 29 years)				
30 – 39 years				. 066(.140)
40 – 49 years				-.041 (.132)
50- 59 years				-.088 (.124)
60 – 69 years				.026 (.120)
70 – 79 years				.122 (.135)
80 > years				.635** (.231)
Education level (ref. primary school)				
VMBO				-.304(.226)
HAVO/VWO				-.039 (.229)
MBO				-.155 (.218)
HBO				-.133 (.221)
University				-.137 (.240)
Constant	5.244	4.875	3.036	3.245
Observations	4,579	4,579	4,579	4,579
R-Squared	.04	.12	.192	.196

Note: Dependent variable is the liveability score, addresses per km (10,000) is the variable addresses per km2 / 10,000, Additionally Distance (50km) is the variable of distance per km / 50. *, **, *** Significance at 10%, 5%, and 1%, respectively

A4 Decomposition with OLS regression

Table 7
Decomposition liveability score

	Place to live (A)	Place to find a job (B)	Place to grow old (C)
Addresses km2 (10,000)	-.464* (.252)	-.553** (.242)	-.208 (.265)
Distance km (50)	-.081***(.031)	-.199***(.029)	-.058* (.030)
Safe (ref. No)	.711***(.091)	.432***(.085)	.724***(.092)
Green (ref. No)	.201* (.110)	-.042 (.104)	.391***(.110)
Spacious (ref. No)	.217* (.140)	.064 (.136)	.485**(.137)
Healthy (ref. No)	.992***(.101)	.834***(.093)	1.202***(.104)
Well Connected (ref. No)	.905***(.071)	1.148***(.066)	.763***(.071)
Visited_Gro (ref. No)	.684***(.086)	.317***(.082)	.620***(.087)
Socioeconomic controls	Yes	Yes	Yes
R-Squared	.166	.154	.188

Note: Dependent variables change per model as A) Place to Live, B) Place to find a job and C) Place to grow old respectively. Number of observations for each panel (A, B, C) is 4,579. Addresses per km (10,000) is the variable addresses per km2 / 10,000, Additionally Distance km (50km) is the variable of distance per km / 50. Socioeconomic controls consist of respondent Sex, Age and Education Level. *, **, *** Significance at 10%, 5%, and 1%, respectively

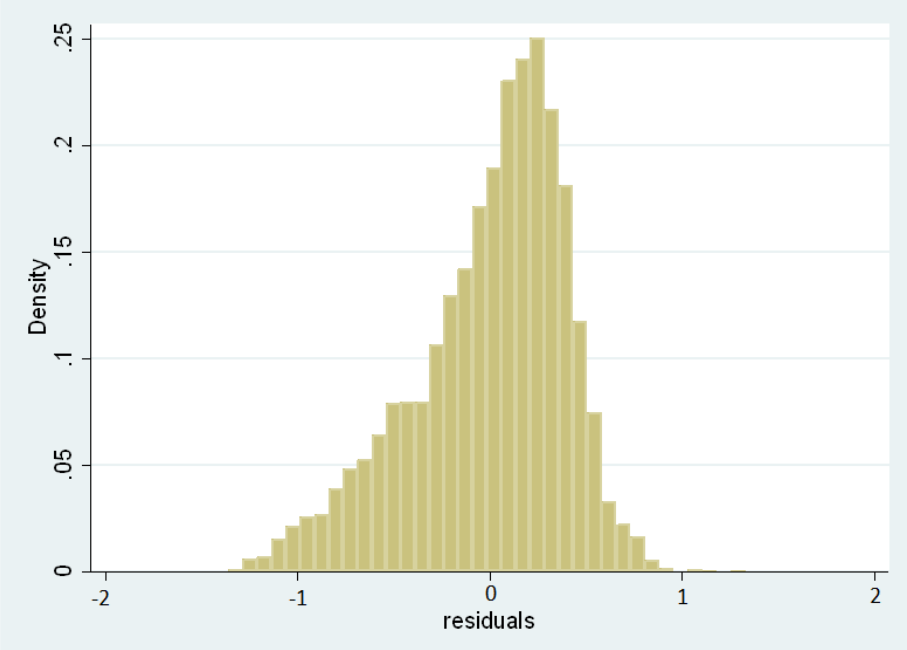
A5 Drivers of liveability perception in Noord-Holland

