# Ethnic neighbourhood segregation and incomes of Polish immigrants in the Netherlands

*Exploring patterns of neighbourhood segregation and its consequences by using individualised, scalable neighbourhoods*

Simon Douwe Venema
s2240734

Master in Social Demography

Supervisors: prof. dr. Helga de Valk & prof. Bruno Arpino
Daily supervision: Jeroen Ooijevaar & dr. Bart Sleutjes

University of Groningen, Faculty of Spatial Sciences, Department of Demography
Universitat Pompeu Fabra, Faculty of Political and Social Science, Department of Political and Social Science



29-7-2016

# Contents

# List of Figures



# List of Tables

## Glossary

**GIS** Geographic Information System.

**MAUP** Modifiable Areal Unit Problem.

**SSD** System of Social Statistical Datasets.

**Abstract**

The Polish population in the Netherlands has grown rapidly since the accession of Poland to the European Union in 2004. Previous work on this relatively new migrant group has largely focused on their demographic and socio-economic characteristics, but hardly on their settlement patterns. This thesis analyses the degree to which Polish migrants in the Netherlands live in ethnically segregated neighbourhoods, the location of these neighbourhoods and the consequences for the incomes of Polish migrants who live in these neighbourhoods by applying the method of individualised neighbourhoods. Geocoded register data from Statistics Netherlands of 2012 allows for computation of individualised neighbourhoods on different scales, based on an individual's 50 to 51,200 nearest neighbours. The results indicate that strong concentrations of Polish migrants can be found in the western, and to a lesser degree in the southern and south-eastern part of the country, and that the role of scale in segregation patterns varies across municipalities. Fixed effects regression analyses applied on a total of 62,197 Polish migrants show that there is a negative relationship between small scale ethnic segregation and the incomes of Polish migrants which can't be captured when using administratively defined neighbourhoods. The analyses provide support for the notion that segregation patterns and neighbourhood effects are of a multi-scalar nature.

**Keywords:** *neighbourhood segregation, individualised neighbourhoods, Polish migrants, income, scale, multi-scalar segregation, micro-level segregation*

# 1 Introduction

Since the accession of Poland to the European Union in May 2004, immigration rates of Polish migrants to the Netherlands have surged. In 2014, the Netherlands received 24 thousand Polish immigrants, which makes Polish migrants the largest new immigrant group in the country (Statistics Netherlands, 2015). Recent research shows that Polish migrants often work in industry, construction and agriculture and generally have low-wage jobs characterized by precarious working conditions (Dagevos, 2011; Engbersen, Leerkes, Ilies, Snel, & Meij, 2011; Gijsberts & Lubbers, 2015; Van den Brakel et al., 2014). The experiences of Polish migrants in the Netherlands are not unique. Countries such as Germany, Ireland, Sweden, Norway and the United Kingdom also have a substantial Polish population with similar labour market characteristics (Drinkwater, Eade, & Garapich, 2009; Engbersen, Okolski, Black, & Pantîru, 2010). Despite relatively the high rates of return migration among Polish migrants (Nicolaas, 2011), about half of the Polish migrants have the intention to settle in the Netherlands (Gijsberts & Lubbers, 2015). Little research, however, addresses the location where Polish migrants settle in the Netherlands.

The question arises whether Polish migrants, like many other immigrant groups in the Netherlands live in ethnically segregated neighbourhoods. Much research has been dedicated to studying the settlement patterns of migrants, which shows that migrants are generally not equally spread over countries and cities, but live in rather ethnically concentrated areas (Musterd, 2005). Living in these ethnically segregated neighbourhoods is thought to have negative consequences for the socio-economic integration of immigrants (Musterd, 2003). It is unclear, however, whether this is also true for Polish migrants. This thesis assesses the extent to which Polish migrants live in ethnically segregated neighbourhoods and analyses the consequences of living in such neighbourhoods for the incomes of Polish migrants.

The idea that "living in deprived neighbourhoods has a negative effect on residents' life chances over and above the effect of their individual characteristics" (Van Ham, Manley, Bailey, Simpson, & MacLennan, 2011, p. 1) are called neighbourhood effects. Despite the strong theoretical foundations of neighbourhood effects, the empirical work on neighbourhood effects hardly provide a uniform picture of its impact on individual outcomes (R. Andersson, Musterd, Galster, & Kauppinen, 2007; E. K. Andersson & Malmberg, 2015; Musterd, 2003; Musterd, Andersson, Galster, & Kauppinen, 2008; Van Ham et al., 2011). Results often contradict each other, and the size of the effects are generally moderate in magnitude. Understanding whether neighbourhood effects are present and how they operate is of key importance given the ample attention that neighbourhood segregation and its assumed negative effects receives by scholars, policy makers and politicians on both sides of the Atlantic (Musterd, 2005).

A recent attempt to shed light on the ambiguous results regarding neighbourhood effects has been carried out by Andersson and Malmberg (2015). They argue that the lack of quantitative support for neighbourhood effects may be due

to methodological errors in the measurement of neighbourhood effects. In the majority of studies, neighbourhoods are defined by administrative areas, such as census tracts or administrative neighbourhoods. Andersson and Malmberg (2015) propose a new methodological technique called individualised neighbourhoods, in which neighbourhoods are defined as a buffer around an individual on the basis of the individual's $k$ nearest neighbours, rather than on administrative areas. The value of $k$ is allowed to vary in terms of the amount of nearest neighbours included in the individualised neighbourhood, allowing the size of the buffer created around the individual to vary by scale. This methodological approach is in line with the notion that neighbourhood segregation and neighbourhood effects are phenomena that take place and have different characteristics on different scales (W. Clark, Andersson, Östh, & Malmberg, 2015; Fowler, 2016; Jones, Johnston, Manley, Owen, & Charlton, 2015; Sleutjes & De Valk, 2015). In their analysis, Andersson and Malmberg (2015) report that the methodology of individualised neighbourhoods finds neighbourhood effects which are three times stronger in magnitude compared to conventional area-based measurements of neighbourhoods, in which neighbourhoods are measured by administratively defined areas.

The method of individualised neighbourhoods has not yet been applied to analyse how the composition of the neighbourhood in terms of ethnic segregation has economic consequences for immigrants (for a notable exception see Van Ham, Hedman, Manley & Östh, 2014). This thesis will add to the literature by applying the method of individualised neighbourhoods to the case of Polish immigrants in the Netherlands. The use of the individualised neighbourhoods can extend our understanding of the relatively unexplored role of scale in neighbourhood segregation, an aspect often overlooked in the neighbourhood effects literature (Lupton & Kneale, 2012; Manley & Van Ham, 2012). Moreover, the use of individualised neighbourhoods circumvents the problem of spatial dependency of segregation measures (known as the modifiable areal unit problem, or MAUP) by defining personal neighbourhoods in the same way for all individuals (Östh, Malmberg, & Andersson, 2014). Overcoming this problem allows for a sound comparison of neighbourhood segregation measures over different areas.

The case of Polish migrants in the Netherlands first of all provide a unique opportunity to apply this method to a relatively new immigrant group, for which limited information on their residential patterns is available. Secondly, disentangling whether the degree of ethnic segregation of the neighbourhood affects the incomes of Polish migrants will add to the knowledge of the economic integration experience of this new and rapidly growing migrant group.

The objective of this thesis is threefold. First, residential patterns in terms of ethnic segregation of Polish migrants in the Netherlands are explored. Secondly, this thesis analyses how the method of individualised neighbourhoods relates to conventional area-based measurements of segregation for this migrant group in the Netherlands. Lastly, the influence of ethnic neighbourhood segregation on the incomes of Polish migrants is analysed by using the individualised neighbourhoods.

The three research questions of this paper are thus:

1. Where in the Netherlands do Polish immigrants live and to what extent are the neighbourhoods they live in ethnically segregated?

2. Does the identification of ethnic segregation depend on the measurement of neighbourhoods and to what extent are different patterns of segregation found by using the individualised neighbourhoods-approach compared to using administrative neighbourhoods in the case of Polish migrants in the Netherlands?

3. How and to what extent does ethnic neighbourhood segregation affect the incomes of Polish migrants in the Netherlands and how is this relationship affected by the use of individualised neighbourhoods compared to using administrative neighbourhoods?

To address the research questions data from the Systems of social statistical datasets (SSD) that cover the entire population is used (Bakker, Van Rooijen, & Van Toor, 2014). Having high quality register data is a prerequisite to perform an analysis based on the method of individualised neighbourhoods, which requires geocoded data for all individuals. The register data allows for the identification and a detailed description of the residential areas in which Polish migrants live. The effect of ethnic neighbourhood segregation of Polish migrants on their incomes will be analysed by performing a fixed effects regression.

# 2 Context: Polish immigrants in the Netherlands

Due to the vastly increasing Polish population in the Netherlands since the accession of Poland to the European Union in May 2004, a growing body of literature addressing the position of Polish migrants in the Dutch society after migration is emerging. This literature covers various aspects of the Polish immigrants' lives, such as the motivation for migration, living arrangements, labour market position, demographic behaviour, education and proficiency in the Dutch language. The following paragraph provides a brief overview on these topics.

Many Polish migrants arrive in the Netherlands through special recruitment agencies, which are often located in Poland (Engbersen et al., 2011). Indeed, the majority of migration from Poland to the Netherlands is driven by labour motivations, although family motivated migration seems to be slowly increasing (Gijsberts & Lubbers, 2013). It is not uncommon that employment agencies also provide housing for the migrants. A prime example are so-called 'Polish hotels', referring to accommodations that are specifically set up for the purpose of housing Polish migrants, which are mainly located in rural areas (Engbersen et al., 2012). The majority of Polish migrants in the Netherlands live in areas of high agricultural activity (Gijsberts & Lubbers, 2013). Polish migrants who live in more urban areas, however, tend to live in racially mixed areas, characterized by high degrees of non-western immigrant concentrations and high levels

of socio-economic deprivation (Snel, 2011). The available research, however, does not explicitly address whether Polish migrants live in concentrated areas amongst themselves.

These locational patterns are reflected by the sectors in which Polish migrants are employed, as the majority of Polish migrants in the Netherlands work in agriculture, industry and construction (Gijsberts & Lubbers, 2013, 2015). These are generally temporary, low wage and low prestige jobs (Dagevos, 2011; Gijsberts & Lubbers, 2015). The rate of unemployment and self-employment among Polish migrants is relatively low. When Polish migrants are unemployed re-entry into the labour market takes place rather quick (Gijsberts & Lubbers, 2015). At the same, time sequence analyses on Dutch register data also show that unemployment is an important factor in return migration (Kleinepier, de Valk, & Van Gaalen, 2015). Of the Polish migrants that that stay in the Netherlands, however, over 70% are able to find a new job within a year (Gijsberts & Lubbers, 2015). Furthermore, Gijsberts and Lubbers (2015) find that being in employment is positively affected by a higher frequency of contact with the own ethnic group. The majority of Poles state that co-ethnics are an important resource for finding a job (Gijsberts & Lubbers, 2013, 2015).

The literature on Polish migrants in the Netherlands also addresses various demographic characteristics of this immigrant group. Recent Polish migrants are generally rather young. The majority of the Polish migrants are between the ages 18 and 35 (Gijsberts & Lubbers, 2013). Survey data shows that 87% of all Polish migrants are either in a stable relationship, or are married (Gijsberts & Lubbers, 2015). The large majority of Polish migrants are in relationships with co-ethnics (Kleinepier et al., 2015) and around half of the Polish migrants in the Netherlands have children living in the household (Gijsberts & Lubbers, 2015). Those who have children are less likely to return to Poland than those without children in the household (Kleinepier et al., 2015). Cohabitation and marriage are equally common amongst Polish migrants in the Netherlands, which is remarkable given that cohabitation is a relatively uncommon living arrangement in Poland (Kleinepier et al., 2015).

Furthermore, the Polish migrants tend to be higher educated than other immigrant groups in the Netherlands (Gijsberts & Lubbers, 2013). Various survey data show that half of the Polish migrants has achieved secondary education, and about one-fifth has achieved tertiary education (Weltevrede, De Boom, Rezai, Zuijderwijk, & Engbersen, 2009; Engbersen et al., 2011; Gijsberts & Lubbers, 2013).

Lastly, Polish migrants in the Netherlands have relatively low proficiency in the Dutch language (Gijsberts & Lubbers, 2013). There are indications, however, that language proficiency among Polish migrants increases in the years following the event of migration, particularly when contact with natives is high (Gijsberts & Lubbers, 2015). The majority of Polish migrants makes efforts to learn the Dutch language in the years following migration (Gijsberts & Lubbers, 2015).

# 3 Theoretical framework

Despite the growing body of literature on the Polish migrant group in the Netherlands, no research has yet addressed the extent to which Polish migrants live in ethnically segregated neighbourhoods. Consequently, it is unclear to what extent ethnic neighbourhood segregation affects the incomes of Polish immigrants.

Moreover, there seems to be no satisfying empirical answer to whether ethnic neighbourhood segregation affect immigrants' labour market position in general. Since Wilson's *The truly disadvantaged* (1987), the amount of research analysing neighbourhood effects has grown rapidly (Van Ham et al., 2011). In the United States, relatively strong negative neighbourhood effects on individual socio-economic outcomes have been found (Jencks & Mayer, 1990; Sampson, Morenoff, & Gannon-Rowley, 2002). Quantitative European studies, however, often do not find these strong neighbourhood effects (Musterd, 2003; Musterd et al., 2008; Van Ham et al., 2011; E. K. Andersson & Malmberg, 2015). Qualitative researchers, on the other hand, do find indications for the presence of neighbourhood effects, also in a European setting (Pinkster, 2007; Atkinson & Kintrea, 2001). The discrepancy between the available qualitative and quantitative evidence is puzzling, particularly given the strong theoretical foundations of neighbourhood effects.

According to the existing literature neighbourhood segregation is thought to affect the incomes of migrants in various ways and these may also apply to Polish migrants in the Netherlands. Three mechanisms that link ethnic neighbourhood segregation to the individual incomes of Polish migrants are mechanisms of human capital externalities, the ethnic economy mechanism and the linguistic concentration mechanism.

The first mechanism deals with human capital externalities. Borjas (1992, 1995) argues that the effect of ethnic neighbourhood segregation on immigrants' economic outcomes depends on the stock of human capital within the ethnic community. If the ethnic enclave is highly skilled, disadvantaged recent immigrants could benefit from living in an ethnic segregated neighbourhood. When the level of human capital within the ethnic community is low, on the other hand, the opposite effect may be true (Borjas, 1995; Edin, Fredriksson, & Åslund, 2003). It has been argued, however, that a migrant's human capital obtained in the country of origin is less valued than human capital acquired in the host country, a notion which also finds empirical support in the Dutch case (Kanas & van Tubergen, 2014). Also for Polish migrants, it can be argued that the economic returns to human capital obtained in Poland are relatively low after migrating to the Netherlands, as Polish migrants generally work in low prestige and low wage jobs, despite their relatively high levels of education (Dagevos, 2011; Engbersen et al., 2011; Gijsberts & Lubbers, 2015). Due to the low returns of education for Polish migrants, the relationship between the labour market position and educational attainment of Polish migrants is arguably fairly weak within the Dutch context. Indeed, this has also shown to be the case in

the case of Polish migrants in the United Kingdom (Drinkwater et al., 2009). There is thus limited opportunity for Polish migrants to experience the economic benefits from their co-ethnics' human capital. Human capital externalities are therefore likely to establish a negative relationship between ethnic neighbourhood segregation and the income of Polish migrants in the Netherlands.

Secondly, strong ethnic neighbourhood concentrations could be associated to the formation of 'ethnic enclave economies' (Galster, Metzger, & Waite, 1999), a phenomenon which refers to firms located in geographically bounded areas with high concentrations of other co-ethnic enterprises (K. L. Wilson & Portes, 1980). Given the occupational concentration of Polish migrants in the certain sectors such as industry, construction and agriculture (Dagevos, 2011; Engbersen et al., 2011; Van den Brakel et al., 2014), it is likely that such ethnic economies exist. Qualitative research has shown that a living in an ethnically segregated neighbourhood is associated with higher employment rates of immigrants in ethnic enclave economies through social network mechanisms, as informal job networks among immigrants within ethnic enclave economies are often shaped along ethnic lines (Pinkster, 2007) Employment resulting from these immigrant informal job networks, however, generally consists of low-wage jobs which provide little opportunities for upward social mobility (Sanders & Nee, 1987; Pinkster, 2007). In line with these findings, quantitative research shows migrants who live in ethnically segregated neighbourhoods have less interethnic ties (Martinovic, Van Tubergen, & Maas, 2009), and that immigrants who have more native contacts have a higher occupational status and higher incomes than immigrants with less contacts with natives (Kanas, Chiswick, Lippe, & Tubergen, 2012). In this way, living in an ethnically concentrated neighbourhood may hamper the immigrant's access to the generally more lucrative jobs of the mainstream economy (Sanders & Nee, 1987). These observations lead to the hypothesis that higher degrees of ethnic segregation of Polish migrants negatively affect the incomes of Polish migrants.

The last mechanism relates to the formation of linguistic enclaves, which are linked to ethnic enclaves by two pathways. First, migrants who live in ethnically concentrated areas have more opportunities to speak their mother tongue, as they have the opportunity to meet co-ethnics more frequently than those migrants living in areas with less co-ethnics (Stevens, 1992; Chiswick & Miller, 1996). This reduces the amount of Dutch used in daily life and therefore negatively influences the immigrant's Dutch language skills. Secondly, the incentives for learning the host country's language become relatively lower as the size of an ethnic concentration increases, in particular when migrants work in an ethnic enclave economy where they can rely on their native language (Stevens, 1992; Bauer, Epstein, & Gang, 2005). Living in an ethnically segregated area is therefore thought to be associated with lower degrees of proficiency in the host language (Chiswick & Miller, 2002) Indeed, a larger immigrant group size has been found to relate to lower language proficiency of the migrant group (Van Tubergen & Kalmijn, 2005, 2009). Chiswick and Miller (1995, 2002) argue that limited destination language proficiency reduces the earnings potential of immigrants, as it decreases job search efficiency, access to jobs and may lower

productivity. Living in areas with higher degrees of ethnic neighbourhood segregation is therefore hypothesized to negatively affect the incomes of Polish migrants.

# 4 Measuring segregation in neighbourhoods

In order to test the proposed relationship between segregation and its effects an accurate measurement of the neighbourhood first is needed. Andersson and Malmberg (2015) argue that the lack of empirical support for neighbourhood effects is attributable to the lack of a rigid methodological approach for measuring neighbourhood segregation. They propose a method called individualised neighbourhoods, where neighbourhoods are defined as a buffer around an individual in a circular fashion based on his or her $k$ nearest neighbours. The parameter $k$ can take on different values, allowing the neighbourhoods to vary by scale.

Defining neighbourhoods as buffers around the individual has three major advantages over conventional area-based neighbourhood measurements, which are usually defined as administrative areas. First, area-based measurements of segregation are plagued by the modifiable areal unit problem MAUP, which refers to the phenomenon that aggregated segregation measures are affected by the composition of a geographical unit and by how its boundaries are drawn (Openshaw, 1984; Wong, Lasus, & Falk, 1999). Wong (1993) has shown that levels of segregation tend to decline as the scale of the areal unit increases. Consequently, comparing levels of segregation between neighbourhoods which differ in terms of geographical size and population composition can be misleading. The issue of MAUP can be circumvented when neighbourhoods are defined as a buffer constructed based on the $k$ nearest neighbours around the individual, as the neighbourhoods are then defined equally for all individuals.

A second drawback of conventional neighbourhood measurements is that such measurements refer to an aggregate measure of an abstract spatial structure rather than to a phenomenon which affects individuals (Östh et al., 2014). As a result, geographically defined measures of spatial segregation are based on the assumption that individuals do not interact across the boundaries of the areal units in which they reside. Neighbourhood effects are not bounded by geographically defined areas, as individuals' social lives extend well beyond boundaries of administrative neighbourhoods. Defining neighbourhoods as buffers around individuals based on their nearest neighbours is therefore a more theoretically sound approach of measuring neighbourhood segregation and its possible effects on individuals.

A final advantage of the measurement of individualised neighbourhoods over conventional neighbourhood measurements is that it allows for a more detailed analysis of the role of scale in neighbourhood research, an aspect which is frequently omitted in the empirical literature (Manley & Van Ham, 2012; Östh, Clark, & Malmberg, 2015). Patterns of segregation can strongly differ across various geographical scales (Lee et al., 2008). Differences in micro- and macro-

levels of segregation are not captured by conventional administrative neighbourhood measurements, as the role of scale is not accounted for. Lee and colleagues (2008) argue that there is not one 'right' scale size to measure neighbourhood segregation, and that different scales of neighbourhood measurements should explicitly be taken into account.

When accounting for the role of scale, it is important to acknowledge that neighbourhood segregation and neighbourhood effects may operate differently at different scales (Van Ham et al., 2011). Neighbourhood segregation and neighbourhood effects should therefore be approached multi-scalar phenomena. Different mechanisms of neighbourhood effects may operate at different scales (Östh et al., 2015). When testing hypotheses regarding neighbourhood effects, it therefore important to specify the neighbourhood measurements at the correct scale. Certain neighbourhood effects may not be identified if they are measured at the wrong scale (Manley & Van Ham, 2012).

It should be noted, however, that the method of individualised neighbourhoods does have its drawbacks. The method cannot take natural borders such as rivers and roads into account. For this reason, the neighbourhood an individual perceives to live in may differ from the neighbourhood computed with the individualised neighbourhoods method. As calculation of the individualised neighbourhoods are based on the nearest neighbours, this consequently means that the geographical sizes of individualised neighbourhoods may differ strongly between people living in rural and in urban areas. Lastly, the choice for the population size of the individualised neighbourhoods is relatively arbitrary. As noted, there is no one 'right' size for the measurement of neighbourhoods (Lee et al., 2008).

# 5 Conceptual model

The conceptual model which follows from the previous paragraphs is depicted in Figure 1, which outlines the expected relationships between the introduced concepts. First, the use of individualised, scalable neighbourhoods is thought to improve measurements of ethnic neighbourhood segregation compared to administrative neighbourhoods, by circumventing the MAUP, being theoretically sound and providing insight in the role of scale. Second, the hypothesized negative relationship between ethnic neighbourhood segregation of Polish migrants and the incomes of these migrants is thought to be strengthened by the use of the individualised neighbourhoods approach compared to administrative neighbourhoods.