Table 9
Drivers of liveability perception in Noord-Holland

	Ordered Logistic Regression (A)	OLS Regression (B)
Safe (ref. No)	1.493**(.279)	.476** (.230)
Green (ref. No)	1.053 (.250)	.072 (.272)
Spacious (ref. No)	1.543 (.143)	.603 (343)
Healthy (ref. No)	2.179*** (.171)	1.011***(.256)
Well Connected (ref. No)	2.298*** (.449)	.958***(.169)
Visited_Gro (ref. No)	1.757*** (.310)	.641**(.210)
Socioeconomic controls	Yes	Yes
Constant		2.448
Observations	691	
Model Chi2	0.0000	
Log likelihood	-1,036.3	
R-Squared		0.182

Note: Dependent variable is the liveability score categorized in model A for ordered logistic regression and the liveability score in model B for the OLS regression. The coefficients in model 1 represent odds ratios Socioeconomic controls consists of respondent Sex, Age and Education Level. Within this in-depth analysis only the respondents of the province of Noord-Holland are used. *, **, *** Significance at 10%, 5%, and 1%, respectively

A6 Histogram residuals OLS regression



A7 Stata do-file

```
clear all
cd "C:\Scriptie\Data"
import spss "NietGroningers.sav", clear

* see distribution of respondents*
tabulate v5provincie

*drop unnecessary variables*
drop firstname lastname email emailstatus language sent remindersent remindercount usesleft attribute_1 submitdate lastpage startlanguage startdate datestamp

tabulate v9

tabulate v11_SQ001

drop v11_SQ002 v11_SQ003 v11Extra4_SQ004 v11Extra5_SQ005 v11Extra6_SQ006 v11Extra7_SQ007 v11Extra8_SQ008 v11Extra9_SQ009 v11Extra10_SQ010

tabulate v12Extra1_SQ001

drop v12Extra2_SQ002 v12Extra3_SQ003 v12Extra3_SQ003 v12Extra4_SQ004 v12Extra5_SQ005 v12Extra6_SQ006 v12Extra7_SQ007 v12Extra8_SQ008 v12Extra9_SQ009 v12Extra10_SQ010

drop if missing(v13_SQ001)
drop if missing(v10)

drop v14_SQ001 v15_SQ001
drop v17_SQ002 v17_SQ003 v17_SQ004 v17_SQ005 v17_SQ006 v17_SQ007 v17_SQ012 v17_SQ013 v17_SQ014 v17_SQ016 v17_SQ019 v17_SQ020 v17_SQ021 v23Anders v5GrunWonenDuur v6extra v7extra v4

*eyeballing the data*
tabulate v13_SQ001
tabulate v23
tabulate v24
tabulate v5GrunWonen
tabulate v6
tabulate v14_SQ002
tabulate v12Extra1_SQ001
tabulate v16_SQ001

sum v14_SQ002

*summing "liveability" *
sum v16_SQ001 v16_SQ002 v16_SQ003 v16_SQ004 v16_SQ005 v16_SQ006
```

```
drop if missing(v16_SQ001, v16_SQ002, v16_SQ003, v16_SQ004, v16_SQ005, v16_SQ006)

tabulate v17_SQ001
tabulate v17_SQ008
tabulate v17_SQ009
tabulate v17_SQ010
tabulate v17_SQ011
tabulate v17_SQ015
tabulate v17_SQ017

*renaming variables*
rename v1 Sex
rename v2 Age
rename v5provincie Province
rename v3 Education_level
rename v9 Household_Composition
rename v10 Zipcode
rename v16_SQ001 Place_to_Live
rename v16_SQ002 Place_to_grow_up
rename v16_SQ003 Place_to_study
rename v16_SQ004 Place_to_grow_old
rename v16_SQ005 Place_to_find_a_job
rename v16_SQ006 Place_to_undertake
rename v16_SQ007 Place_to_relax
rename v16_SQ008 Place_to_experience_culture
rename v17_SQ001 Spacious
rename v17_SQ008 Pleasant
rename v17_SQ009 Green
rename v17_SQ010 Quality_living
rename v17_SQ011 Sustainable
rename v17_SQ015 Healthy
rename v17_SQ017 Safe
rename v17_SQ018 Young
rename v17_SQ022 Bustling
rename v17_SQ023 Talent
rename v17_SQ024 Well_Connected

*merging with spatial data*
merge m:1 Zipcode using zipcodedata

*dropping missing values*
drop if missing(id)
drop if missing(sted)