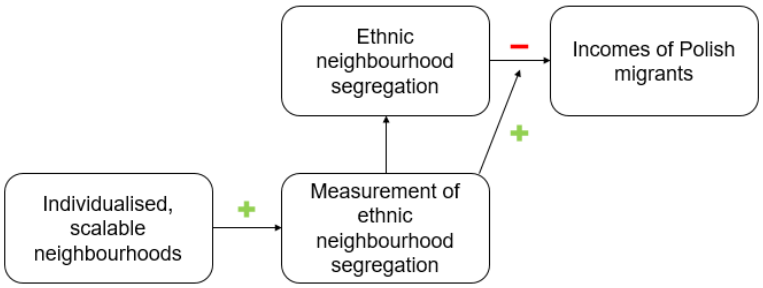


Figure 1: Conceptual model

# 6 Data and methods

## 6.1 Data

To analyse the research questions and the hypothesis register-based data from Statistics Netherlands is used. These data are constructed by using information from the System of social statistical datasets (SSD). Bakker, Van Rooijen and Van Toor (2014) refer to the SSD as a "a system of linked statistical registers and surveys which cover a broad range of demographic and socio-economic subjects" (p. 412). The data is constructed by combining data from various administrative registers into one dataset (Bakker et al., 2014). These data are linked on the basis of the citizen service number, which is the unique personal identifier of all Dutch citizens. After linking the various datasets, the SSD undergoes a procedure of anonymization. Data such as the SSD, which are based on population registers, are widely recognised as high quality data, as they overcome problems such as non-response, recall errors or small sample sizes. The data refers to the situation in the Netherlands at the 31st of December, 2012.

The SSD contains geocoded information of the households of all registered individuals, which is used to calculate neighbourhoods based on an individual's $k$ nearest neighbours. This computation is done by using EquiPop, a software package specifically designed to calculate individualised neighbourhoods (Östh, 2014). As the dataset is very large, calculating these individualised neighbourhoods is computationally demanding. Therefore, the geocoded data is transformed to grids of 100 by 100 meters, a computational technique which is in line with Östh, Malmberg and Andersson (2014) and Östh, Clark and Malmberg (2015). Individualised neighbourhoods are then calculated by counting the amount of individuals within a grid, and subsequently adding neighbours from adjacent grid cells in a circular fashion until the specified level of $k$ nearest neighbours has been reached (Östh et al., 2014). The programme then provides aggregate statistics for these grids. EquiPop calculates the proportion of people who are Polish immigrants, defined as those individuals who were born in Poland, within the individualised neighbourhood. EquiPop also provides a variable indicating the distance needed to reach the specified amount of nearest

neighbours for every grid. Ethnic neighbourhood segregation is measured on the on eleven different scales by calculating the percentage of Polish immigrants within every $k$-level. These scales are 50, 100, 200, 400, 800, 1,600, 3,200, 6,400, 12,800, 25,600 and 51,200 nearest neighbours. The dataset is restricted to the working ages of 18 until 64.

## 6.2 Method

The analysis consists of three parts, which correspond to the three objectives of the study. In the first part of the analyses, descriptive statistics on the degree to which Polish migrants live in segregated residential areas are presented, which allow for the identification of the extent to which Polish migrants live in segregated areas. To explore the location of these areas in the Netherlands, a geographic analysis is conducted using Geographical Information Systems (GIS). Presenting the degree ethnic segregation on various $k$-levels for three selected municipalities allows for the distinction between micro- and macro-level patterns of segregation.

In the second part of the analyses, the individualised neighbourhood approach is compared to the conventional measurement of administrative neighbourhoods by means of the isolation index. The isolation index measures the probability that a minority member, in this case Polish immigrants, meets another minority member if contacts are picked randomly within a neighbourhood (Östh et al., 2014). The intuitive interpretation of the isolation index makes this measure a popular choice for measuring neighbourhood segregation. The isolation index is calculated for Polish migrants on the eleven selected $k$-levels, which are then juxtaposed to the isolation index based on administrative neighbourhoods available in the SSD. This analysis is carried out for three selected municipalities. The isolation index calculated based on administrative neighbourhoods is calculated as:

$$\sum_{i=1}^{n} \left[ \left( \frac{x_i}{X} \right) \left( \frac{x_i}{t_i} \right) \right] \tag{1}$$

where $x_i$ is the minority population of area $i$, $X$ is the total minority population, and $t_i$ is the total population of area $i$ (Iceland & Weinberg, 2002).

Calculating the isolation index on the basis of the individualised neighbourhoods method requires a small modification of the original formula:

$$\frac{\sum_{i=1}^{n} \left( x_i * \frac{x_{i,k}}{k} \right)}{\sum_{i=1}^{n} x_i} \tag{2}$$

where $k$ is the amount of nearest neighbours included in the individualised neighbourhood (Östh et al., 2015).

In the last part of the analysis the relationship between ethnic neighbourhood segregation and the incomes of Polish migrants is tested by estimating eleven different fixed effects regression models on all eleven specified $k$-levels. This analysis is carried out on the national level. The eleven fixed effects regression models are identical apart for the value of the specified number of $k$ nearest neighbours included in the individualised neighbourhoods, effectively testing the hypothesized relationship on eleven different scales. In the fixed effects models a dummy for every administrative neighbourhood in which at least one Polish individual lives is included in the model (6,454 dummies in total). Hereby, the model accounts for the variation between administrative neighbourhoods, reducing the risk of omitted variable bias on the neighbourhood level.

To test the hypothesized negative relationship between ethnic segregation and the incomes of Polish migrants, the proportion of Polish migrants within the individualised neighbourhoods as calculated by EquiPop is included as the main independent variable in the fixed effects regression model. Due to the inclusion of the 6,454 dummies for every administrative neighbourhood, the estimated coefficient for the proportion of Polish migrants within $k$ effectively shows the added effect of ethnic segregation on income as calculated by the individualised neighbourhoods method on top of the administrative neighbourhoods. To correct for the spatial autocorrelation between individuals living in similar residential contexts, clustered standard errors based on the neighbourhood level are applied.

The fixed effects regression analysis is restricted to first generation Polish migrants in the working ages and who participate in the labour market, meaning that they are between the ages of 18 and 64 and have a job or are looking for a job at the time of measurement. These are 62,356 individuals in total. The amount of missing data for these cases is minimal. For 120 Polish migrants no geocoded data is available. 39 Polish migrants have missing values on one of the variables in the model (19 missing values on income and 20 missing values for length of stay in the Netherlands), leaving the fixed effects regression analysis with a total of 62,197 cases in the final model.

## 6.3 Dependent variable of the fixed effects regression analysis

The dependent variable is measured as the absolute income of individuals. In the SSD this is measured as an individual's so-called 'personal' income, which is defined as the yearly gross income from work, a business and government welfare payments from income insurances and social insurances (not including childcare benefits). The measurement does not include a household's income sources or subsidies which cannot directly be traced back to individual members of a household, therefore only measuring individual income. In research in which income is the dependent variable, it is a standard procedure to perform a natural logarithmic transformation to the dependent variable to account for the positive skewness of the income distribution and the violation of the homoscedasticity-assumption in regression analyses (Mincer, 1974; Chiswick &

Miller, 2002; Kanas et al., 2012; Drinkwater et al., 2009). In this thesis, however, the dependent variable remains untransformed due to the fact that the used income measure can take on negative values (which are values usually present among people who are self-employed who own a business which records losses), making a natural logarithmic transformation impossible. The possibility of adding a constant value which is slightly higher than the minimum observed value in the income distribution to the dependent variable has been explored, but has been found to distort the data. To account for the distributional issues of the dependent variable and its consequences, a sensitivity analysis which uses a relative rather than an absolute specification of the dependent variable is carried out to test the robustness of the model.

## 6.4 Independent variables of the fixed effects regression analysis

The main independent variable of interest measuring the concept of ethnic segregation which is used to test the hypotheses is measured as follows.

*Percentage of Polish migrants within k*: This variable is calculated by EquiPop, indicating the percentage of Polish migrants within each individualised neighbourhood at the specific $k$-level. In the fixed effects regression, the parameter for this variable must be interpreted as the extra effect of using the individualised neighbourhood method on top of using conventional administrative neighbourhoods, as the administrative neighbourhoods are already included as dummies in the model. To ease the interpretation of the coefficient, the variable is transformed in such a way that the coefficient gives the estimated difference in income between the minimum and maximum observed values of ethnic segregation.

The model also contains the following set of individual level control variables.

*Distance*: The distance variable represents the Euclidian distance between the grid in which the individual's household is located and the grid furthest removed from the individual's grid needed to reach the specified $k$-level as calculated by EquiPop (Östh, 2014). This variable accounts for the variation in the distance needed to reach the $k$-level due to differences in the population densities of residential areas.

*Age*: This variable indicates the age of the individual on the 31st of December 2012. To account for the non-linearity of age-income profiles a quadratic specification is included in the model (Mincer, 1974).

*Sex*: The variable sex is included in the model as a dummy variable, taking a value of 1 if the individual is a man, and value of 2 if the individual is a woman.

*Length of stay*: The length of stay of the Polish migrant in the Netherlands is split in four categories. The first category includes migrants who arrived in the Netherlands 8 or more years ago (corresponding to the time before Poland entered the European Union), the second category includes those who arrived between 5 and 8 years before the moment of observation, the third comprises those who arrived between 1 and 4 years ago and the last category arrived to the Netherlands less than one year before the moment of observation.

*Household characteristics*: The variable indicating the household characteristics is split into nine categories: living in a household without a partner (either marriage or cohabitation) and without children, living in a household with a partner and with one child, living in a household with a partner and with two children, living in a household with a partner and with three or more children, living in a household with a partner and no children, living in a household without a partner and with one child, living in a household without a partner and with two children, living in a household without a partner and with three or more children and one rest-category, which is a small category made up out of children who live in their parents' household (52.41%), a person living in a household who is not the partner, the parent or a child of someone else in the household (39%) or a person living in an institutional household (8.59%).

*Self-employed*: This dummy variable takes a value of 1 if the Polish individual is self-employed, and 0 if this is not the case.

# 7 Results

## 7.1 Descriptive analysis

To illustrate the extent to which Polish migrants live in segregated areas, Table 1 shows the proportion of Polish migrants in the individualised neighbourhoods for selected percentile ranks on various $k$-levels. For a $k$-level of 50, it can be seen that half of the Polish migrants in the Netherlands live in an area with 3.8% co-ethnics or less. This means that half of the Polish migrants roughly have two or less Polish-born neighbours in an area of 50 nearest neighbours. For $k = 200$, the value of 0.09 for the 90th decile implies that for ten percent of the Polish migrants in the Netherlands at least 18 out of the 200 nearest neighbours are of Polish descent. One percent of the Polish migrants between ages 18 and 64 lives in small segregated areas which almost only consist of Polish-born. The 99th percentile has a value of 0.957 on $k = 50$, implying that out of the 50 neighbours, 48 are Polish-born. As the $k$-level increases the highest observed value naturally decreases. The highest observed value for $k = 51{,}200$ is 0.048, meaning that the highest observed number of Polish-born within an area of 51,200 nearest neighbours is 2,458. These descriptive results show that, although there are some areas where Polish migrants live in highly ethnically segregated neighbourhoods, the larger share of the Polish migrants does not live in such areas.

To explore where these highly segregated areas are, the ratios of Polish migrants on a $k$-level of 25,600 are shown on a map of the Netherlands, which is presented in Figure 2. The specific $k$-level of 25,600 is chosen because it represents the size of a medium sized city or a metropolitan area, which allows for the identification of general residential patterns of Polish migrants on a macro-level of segregation. The colours indicate the proportions of Polish migrants within the buffer of $k$ nearest neighbours in quintiles.

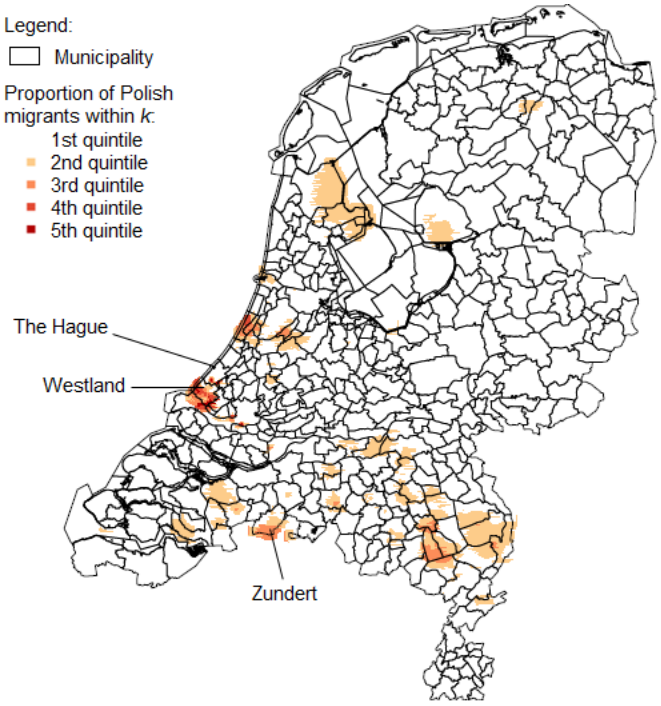


Figure 2: Percentages of Polish migrants on $k$ = 25,600

Table 1: Observed proportion of Polish-born nearest neighbours for selected $k$-levels for various percentiles across the Polish population aged 18-64 ($N$ = 62,236)

| | $k = 50$ | $k = 200$ | $k = 800$ | $k = 3,200$ | $k = 12,800$ | $k = 52,200$ |
|---|---|---|---|---|---|---|
| Minimum value | 0.002 | 0.002 | 0.000 | 0.000 | 0.000 | 0.001 |
| 10th percentile | 0.013 | 0.004 | 0.003 | 0.003 | 0.004 | 0.004 |
| 50th percentile | 0.038 | 0.021 | 0.014 | 0.011 | 0.010 | 0.009 |
| 75th percentile | 0.076 | 0.045 | 0.030 | 0.023 | 0.019 | 0.017 |
| 90th percentile | 0.143 | 0.090 | 0.060 | 0.045 | 0.036 | 0.029 |
| 95th percentile | 0.233 | 0.132 | 0.090 | 0.063 | 0.050 | 0.036 |
| 99th percentile | 0.957 | 0.488 | 0.228 | 0.103 | 0.075 | 0.041 |
| Maximum value | 1 | 0.996 | 0.365 | 0.154 | 0.089 | 0.048 |

Source: own calculations.

Figure 2 shows that the largest Polish concentrations are found in the western part of the country, near to the city of The Hague in the municipality Westland. Smaller concentrations are also found in the municipality of Zundert, close to the Belgian border, and close to Eindhoven in the province of North Brabant (the municipalities of Asten and Someren in particular). To explore segregation patterns within particular municipalities, the following section focusses on three municipalities with high concentrations of Polish migrants which have been identified in Figure 2. These municipalities are Westland, Zundert and The Hague. The municipality of Westland is chosen as it is the municipality with the highest percentage of Polish migrants in the Netherlands on the 31st of December 2012. Zundert is included to also explore segregation patterns in smaller municipalities with strong concentrations of Polish migrants. The motivation to include The Hague in the in-depth analysis is that this is the municipality with highest absolute number of Polish migrants in the Netherlands.

It should be noted that the statistics presented in the Figures 3, 4 and 5 are based on national calculations.

Figure 3 presents the municipality of Westland. Out of the 63,509 total inhabitants aged between 18 and 64, 2,910 are Polish migrants, making up 4.6% of the total population of Westland. For $k = 50$, very strong concentrations of Polish migrants can be found in Westland, as shown in Figure 3. There are numerous grids in which 80 to 100% of the 50 nearest neighbours are Polish migrants. As the $k$-level gradually increases to 400, patterns of ethnic segregation become less pronounced, as some of the very strong concentrations disappear when the value of $k$ increases. This shows that patterns of segregation in the municipality of Westland are particularly pronounced on smaller scales, and become gradually less pronounced when higher scale levels are introduced.

The same analysis is presented for the municipality of Zundert in Figure 4, which is a much smaller municipality. The 415 Polish migrants living in this


Figure 3: Proportion of Polish migrants in an individualised neighbourhood on various $k$-levels in the municipality of Westland

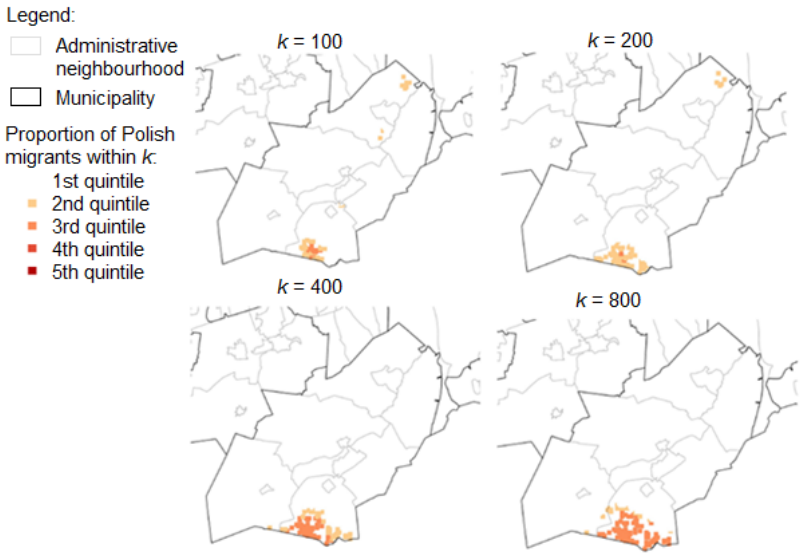


Figure 4: Proportion of Polish migrants in an individualised neighbourhood on various $k$-levels in the municipality of Zundert

municipality make up 3.1% of the total population in the age range of 18 until 64. In Zundert, some fairly strong concentrations of Polish migrants can be found. For $k = 100$, there are concentrations of 40 to 60 Polish migrants in the southern part of the municipality. As the $k$-level increases, the weaker northern concentrations disappear, while the area of the concentrations in the southern part of the municipality increase in size. This is likely to be related to the low population density in this municipality. The size of the ethnic concentration naturally increases as the number of households within the grids is low. This outcome indicates that there is geographical variation in the role of scale in neighbourhood segregation between municipalities. It also implies that the use of higher $k$-levels may be less informative for rural areas with a low population density such as Zundert.

In the municipality of The Hague, another relationship between scale and ethnic segregation emerges. The Hague is the municipality with the highest absolute number of Polish migrants (6,575) in the Netherlands. The Polish residents in The Hague account for 8.7% of the total Polish population in the Netherlands aged between 18 and 64. Figure 5 shows that on the smallest $k$-levels presented in the figure ($k = 400$), virtually no segregation of Polish migrants is visible. As the $k$-level gradually increases, however, the patterns of segregation strongly change. At the highest $k$-level presented in the figure ($k = 3{,}200$), relatively high levels of segregation are found in the administrative neighbourhood of Laakkwartier in the eastern part of the municipality. Levels of

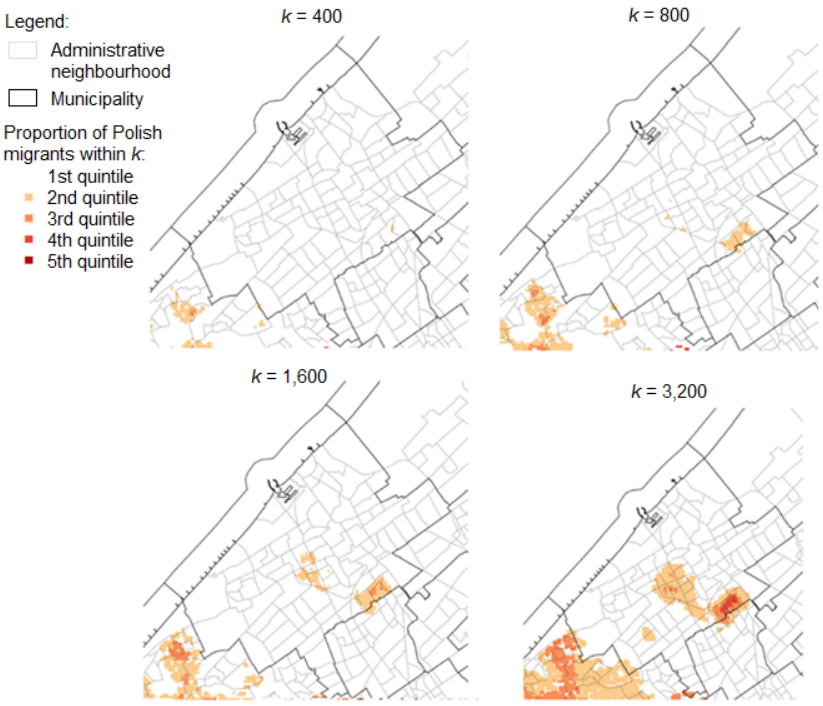


Figure 5: Proportion of Polish migrants in an individualised neighbourhood on various $k$-levels in the municipality of The Hague

ethnic segregation become more pronounced as the size of the buffer constructed around the grids increases.

In sum, the results from the GIS analysis indicate that the role of scale is an important factor to take into account when exploring patterns of neighbourhood segregation, as the observed degree of segregation can change when a different scale-level is specified. It is also important to note that there appears to be geographical variation in the scale at which neighbourhood segregation is most pronounced. In the municipality of Westland, segregation takes place on the smallest level identified in the analysis, whereas in the municipality of The Hague patterns of segregation appear not to be pronounced on smaller scales, but do take place on larger scales. These results may relate to differences in the population density of the two municipalities. Lastly, the results indicate that for smaller, less densely populated municipalities such as Zundert, higher $k$-levels may be less informative than in more densely populated municipalities.

## 7.2 Comparing isolation indices on different $k$-levels

In the following section the isolation index based on the individualised neighbourhoods measurement is compared to the same index calculated by means

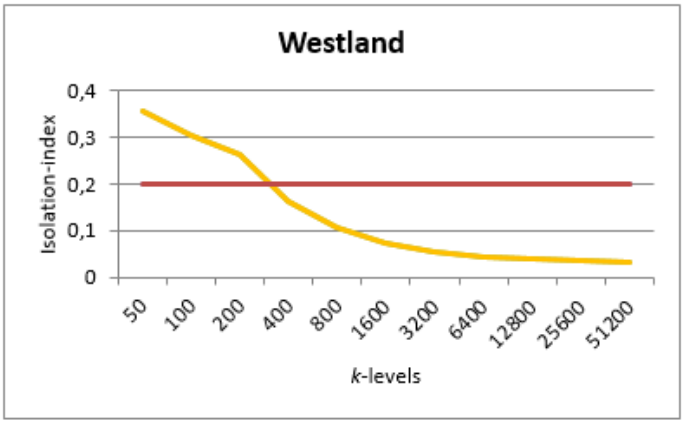


Figure 6: Isolation index calculated for different $k$-levels (yellow line) and based on administrative neighbourhoods (red line) for the municipality of Westland

of administrative neighbourhoods. Comparing the isolation index calculated by using both neighbourhood measurements allows for the analysis of how the measurement of segregation may depend on the methodology used for measuring neighbourhoods. It also allows for the direct comparison of how individualised neighbourhood approach relates to the conventional administrative neighbourhoods. The analysis is carried out for the three selected municipalities of Westland, Zundert and The Hague.