drop if missing(Spacious)
drop if missing(Pleasant)
drop if missing(Green)
drop if missing(Quality_living)
drop if missing(Sustainable)
drop if missing(Healthy)
drop if missing(Safe)
```

```
drop if missing(Young)
drop if missing(Bustling)
drop if missing(Talent)
drop if missing(Well_Connected)
drop if missing(v7)
drop if missing(v23)
drop if missing(v5GrunWonen)
drop if Sex == 3
drop if Sex == 4

*destring distance variable*
split near_dist, destring p(,)

*dropping unnecessary varialbles*
drop aantal opp_totaal near_fid opp_woonfu opp_winkel opp_sportf opp_overig opp_onderw opp_logies opp_bij aantal_bij opp_gezond opp_indust opp_kantoo opp_celfun oid_ token aantal_ove aantal_log aantal_cel aantal_vbo aantal_sta aantal_lig fid_ pc4 _merge near_dist2 won_hcorp won_nbew v13_SQ001 woning v24

*renaming variables*
rename inwoner Inhabitants
rename man Male
rename vrouw Female
rename aantal_woo Amount_houses
rename p_koopwon p_owneroccupied
rename p_huurwon p_rental
rename sted Urbanisation
rename near_dist1 Distance
rename v5GrunWonen Lived_in_Gro
rename v6 Work_visit_Gro
rename v7 Visited_Gro
rename v11_SQ001 First_association
rename v14_SQ002 Change_image_past
rename v15_SQ002 Change_image_future
rename v21_SQ001 Pride_place_res
rename v22_SQ001 Pride_province_res
rename aantal_gez Am_healthcare
rename aantal_ind Am_industrial
rename aantal_kan Am_office
rename aantal_ond Am_education
rename aantal_spo Am_sport
rename aantal_win Am_retail
rename oad Adresses_km2

sum Distance

*dropping last unnecessary variables*
drop near_dist
drop if p_owneroccupied < 0
drop if p_rental < 0
```

```
drop if wozwoning < 0
drop Work_visit_Gro
drop Amount_houses
drop Inhabitants
drop Male
drop Female
drop if missing(Pride_place_res)
drop if missing(Pride_province_res)
drop Place_to_relax
drop Place_to_experience_culture

*create distance in km variable*
gen Distance_km = Distance / 1000

gen Distance_km10 = Distance_km / 10

gen Distance_km50 = Distance_km / 50

gen Distance_km100 = Distance_km / 100

* generate higher variable with urbanisation degree*
gen Urbanisation1000 = Adresses_km2/ 1000

gen Urbanisation10000 = Adresses_km2/ 10000

summarize
tab Province

*correlation matrix*
corr Place_to_Live Place_to_find_a_job Place_to_grow_old Place_to_study

egen mean_score2 = rowmean(Place_to_Live Place_to_find_a_job Place_to_grow_old)

* label urbanisation variable*
label define urbanisation_label 1 "Very Highly Urban" 2 "Highly Urban" 3 "Moderately Urban" 4 "Little Urban" 5 "Not Urban"

label values Urbanisation urbanization_label

* create natural log of distance*
gen ln_Distance_km = ln(Distance_km)

* create natural log of y variable*
gen ln_mean_score2 = ln(mean_score2)

* create catagorical 5 point y variable with labels*
egen Place_to_Live_cut = cut(mean_score2), group(5)
label define attractiveness_label 0 "Very Unattractive" 1 "Unattractive" 2 "Neutral" 3 "Attractive" 4 "Very Attractive"
label values Place_to_Live_cut attractiveness_label
```

```
* create catagorical 4 point y variable with labels*
egen Place_to_Live_cut4 = cut(mean_score2), group(4)
label define attractiveness_label4 0 "Very Unattractive" 1 "Unattractive" 2 "Attractive" 3 "Very Attractive"

label values Place_to_Live_cut4 attractiveness_label4

*make other labes*
label variable Place_to_Live "Place to Live"