The results of the comparison are presented in Figures 6, 7 and 8. The yellow line depicts the isolation-index based on the $k$ nearest neighbours, whereas the red line represents the same measurement calculated on the basis of administrative neighbourhoods. The yellow lines indicate that the chance that a random contact between Polish individuals is made decreases for higher values of $k$, providing support for the idea that segregation is a scale-dependent phenomenon. The point where the two lines intersect indicate the effective population size of the administrative neighbourhoods in the given municipality, and thus the level of segregation which is most strongly captured when using conventional area-based segregation measures. The figures show that segregation measurements which are based on administrative neighbourhood measurements do not reveal the level of ethnic segregation on scales lower than the population size of administrative neighbourhoods. When measuring neighbourhoods by using administrative neighbourhoods, one would find an isolation index of 0.20 for the municipality of Westland, whereas when neighbourhoods are defined by the 50 nearest neighbours, the isolation index has a substantial higher value of 0.35. These results are in line with the results found by Östh, Malmberg and Andersson (2014) for visible minorities in Sweden.

By circumventing the MAUP, the method of individualised neighbourhoods also allows for a meaningful comparison of levels of segregation over different $k$-levels. This is done for the municipalities of Westland, Zundert and The Hague,

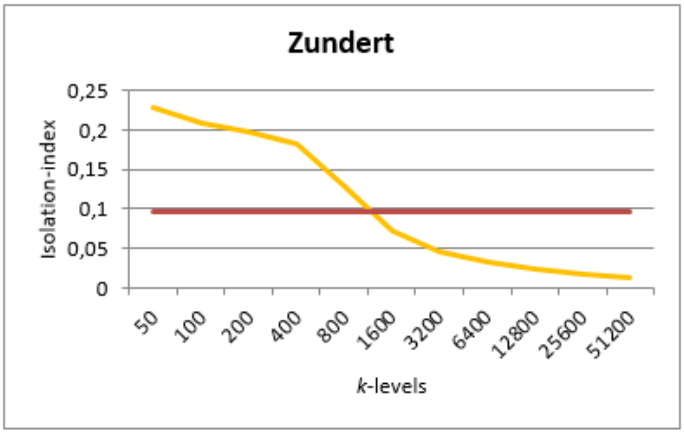


Figure 7: Isolation index calculated for different $k$-levels (yellow line) and based on administrative neighbourhoods (red line) for the municipality of Zundert


Figure 8: Isolation index calculated for different $k$-levels (yellow line) and based on administrative neighbourhoods (red line) for the municipality of The Hague

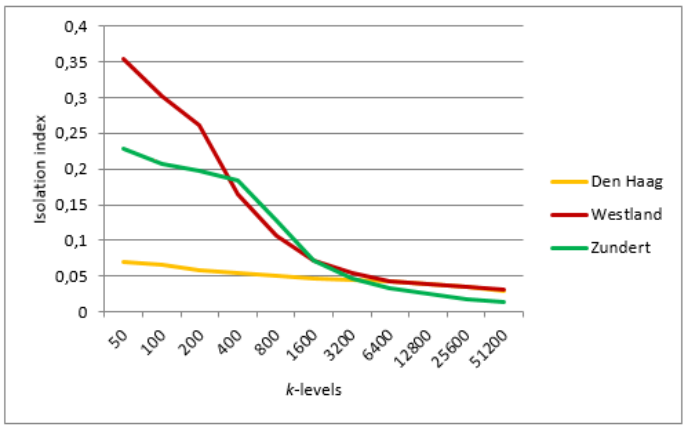


Figure 9: Comparing isolation indices on different $k$-levels for the municipalities of Westland, Zundert and The Hague

as presented in Figure 9. The results indicate that ethnic segregation of Polish migrants is strongest in Westland for $k = 50$. The values of the isolation index of Westland and Zundert converge at $k = 400$, implying that levels of segregation are similar at this scale. When comparing the isolation indices based on the geographically defined neighbourhoods, one would overlook this convergence and simply conclude that in the municipality of Westland, Polish migrants live in more segregated neighbourhoods than in Zundert.

The analysis shows that different municipalities have different effective neighbourhood population sizes. To compare neighbourhoods in a meaningful way, neighbourhoods should be calculated on the basis of the same population size, which can be achieved by using the individualised neighbourhoods methodology. As we have seen in the previous part of the analysis, however, municipalities differ in terms of the scale at which segregation is most pronounced. When taking both these notions into account, the choice of the correct scale for meaningfully comparing two different areas is not straightforward.

## 7.3 Fixed effect regression analysis

To explore whether there is a statistical relationship between income and the level of ethnic segregation of Polish-born migrants within a neighbourhood consisting of $k$ nearest neighbours on the national level, the correlations between income and ethnic segregation are presented in Table 2. From the analysis, there appears to be a moderately negative correlation between income and higher levels of ethnic segregation, meaning that incomes of Polish migrants who live in more ethnically segregated areas seem to be lower than the incomes of Polish migrants living in areas with lower degrees of ethnic segregation. The strength of this negative correlation is rather stable over various $k$-levels. Naturally, the degree of segregation strongly correlates with segregation measures on scales close

to a particular $k$-level. This correlation decreases in strength as the $k$-level is further removed from the particular scale. The correlation between ethnic segregation on $k = 50$ and $k = 100$, for example, is 0.960, whereas the correlation between the scales of $k = 50$ and $k = 12{,}800$ is 0.302.

Table 2: Bivariate correlations between income and the proportion of Polish migrants within the individualised neighbourhoods on 11 scales on the national level ($N = 62{,}197$)

| | *Income* | k = 50 | k = 100 | k = 200 | k = 400 | k = 800 | k = 1,600 | k = 3,200 | k = 6,400 | k = 12,800 | k = 25,600 | k = 51,200 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Income* | 1 | | | | | | | | | | | |
| *k = 50* | -0.090 | 1 | | | | | | | | | | |
| *k = 100* | -0.086 | 0.960 | 1 | | | | | | | | | |
| *k = 200* | -0.080 | 0.901 | 0.962 | 1 | | | | | | | | |
| *k = 400* | -0.089 | 0.863 | 0.916 | 0.951 | 1 | | | | | | | |
| *k = 800* | -0.096 | 0.782 | 0.833 | 0.868 | 0.954 | 1 | | | | | | |
| *k = 1,600* | -0.100 | 0.661 | 0.703 | 0.732 | 0.839 | 0.937 | 1 | | | | | |
| *k = 3,200* | -0.096 | 0.502 | 0.536 | 0.559 | 0.676 | 0.807 | 0.927 | 1 | | | | |
| *k = 6,400* | -0.090 | 0.384 | 0.414 | 0.435 | 0.558 | 0.703 | 0.851 | 0.955 | 1 | | | |
| *k = 12,800* | -0.086 | 0.302 | 0.325 | 0.340 | 0.464 | 0.615 | 0.780 | 0.894 | 0.961 | 1 | | |
| *k = 25,600* | -0.085 | 0.268 | 0.289 | 0.300 | 0.416 | 0.563 | 0.725 | 0.837 | 0.911 | 0.969 | 1 | |
| *k = 51,200* | -0.085 | 0.261 | 0.278 | 0.288 | 0.394 | 0.523 | 0.663 | 0.759 | 0.825 | 0.890 | 0.954 | 1 |

Source: own calculations.

To assure the moderately negative correlation between ethnic segregation and income is no spurious relationship which is influenced by other confounding variables, the fixed effects regression includes various individual level variables. Descriptive statistics for these individual level variables for Polish-born immigrants, who are participating in the labour market and are within the age range of 18-64, are presented in Table 3. These descriptive statistics are calculated on the national level. The table also shows the median income for different values of the independent variables.

The descriptive statistics indicate that there are differences between the incomes of Polish migrants for different demographic characteristics. The majority (58.66%) of Polish migrants are between the ages 25 and 40. This group also has the highest median annual income (21,606 euro), which is 5,679 euros higher than those in the youngest age category of 18-24 and 1,185 euros higher than those aged 41-64. The data also shows that there are slightly more female Polish

immigrants in the Netherlands than male immigrants (53.28% is female). Men, however, have a higher annual median income than women (median income = 22,805 euro for men and 18,558 euro for women). Furthermore, it appears that almost half (46.19%) of the Polish migrants have been living in the Netherlands for a period between one and four years and 19.59% arrived in the Netherlands less than a year before the 31st of December 2012. Those who have resided in the Netherlands for longer periods of time have a higher median annual income than those who have arrived more recently (22,532 euro for those have lived in the Netherlands for a period longer than 8 years, and 17,733 euro for those who arrived less than a year ago). In terms of household characteristics, it seems that the most common living arrangement for employed Polish immigrants is to live in a household without a partner and without any children (33.11%), who have a median annual income close to the that of the total Polish population (19,856 euro). Almost one in ten of the Polish migrants are self-employed, who do not seem to have different income levels than other Poles.

In Table 4 the results of the fixed effects regression model with clustered standard errors calculated on the neighbourhood level are presented for all $k$-levels. As mentioned, the coefficients of the variable measuring the proportion of Poles within the individualised neighbourhoods of $k$ nearest neighbours should be interpreted as the added effect of using the individualised neighbourhoods method on top of administrative neighbourhoods, due to the inclusion of a dummy-variable for all the administrative neighbourhoods. Figure 10 displays the coefficients of all eleven models in a line plot, including the confidence intervals based on the clustered standard errors.

The results indicate that for Polish migrants in the Netherlands, there is a negative relationship between ethnic neighbourhood segregation and income. The relationship between ethnic segregation as measured by the individualised neighbourhood method and income is strongest on lower $k$-levels. On $k = 400$, the estimated difference between the least and most segregated Polish migrants is 7,708.33 euros per year, while controlling for other individual level factors and neighbourhood variation.

The relationship is no longer significant for $k = 3{,}200$ and higher $k$-levels. The absence of a significant relationship for $k$-levels of 3,200 and higher is likely to be accounted for by the overlap between individualised neighbourhoods and administrative measures of neighbourhoods at higher$k$-levels, which are included in the model as fixed effects. From the model, the main added value of using the individualised neighbourhoods approach appears to be that the effects of ethnic segregation on scales smaller than the effective population size of administrative neighbourhoods can be taken into account.

The control variables, lastly, generally show the same results in all eleven models. First, migrants who have resided in the Netherlands for a shorter period of time have lower earnings than those who have arrived earlier, which is in line with earlier findings for Poles in the United Kingdom (Drinkwater et al., 2009). Secondly, the model implicates a non-linear, inverted U-shape relationship between age and income, which corresponds to the human capital earnings function (Mincer, 1974; Chiswick & Miller, 2002; Drinkwater et al.,

Table 3: Descriptive statistics for the individual level variables for Polish-born immigrants between ages 18-64 who are active in the labour market on the national level ($N$ = 62,337)

| | **N** | **%** | **Median annual income (in euros)** |
|---|---|---|---|
| Total | 62,337 | 100.00 | 20,422.00 |
| *Age* | | | |
| 18-24 | 9563 | 15.34 | 15,927.00 |
| 25-40 | 36,569 | 58.66 | 21,606.00 |
| 41-64 | 16,205 | 26 | 20,421.00 |
| | Mean age: 35.05 | | |
| Sex | | | |
| Male | 29,122 | 46.72 | 22,805.00 |
| Female | 33,215 | 53.28 | 18,558.00 |
| *Length of stay* | | | |
| longer than 8 years | 11,485 | 18.42 | 22,532.00 |
| between 5 and 8 years | 10,453 | 16.77 | 22,406.00 |
| between 1 and 4 years | 28,566 | 45.83 | 20,820.50 |
| less than 1 year | 11,813 | 18.95 | 17,733.00 |
| *Household characteristics* | | | |
| without a partner, no children | 20,642 | 33.11 | 19,856.00 |
| without a partner, 1 child | 2324 | 3.73 | 18,998.00 |
| without a partner, 2 children | 702 | 1.13 | 18,939.00 |
| without a partner, 3 or more children | 156 | 0.25 | 18,266.50 |
| with a partner, no children | 18,439 | 29.58 | 21,542.00 |
| with a partner, 1 child | 10,219 | 16.39 | 21,263.00 |
| with a partner, 2 children | 5,713 | 9.16 | 21,487.00 |
| with a partner, 3 or more children | 1,189 | 1.91 | 20,675.00 |
| other | 2,953 | 4.74 | 14,239.00 |
| *Self-employed* | | | |
| Yes | 5,692 | 90.87 | 20,702.50 |
| No | 56,645 | 9.13 | 20,405.00 |

Source: own calculations.

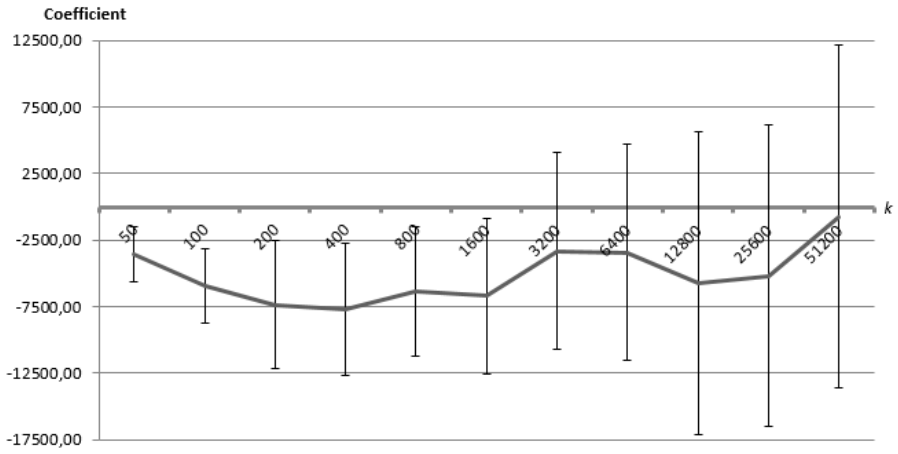


Figure 10: Estimated coefficients for the proportion of Polish migrants within the individualised neighbourhoods from the fixed effects model with 95% confidence interval based on clustered standard errors for 11 $k$-levels

2009). Third, the results show that Polish women earn significantly less than Polish men, a difference which has also been found among recent migrants from the Eastern European countries which accessed the European Union in 2004 in the United Kingdom (K. Clark & Drinkwater, 2008). Fourth, Polish migrants with a partner earn more than those without a partner, which is in line the relationship between higher labour market activity and being in a marital or cohabitational union found by Kleinepier and colleagues (2015). When a migrant with a partner lives with three children or more, however, the effect is opposite to when living with a partner and without children, which may be driven by lower labour market activity of Polish women after childbirth (Kleinepier et al., 2015). Compared to living without a partner and without children, living without a partner but with two or more children also negatively relates to the incomes of Polish migrants, indicating a relationship between single parenthood and lower incomes. Lastly, Polish migrants who are self-employed have significantly lower incomes than those who are not self-employed, which is notable given the absence of such a difference in the descriptive analysis.

## 7.4 Sensitivity analyses

The results have proven to be robust in a variety of sensitivity analyses. Removing outliers on the right- and left-hand side of the income distribution does not alter the interpretation and the results of the model, although coefficients do slightly change. When using household income rather than individual income as the dependent variable, the pattern and significance of the results remain the same. Also, transforming the dependent variable to a relative rather than an

Table 4: Fixed effects regression analyses predicting incomes of Polish migrants in the Netherlands using clustered standard errors ($N$ = 62,197)

| | $k$ = 50 | | $k$ = 100 | | $k$ = 200 | | $k$ = 400 | | $k$ = 800 | | $k$ = 1,600 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Coefficient | Cl. st. err. | Coefficient | Cl. st. err. | Coefficient | Cl. st. err. | Coefficient | Cl. st. err. | Coefficient | Cl. st. err. | Coefficient | Cl. st. err. |
| Proportion of Polish migrants within $k$ | -3567.63*** | 1059.51 | -5914.38*** | 1428.17 | -7684.35** | 2453.19 | -7708.83** | 2554.76 | -6360.41** | 2478.46 | -6713.53* | 2987.32 |
| Distance needed to reach $k$ | 2.10* | 0.94 | 2.30*** | 0.85 | 0.90 | 0.59 | 1.58** | 0.55 | 1.72* | 0.58 | 1.39* | 0.60 |
| Length of stay (more than 8 years = ref.) | | | | | | | | | | | | |
| between 5 and 8 years | -3955.70*** | 337.91 | -3951.11*** | 337.45 | -3956.17*** | 337.24 | -3954.35*** | 336.81 | -3962.16*** | 336.71 | -3971.03*** | 337.05 |
| between 1 and 4 years | -6390.26*** | 287.86 | -6387.67*** | 287.09 | -6396.69*** | 286.58 | -6401.59*** | 286.05 | -6409.36*** | 285.49 | -6418.16*** | 285.72 |
| less than 1 year | -10363.24*** | 294.38 | -10359.10*** | 293.72 | -10376.04*** | 293.40 | -10385.66*** | 292.54 | -10395.13*** | 291.95 | -10406.24*** | 292.55 |
| Age | 1426.33*** | 56.71 | 1425.50*** | 56.74 | 1427.92*** | 56.67 | 1426.53*** | 56.78 | 1427.29*** | 56.73 | 1426.43*** | 56.82 |
| Age squared | -17.47*** | 0.74 | -17.46*** | 0.74 | -17.48*** | 0.74 | -17.47*** | 0.74 | -17.47*** | 0.74 | -17.47*** | 0.74 |
| Sex (male = ref.) | -6472.44*** | 176.90 | -6472.54*** | 177.04 | -6469.93*** | 176.94 | -6467.89*** | 176.84 | -6467.07*** | 176.54 | -6464.98*** | 176.46 |
| Household characteristics (without a partner, no children = ref.) | | | | | | | | | | | | |
| with a partner, no children | 1172.67*** | 183.55 | 1171.52*** | 183.13 | 1171.05*** | 183.38 | 1179.53*** | 183.29 | 1184.85*** | 183.14 | 1185.17*** | 183.06 |
| with a partner, 1 child | -81.30 | 256.79 | -82.90 | 256.40 | -86.52 | 256.36 | -76.32 | 256.13 | -74.36 | 256.12 | -73.97 | 256.17 |
| with a partner, 2 children | -703.07 | 371.45 | -712.68 | 371.61 | -715.80 | 371.61 | -710.05 | 371.64 | -704.46 | 371.91 | -701.62 | 371.96 |
| with a partner, 3 or more children | -2142.69*** | 569.79 | -2143.15*** | 569.97 | -2152.91*** | 570.84 | -2141.98*** | 570.39 | -2153.22*** | 571.98 | -2150.36*** | 572.05 |
| without a partner, 1 child | -160.2 | 407.79 | -168.55 | 407.54 | -165.28 | 408.04 | -154.85 | 407.93 | -154.91 | 407.91 | -152.13 | 407.57 |
| without a partner, 2 children | -1860.49*** | 584.60 | -1866.59*** | 584.25 | -1865.08*** | 584.70 | -1854.50*** | 584.52 | -1847.75*** | 585.08 | -1840.51*** | 583.98 |
| without a partner, 3 or more children | -4665.47*** | 890.79 | -4679.43*** | 889.69 | -4664.75*** | 889.82 | -4671.16*** | 889.52 | -4632.20*** | 890.19 | -4599.85*** | 890.52 |
| other | -6236.16*** | 305.56 | -6258.66*** | 306.12 | -6252.56*** | 306.75 | -6261.18*** | 307.24 | -6256.25*** | 307.88 | -6258.85*** | 308.15 |
| Self-employed | -2693.96*** | 445.31 | -2693.51*** | 444.91 | -2695.30*** | 444.81 | -2691.17*** | 445.02 | -2688.54*** | 444.56 | -2688.72*** | 444.46 |
| Constant | 5044.05*** | 1083.80 | 4998.01*** | 1068.69 | 5091.71*** | 1092.20 | 4916.46*** | 1117.50 | 4664.56*** | 1153.17 | 4572.99*** | 1210.71 |
| $R^2$ (adj.) | 0.182 | | 0.182 | | 0.182 | | 0.182 | | 0.182 | | 0.182 | |

$***p \leq 0.001$, $**p \leq 0.01$, $*p \leq 0.05$

Table 4: Continued

| | $k$ = 3,200 | | $k$ = 6,400 | | $k$ = 12,800 | | $k$ = 25,600 | | $k$ = 51,200 | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Coefficient | Cl. st. err. | Coefficient | Cl. st. err. | Coefficient | Cl. st. err. | Coefficient | Cl. st. err. | Coefficient | Cl. st. err. |
| Proportion of Polish migrants within $k$ | -3312.76 | 3768.29 | -3409.97 | 4145.30 | -5716.30 | 5801.35 | -5168.20 | 5779.67 | -727.28 | 6560.63 |
| Distance needed to reach $k$ | 1.71** | 0.54 | 1.53** | 0.51 | 0.97* | 0.44 | 0.18 | 0.42 | 0.03 | 0.44 |
| Length of stay (more than 8 years = ref.) | | | | | | | | | | |
| between 5 and 8 years | -3969.36*** | 336.34 | -3968.15*** | 336.68 | -3971.37*** | 337.54 | -3978.31*** | 338.78 | -3977.83*** | 338.73 |
| between 1 and 4 years | -6425.27*** | 286.07 | -6425.04*** | 286.26 | -6435.48*** | 286.95 | -6440.53*** | 289.27 | -6439.13*** | 289.35 |
| less than 1 year | -10412.75*** | 291.76 | -10411.71*** | 292.53 | -10420.50*** | 294.12 | -10423.88*** | 294.44 | -10427.21*** | 295.81 |
| Age | 1426.58*** | 56.83 | 1426.44*** | 56.80 | 1428.42*** | 56.56 | 1429.47*** | 56.61 | 1429.43*** | 56.64 |
| Age squared | -17.47*** | 0.74 | -17.46*** | 0.74 | -17.49*** | 0.74 | -17.50*** | 0.74 | -17.50*** | 0.74 |
| Sex (male = ref.) | -6463.07 | 176.33 | -6462.43 | 176.71 | -6464.69 | 176.42 | -6463.53 | 176.10 | -6462.02 | 176.24 |
| Household characteristics (without a partner, no children = ref.) | | | | | | | | | | |
| with a partner, no children | 1186.21*** | 182.20 | 1188.51*** | 181.28 | 1189.73*** | 180.72 | 1193.00*** | 180.90 | 1191.88*** | 181.34 |
| with a partner, 1 child | -75.92 | 256.43 | -74.41 | 255.72 | -72.39 | 255.26 | -67.75 | 254.84 | -69.04 | 255.08 |
| with a partner, 2 children | -696.81 | 371.42 | -694.88 | 370.66 | -697.24 | 370.41 | -694.63 | 370.67 | -693.69 | 369.93 |
| with a partner, 3 or more children | -2146.23*** | 572.38 | -2165.54*** | 571.64 | -2163.95*** | 572.72 | -2145.89*** | 571.58 | -2134.75*** | 569.90 |
| without a partner, 1 child | -151.16 | 407.39 | -155.43 | 407.93 | -154.31 | 407.65 | -150.18 | 407.01 | -150.68 | 406.26 |
| without a partner, 2 children | -1838.81*** | 582.96 | -1844.77*** | 583.91 | -1852.34*** | 583.65 | -1840.42*** | 583.12 | -1842.72*** | 582.90 |
| without a partner, 3 or more children | -4603.78*** | 886.5 | -4604.32*** | 885.65 | -4595.46*** | 885.25 | -4607.40*** | 887.55 | -4624.52*** | 889.50 |
| other | -6252.26*** | 306.85 | -6253.65*** | 306.98 | -6241.71*** | 305.40 | -6228.36*** | 304.93 | -6230.36*** | 304.81 |
| Self-employed | -2688.66*** | 444.02 | -2686.59*** | 443.73 | -2696.54*** | 442.71 | -2687.61*** | 445.27 | -2687.04*** | 445.79 |
| Constant | 3664.89** | 1357.26 | 3152.10* | 1588.76 | 3588.62 | 1965.70 | 5346.16* | 2310.05 | 4939.70 | 2856.98 |
| $R^2$ (adj.) | 0.182 | | 0.182 | | 0.182 | | 0.182 | | 0.182 | |