label variable Place_to_grow_old "Place to grow old"
label variable Place_to_find_a_job "Place to find a job"
label variable Place_to_study "Place to study"

* create dummies for provinces*

tabulate Province, gen(Province_dummy)
label define province_labels 1 "Drenthe" 2 "Groningen" 3 "Flevoland" 4 "Friesland" 5 "Gelderland" 6 "Limburg" 7 "Noord-Brabant" 8 "Noord-Holland" 9 "Overijssel" 10 "Utrecht" 11 "Zeeland" 12 "Zuid-Holland"
rename Province_dummy1 Dummy_Drenthe
rename Province_dummy2 Dummy_Flevoland
rename Province_dummy3 Dummy_Friesland
rename Province_dummy4 Dummy_Gelderland
rename Province_dummy5 Dummy_Limburg
rename Province_dummy6 Dummy_Noord_Brabant
rename Province_dummy7 Dummy_Noord_Holland
rename Province_dummy8 Dummy_Overijssel
rename Province_dummy9 Dummy_Utrecht
rename Province_dummy10 Dummy_Zeeland
rename Province_dummy11 Dummy_Zuid_Holland

save scriptiedata, replace

*descriptives
tab Safe
tab Green
tab Spacious
tab Healthy
tab Well_Connected
tab Visited_Gro

*truncated regression*
regress mean_score2 Adresses_km2 Distance_km ib2.Safe ib2.Green ib2.Spacious ib2.Healthy ib2.Well_Connected ib2.Visited_Gro ib1.Sex ib2.Age ib1.Education_level

regcheck

truncreg mean_score2 Urbanisation10000, ul(11) ll(-1) r
```

```
truncreg mean_score2 Urbanisation10000 Distance_km50, ul(11) ll(-1) r

truncreg mean_score2 Urbanisation10000 Distance_km50 ib2.Safe ib2.Green ib2.Spacious ib2.Healthy ib2.Well_Connected ib2.Visited_Gro, ul(11) ll(-1) r

truncreg mean_score2 Urbanisation10000 Distance_km50 ib2.Safe ib2.Green ib2.Spacious ib2.Healthy ib2.Well_Connected ib2.Visited_Gro ib1.Sex ib2.Age ib1.Education_level, ul(11) ll(-1) r

*OLS*
regress mean_score2 Urbanisation10000, r

regress mean_score2 Urbanisation10000 Distance_km50, r

regress mean_score2 Urbanisation10000 Distance_km50 ib2.Safe ib2.Green ib2.Spacious ib2.Healthy ib2.Well_Connected ib2.Visited_Gro, r

regress mean_score2 Urbanisation10000 Distance_km50 ib2.Safe ib2.Green ib2.Spacious ib2.Healthy ib2.Well_Connected ib2.Visited_Gro ib1.Sex ib2.Age ib1.Education_level, r
* ordered logistic regression *

ologit  Place_to_Live_cut Urbanisation10000 , or

ologit  Place_to_Live_cut Urbanisation10000 Distance_km50, or

ologit  Place_to_Live_cut Urbanisation10000 Distance_km50 ib2.Safe ib2.Green ib2.Spacious ib2.Healthy ib2.Well_Connected ib2.Visited_Gro, or

ologit  Place_to_Live_cut Urbanisation10000 Distance_km50 ib2.Safe ib2.Green ib2.Spacious ib2.Healthy ib2.Well_Connected ib2.Visited_Gro ib1.Sex ib2.Age ib1.Education_level, or

*robustness check*

truncreg mean_score2 Urbanisation10000 Distance_km50 ib2.Safe ib2.Green ib2.Spacious ib2.Healthy ib2.Well_Connected ib2.Visited_Gro ib1.Sex ib2.Age ib1.Education_level if Province != 12, ul(11) ll(-1) r

ologit Place_to_Live_cut Urbanisation10000 Distance_km50 ib2.Safe ib2.Green ib2.Spacious ib2.Healthy ib2.Well_Connected ib2.Visited_Gro ib1.Sex ib2.Age ib1.Education_level if Province != 12, or


* in depth analysis*

*1*
regress Place_to_Live Urbanisation10000 Distance_km50 ib2.Safe ib2.Green ib2.Spacious ib2.Healthy ib2.Well_Connected ib2.Visited_Gro ib1.Sex ib2.Age ib1.Education_level, r

regress Place_to_Live Dummy_Drenthe Dummy_Flevoland Dummy_Friesland Dummy_Gelderland Dummy_Limburg Dummy_Noord_Brabant Dummy_Noord_Holland
```

Dummy_Overijssel Dummy_Utrecht Dummy_Zeeland Dummy_Zuid_Holland ib2.Safe ib2.Green ib2.Spacious ib2.Healthy ib2.Well_Connected ib2.Visited_Gro ib1.Sex ib2.Age ib1.Education_level, r

2
regress Place_to_find_a_job Urbanisation10000 Distance_km50 ib2.Safe ib2.Green ib2.Spacious ib2.Healthy ib2.Well_Connected ib2.Visited_Gro ib1.Sex ib2.Age ib1.Education_level, r

regress Place_to_find_a_job Dummy_Drenthe Dummy_Flevoland Dummy_Friesland Dummy_Gelderland Dummy_Limburg Dummy_Noord_Brabant Dummy_Noord_Holland Dummy_Overijssel Dummy_Utrecht Dummy_Zeeland Dummy_Zuid_Holland ib2.Safe ib2.Green ib2.Spacious ib2.Healthy ib2.Well_Connected ib2.Visited_Gro ib1.Sex ib2.Age ib1.Education_level, r

3
regress Place_to_grow_old Urbanisation10000 Distance_km50 ib2.Safe ib2.Green ib2.Spacious ib2.Healthy ib2.Well_Connected ib2.Visited_Gro ib1.Sex ib2.Age ib1.Education_level, r

regress Place_to_grow_old Dummy_Drenthe Dummy_Flevoland Dummy_Friesland Dummy_Gelderland Dummy_Limburg Dummy_Noord_Brabant Dummy_Noord_Holland Dummy_Overijssel Dummy_Utrecht Dummy_Zeeland Dummy_Zuid_Holland ib2.Safe ib2.Green ib2.Spacious ib2.Healthy ib2.Well_Connected ib2.Visited_Gro ib1.Sex ib2.Age ib1.Education_level, r

in depth analysis with province as dummies
regress mean_score2 Urbanisation10000 Dummy_Drenthe Dummy_Flevoland Dummy_Friesland Dummy_Gelderland Dummy_Limburg Dummy_Noord_Brabant Dummy_Noord_Holland Dummy_Overijssel Dummy_Utrecht Dummy_Zeeland Dummy_Zuid_Holland ib2.Safe ib2.Green ib2.Spacious ib2.Healthy ib2.Well_Connected ib2.Visited_Gro ib1.Sex ib2.Age ib1.Education_level, r

ologit Place_to_Live_cut Urbanisation10000 Dummy_Drenthe Dummy_Flevoland Dummy_Friesland Dummy_Gelderland Dummy_Limburg Dummy_Noord_Brabant Dummy_Noord_Holland Dummy_Overijssel Dummy_Utrecht Dummy_Zeeland Dummy_Zuid_Holland ib2.Safe ib2.Green ib2.Spacious ib2.Healthy ib2.Well_Connected ib2.Visited_Gro ib1.Sex ib2.Age ib1.Education_level, or

*in depth analysis of Noord-Holland *
regress mean_score2 ib2.Safe ib2.Green ib2.Spacious ib2.Healthy ib2.Well_Connected ib2.Visited_Gro ib1.Sex ib2.Age ib1.Education_level if Province == 8, r

prediction model
regress mean_score2 Adresses_km2 Distance_km ib2.Safe ib2.Green ib2.Spacious ib2.Healthy ib2.Well_Connected ib2.Visited_Gro ib1.Sex ib2.Age ib1.Education_level, r

estimates store mymodel

```
predict predict_score, xb

gen residuals = mean_score2 - predict_score

sum predict_score


save scriptiedata, replace


drop v8_SQ001 v8_SQ002 v8_SQ003 v8_SQ004 v8_SQ005 v8_SQ006 v8_SQ007 v8_SQ008
v8_SQ009 v8anders
```