$***p \leq 0.001$, $**p \leq 0.01$, $*p \leq 0.05$

absolute income measure, by means of transforming the income variable into 100 groups of equal sizes, does not alter the results of the model. When changing the level of the fixed effects dummies and the clustered standard errors to a higher level of aggregation, such as the district or the municipality level (and all combinations thereof), the results remain robust. Even though the coefficients change, the overall pattern and conclusions of the analyses remain similar. The same holds when excluding those individuals from the active population who currently do not have a job (such as people who are receiving government benefits for unemployment or illness) from the estimations. In a different type of model, where the fixed effects of the administrative neighbourhoods are removed from the model and are replaced by various socio-economic neighbourhood measurements, which correspond to the control variables in the analysis, calculated on the basis of the individualised neighbourhoods method on all eleven $k$-levels, the results remain similar (although issues with estimating the coefficients due to high levels of multicollinearity occur at higher $k$-levels).

# 8 Discussion

The influx of Polish migration to north-western European countries since Poland's accession to the European Union in May 2004 has been an emerging subject of interest for social scientists. This thesis fills the gap of scholarly knowledge on the locational patterns of Polish migrant in the Netherlands in terms of ethnic neighbourhood segregation and assesses the extent to which neighbourhood effects for the incomes of Polish migrants are present by applying the individualised neighbourhood methodology on the basis of geocoded Dutch population register data. In addition to this, the methodology of individualised neighbourhoods is compared to the conventional measure of administrative neighbourhoods. The use of such a methodology may contribute in the methodological shift the neighbourhood effects literature required to provide a more clear-cut answer on how neighbourhood effects operate (Van Ham et al., 2011).
The analyses show that Polish migrants are not equally spread over the country of the Netherlands, but live in rather clustered areas in the west, the south and the south-eastern part of the country. Although the majority of Polish migrants do not live in highly segregated areas, there is a substantial amount of Polish migrants who live in residential contexts with a relatively high number of co-ethnics in their individualised neighbourhoods.

The analyses show that the use of the individualised neighbourhood method allows for a more detailed inspection of micro-level patterns of segregation within municipalities, which would be overlooked when using conventional administrative measures of neighbourhoods. When comparing the isolation index over various scale-levels, the individualised neighbourhood method shows higher degrees of segregation on lower scale levels than on higher scale levels, which is in line with results found by Östh, Malmberg and Andersson (2014). Segregation measures based on administrative neighbourhoods would overlook and underes-

timate levels of ethnic segregation for areas which are smaller than the effective population size of the administrative neighbourhoods.

Moreover, the method of individualised neighbourhoods demonstrates the role of scale in segregation, an aspect often overlooked in the neighbourhood segregation and neighbourhood effects literature (R. Andersson & Musterd, 2010; Lupton & Kneale, 2012; Manley & Van Ham, 2012). The analyses show that the scale of ethnic segregation varies between municipalities. In Westland, segregation is most pronounced the lowest specified scale levels, whereas in the municipality of The Hague the opposite pattern is found. The analyses also show that administrative neighbourhoods have varying effective population sizes in different municipalities, making comparison between municipalities infeasible when using administrative neighbourhood measurements. The individualised neighbourhood approach is not affected by this problem as neighbourhoods are measured in the same way for all individuals.

The role of scale also becomes apparent in the multivariate analyses. The results from the fixed effects regression indicate that the magnitude of neighbourhood effects may vary by scale. In this analysis the relationship between ethnic segregation of Polish migrants based on the individualised neighbourhood method and the income of Polish individuals between the ages 18 and 64 who are participating in the labour market is estimated, while also including fixed effects for administrative neighbourhoods. In this way, the extra effect of using individualised neighbourhoods over conventional methods of measuring neighbourhoods measures can be assessed, while also decreasing the risk of omitted variable bias on the neighbourhood level.

In line with the hypotheses, the results show that there is a negative relationship between ethnic segregation and income. The negative relationship between ethnic segregation and income is particularly pronounced on smaller scale levels. This result corresponds to earlier findings emphasizing that neighbourhood effects on incomes may be stronger on smaller scales (R. Andersson & Musterd, 2010). When the buffer of nearest neighbours around an individual passes the threshold of 3,200 nearest neighbours, the negative relationship between ethnic segregation and income no longer holds. This seems to be the point where administrative neighbourhoods and individualised neighbourhoods overlap. The sensitivity analyses have shown that this relationship has proven to be robust to various specifications of the statistical model.

The analyses indicate that the main added value of the individualised neighbourhoods method is that it allows for the identification of micro-level neighbourhood effects which may be overlooked when using administrative neighbourhoods measures. The results also underline the importance of specifying the measurement of neighbourhoods on the correct scale, as noted by Van Ham and colleagues (2011). When neighbourhoods are specified on the wrong scale, certain neighbourhood effects may be underestimated or even overlooked, as different mechanisms of neighbourhood effects may operate differently on different scales.

The results from the thesis support the notion that neighbourhood segregation and neighbourhood effects are multi-scalar phenomena (W. Clark et al.,

2015; Fowler, 2016; Jones et al., 2015). Approaching neighbourhood segregation and neighbourhood effects as multi-scalar phenomena implies that neighbourhood segregation and neighbourhood effects can operate differently at different scales, as shown by the analyses. The inability to methodologically approach neighbourhood effects as a multi-scalar phenomenon could partially account for the unsatisfactory empirical results of the neighbourhood effects literature.

The conclusions, however, should be interpreted with care, as the analyses have certain drawbacks. The first drawback deals with issues related to the individualised neighbourhood, which cannot take into account the distance needed to achieve the set amount of nearest neighbours. Consequently, the 50 nearest neighbours may cover a larger area in rural areas than it does in urban areas, which may be particularly problematic in descriptive analyses when using higher $k$-levels, as illustrated by the rural municipality of Zundert. Using a distance decay function may provide a solution to overcome this problem. The function is available in EquiPop, but unexplored in this thesis.

The method of individualised neighbourhoods can also not account for natural borders of neighbourhoods. In addition to this, the 'correct' choice of the population size chosen for the individualised neighbourhood is problematic and may differ in different contexts, as illustrated by the case of Westland and The Hague. A sound comparison between the two municipalities, however, requires that segregation is measured on the same scale. When comparing different areas, it is therefore suitable to present segregation measures on different scales. Ultimately, the choice of the scale on which segregation is measured should be based the scale which is of theoretical interest.

A second issue deals with the issue of endogeneity. Residential behaviour has been argued to be subject to selection bias, meaning that the selection mechanism into certain neighbourhoods is not independent from the outcome variable (Hedman & van Ham, 2011). An alternative explanation for the lower incomes of Polish migrants who live in highly segregated areas could be the self-selection of Polish migrants into these areas. As Polish migrants are strongly represented in employment sectors such as agriculture, they may move to areas of high agricultural activity such as Westland. The lower incomes of Polish migrants who live in highly segregated areas such as Westland could then be explained by the fact that agricultural jobs are often not very high-wage jobs.

The use of cross-sectional data and the statistical analyses carried out in this thesis cannot account for this selection bias, meaning that fully causal neighbourhood effects cannot be identified. Furthermore, recent research has emphasized the temporal dimension in neighbourhood effects (Hedman, Manley, Van Ham, & Östh, 2015; Van Ham, Hedman, Manley, Coulter, & Östh, 2014). Cumulative exposure to disadvantaged is shown to have a stronger negative effect on individual outcomes than short-term exposure. Longitudinal data is needed to establish the temporal dimension of neighbourhood effects and its causal relationships.

A third issue regards the "black box" of neighbourhood effects. Underlying mechanisms of neighbourhood effects are often not explicitly researched and remain unclear (Jencks & Mayer, 1990; Galster, 2011). This criticism also applies

to this thesis, as the data used in the analyses does not allow for the explicit testing of the mechanisms behind the theories on neighbourhood effects. As argued Andersson and Musterd (2010), the collaboration of qualitative and quantitative researchers seems to be the most appropriate way to tackle this problem. This research should be aimed at identifying the specific scale at which mechanisms of neighbourhood segregation operate.

The fourth shortcoming of the analysis is the inability to take individual data regarding education. First, the available data on education of in the register data is based on surveys and data from educational institutions. There are high degrees of missing data for education, in particular for older-aged people and for those who obtained their educational degree in a country outside of the Netherlands. Polish migrants belong to this second category. When predicting income, having no data on the education level of Polish migrants may induce omitted variable bias. However, as the relationship between educational attainment income is fairly weak for Polish migrants in the Netherlands, the consequences of omitted variable bias are arguably moderate.

A last drawback of the analysis is the registration and non-registration of Polish migrants in the Netherlands. Different estimates indicate that 50% to 77% of all Polish migrants in the Netherlands are registered (Nicolaas, 2011; Van der Heijden, Cruyff, & Van Gils, 2013). Migrants who do not intend to stay in the Netherlands for a period longer than three months, however, are not required to register. As unregistered migrants are less likely to settle in the Netherlands (Nicolaas, 2011), they are less central to the analyses. It should be noted that there are also municipalities who actively incentivise employers to register their Polish workers, such in the municipality of Westland (Karpinska & Ooijevaar, 2016). Strong concentrations of Polish migrants in Westland could therefore simply be a consequence of a successful registration policy. In line with the local turn identified in migration policy research (Penninx, Kraal, Martiniello, & Vertovec, 2004), research should take into account regional variation in such policies, which can have major consequences in assessing neighbourhood effects when using register data.

Despite these limitations, the results that are presented have proven to be robust to various specifications of the model, indicating that there is a negative relationship between ethnic segregation and income of Polish migrants on small scale levels, which is not identifiable when using administrative neighbourhoods, also while controlling for individual level variables. The case of Polish migrants in the Netherlands has illustrated that neighbourhood segregation and neighbourhood effects are multi-scalar phenomena. For policy makers, this implies that the successfulness of neighbourhood mixing policies may depend on scale at which neighbourhoods are specified.

Future research should focus on whether the effects found for Polish migrants are also generalizable to other migrant groups and other countries. To go beyond correlations and fully establish causal neighbourhood effects longitudinal data is required. Furthermore, an integration of qualitative and quantitative research is needed identify the scale on which mediating mechanisms of neighbourhood effects operate to fully understand the multi-scalar nature of neighbourhood ef-

fects. In this way, we can obtain a thorough understanding of the consequences of ethnic neighbourhood segregation which can be used to develop effective policies to contribute to the economic integration of migrants.

## Acknowledgements

This thesis is part of the ResSegr programme, which is funded by JPI Urban Europe. The claims made in this thesis are made by the author and do not necessarily reflect the views of Statistics Netherlands or the ResSegr project team. Special thanks to Jeroen Ooijevaar (Statistics Netherlands) and dr. Bart Sleutjes (Netherlands Interdisciplinary Demographic Institute) for their useful comments and supervision throughout the project.

## References

Andersson, E. K., & Malmberg, B. (2015). Contextual effects on educational attainment in individualised, scalable neighbourhoods: Differences across gender and social class. *Urban Studies*, *52*(12), 2117-2133.

Andersson, R., & Musterd, S. (2010). What scale matters? Exploring the relationships between individuals' social position, neighbourhood context and the scale of the neighbourhood. *Geografiska Annaler: Series B, Human Geography*, *92*(1), 23–43.

Andersson, R., Musterd, S., Galster, G., & Kauppinen, T. M. (2007). What mix matters? Exploring the relationships between individuals' incomes and different measures of their neighbourhood context. *Housing Studies*, *22*(5), 637-660.

Atkinson, R., & Kintrea, K. (2001). Disentangling area effects: Evidence from deprived and non-deprived neighbourhoods. *Urban Studies*, *38*(12), 2277–2298.

Bakker, B. F., Van Rooijen, J., & Van Toor, L. (2014). The system of social statistical datasets of Statistics Netherlands: An integral approach to the production of register-based social statistics. *Statistical Journal of the IAOS*, *30*(4), 411–424.

Bauer, T., Epstein, G. S., & Gang, I. N. (2005). Enclaves, language, and the location choice of migrants. *Journal of Population Economics*, *18*(4), 649–662.

Borjas, G. J. (1992). Ethnic capital and intergenerational mobility. *The Quarterly Journal of Economics*, *107*(1), 123-150.

Borjas, G. J. (1995). Ethnicity, neighborhoods, and human capital externalities. *The American Economic Review*, *85*(3), 365-390.

Chiswick, B. R., & Miller, P. W. (1995). The endogeneity between language and earnings: International analyses. *Journal of Labor Economics*, *13*(2), 246-288.

Chiswick, B. R., & Miller, P. W. (1996). Ethnic networks and language proficiency among immigrants. *Journal of Population Economics*, *9*(1), 19–35.

Chiswick, B. R., & Miller, P. W. (2002). Immigrant earnings: Language skills, linguistic concentrations and the business cycle. *Journal of Population Economics*, *15*(1), 31–57.

Clark, K., & Drinkwater, S. (2008). The labour-market performance of recent migrants. *Oxford Review of Economic Policy*, *24*(3), 495–516.

Clark, W., Andersson, E., Östh, J., & Malmberg, B. (2015). A multiscalar analysis of neighborhood composition in Los Angeles, 2000-2010: A location-based approach to segregation and diversity. *Annals of the Association of American Geographers*, *105*(6), 1260–1284.

Dagevos, J. (2011). *Poolse migranten. De positie van Polen die vanaf 2004 in Nederland zijn komen wonen* [Polish migrants. The position of Polish who have lived in the Netherlands since 2004]. The Hague: Sociaal en Cultureel Planbureau.

Drinkwater, S., Eade, J., & Garapich, M. (2009). Poles apart? EU enlargement and the labour market outcomes of immigrants in the United Kingdom. *International Migration*, *47*(1), 161–190.

Edin, P. A., Fredriksson, P., & Åslund, O. (2003). Ethnic enclaves and the economic success of immigrants: Evidence from a natural experiment. *The Quarterly Journal of Economics*, *118*(1), 329-357.

Engbersen, G., Leerkes, A., Ilies, M., Snel, E., & Meij, R. (2011). *Arbeidsmigratie in vieren. Bulgaren en Roemenen vergeleken met Polen* [Labour migration in four. Bulgarians and Romanians compared to Polish]. Rotterdam: Erasmus University Rotterdam.

Engbersen, G., Okolski, M., Black, R., & Pantîru, C. (2010). Introduction: working out a way from East to West: EU enlargement and labour migration from Central and Eastern Europe. In R. Black, G. Engbersen, M. Okolski, & C. Pantîru (Eds.), *A continent moving west? EU enlargement and labour migration from Central and Eastern Europe.* Amsterdam: Amsterdam University Press.

Engbersen, G., van de Pol, S., Burgers, J., Snel, E., Ilies, M., van der Meij, R., & Rusinovic, K. (2012). *Poolse arbeidsmigranten in het Westland* [Polish labour migrants in Westland]. Rotterdam: Erasmus University Rotterdam.

Fowler, S. (2016). Segregation as a multiscalar phenomenon and its implications for neighborhood-scale research: The case of south Seattle 1990–2010. *Urban Geography*, *37*(1), 1–25.

Galster, G. C. (2011). The mechanism(s) of neighbourhood effects: Theory, evidence, and policy implications. In M. Van Ham, D. Manley, N. Bailey, L. Simpson, & D. MacLennan (Eds.), *Neighbourhood effects research: New perspectives* (pp. 23–56). Dordrecht: Springer.

Galster, G. C., Metzger, K., & Waite, R. (1999). Neighborhood opportunity structures of immigrant populations, 1980 and 1990. *Housing Policy Debate*, *10*(2), 395–442.

Gijsberts, M. I. L., & Lubbers, M. (2013). *Nieuw in Nederland: Het leven van recent gemigreerde Bulgaren en Polen* [New in the netherlands: The life of recently migrated Bulgarians and Polish]. The Hague: Sociaal en Cultureel Planbureau.

Gijsberts, M. I. L., & Lubbers, M. (2015). *Langer in Nederland: Ontwikkelingen in de leefsituatie van migranten uit Polen en Bulgarije in de eerste jaren na migratie* [Longer in the Netherlands: Changes in the living situation of migrants from Poland and Bulgaria in the first years after migration.] The Hague: Sociaal en Cultureel Planbureau.

Hedman, L., Manley, D., Van Ham, M., & Östh, J. (2015). Cumulative exposure to disadvantage and the intergenerational transmission of neighbourhood effects. *Journal of Economic Geography*, *15*(1), 195-215.

Hedman, L., & van Ham, M. (2011). Understanding neighbourhood effects: Selection bias and residential mobility. In M. Van Ham, D. Manley, N. Bailey, L. Simpson, & D. MacLennan (Eds.), *Neighbourhood effects research: New perspectives* (pp. 79–99). Dordrecht: Springer.

Iceland, J., & Weinberg, D. H. (2002). *Racial and ethnic residential segregation in the United States 1980-2000*. Washington, DC: Bureau of Census.

Jencks, C., & Mayer, S. E. (1990). The social consequences of growing up in a poor neighborhood. In L. E. Lynn & M. F. H. McGeary (Eds.), *Inner-city poverty in the United States* (p. 111-186). Washington, DC: National Academy Press.

Jones, K., Johnston, R., Manley, D., Owen, D., & Charlton, C. (2015). Ethnic residential segregation: A multilevel, multigroup, multiscale approach exemplified by London in 2011. *Demography*, *52*(6), 1995–2019.

Kanas, A., Chiswick, B. R., Lippe, T., & Tubergen, F. (2012). Social contacts and the economic performance of immigrants: a panel study of immigrants in Germany. *International Migration Review*, *46*(3), 680–709.

Kanas, A., & van Tubergen, F. (2014). The conditional returns to origin-country human capital among Turkish and Moroccan immigrants in Belgium. *Social Science Research*, *46*, 130–141.

Karpinska, K., & Ooijevaar, J. (2016). Familierelaties van Polen in Nederland [Family relationships of Polish in the Netherlands]. *Bevolkingstrends*, *2*, 1-16.

Kleinepier, T., de Valk, H. A., & Van Gaalen, R. (2015). Life paths of migrants: A sequence analysis of Polish migrants' family life trajectories. *European Journal of Population*, *31*(2), 155–179.

Lee, B. A., Reardon, S. F., Firebaugh, G., Farrell, C. R., Matthews, S. A., & O'Sullivan, D. (2008). Beyond the census tract: Patterns and determinants of racial segregation at multiple geographic scales. *American Sociological Review*, *73*(5), 766–791.

Lupton, R., & Kneale, D. (2012). Theorising and measuring place in neighbourhood effects research: The example of teenage parenthood in England. In M. Van Ham, D. Manley, N. Bailey, L. Simpson, & D. MacLennan (Eds.), *Neighbourhood effects research: New perspectives* (pp. 121–145). Dordrecht: Springer.

Manley, D., & Van Ham, M. (2012). Neighbourhood effects, housing tenure and individual employment outcomes. In M. Van Ham, D. Manley, N. Bailey, L. Simpson, & D. MacLennan (Eds.), *Neighbourhood effects research: New perspectives* (pp. 147–173). Dordrecht: Springer.

Martinovic, B., Van Tubergen, F., & Maas, I. (2009). Dynamics of interethnic contact: A panel study of immigrants in the Netherlands. *European Sociological Review*, *25*(3), 303–318.

Mincer, J. A. (1974). *Schooling, experience, and earnings.* New York, NY: The National Bureau of Economic Research.

Musterd, S. (2003). Segregation and integration: A contested relationship. *Journal of Ethnic and Migration Studies*, *29*(4), 623–641.

Musterd, S. (2005). Social and ethnic segregation in Europe: Levels, causes, and effects. *Journal of Urban Affairs*, *27*(3), 331–348.

Musterd, S., Andersson, R., Galster, G., & Kauppinen, T. M. (2008). Are immigrants' earnings influenced by the characteristics of their neighbours? *Environment and Planning A*, *40*(4), 785–805.

Nicolaas, H. (2011). Ruim helft Poolse immigranten vertrekt weer [Majority of Polish immigrants to the Netherlands is leaving again]. *Bevolkingstrends*, *1*, 32-36.

Openshaw, S. (1984). *The modifiable areal unit problem.* Norwich: Geobooks.

Östh, J. (2014). *Introducing the EquiPop software: An application for the calculation of k-nearest neighbour contexts/neighbourhoods.* Retrieved from `http://equipop.kultgeog.uu.se/`

Östh, J., Clark, W. A. V., & Malmberg, B. (2015). Measuring the scale of segregation using k-nearest neighbor aggregates. *Geographical Analysis*, *47*(1), 34–49.

Östh, J., Malmberg, B., & Andersson, E. K. (2014). Analysing segregation using individualised neighbourhoods. In C. D. Lloyd, I. G. Shuttleworth, & D. W. S. Wong (Eds.), *Socio-spatial segregation: concepts, processes and outcomes* (p. 135-161). Bristol: Policy Press.

Penninx, R., Kraal, K., Martiniello, M., & Vertovec, S. (2004). *Citizenship in European cities. Immigrants, local politics and integration policies.* Aldershot: Ashgate.

Pinkster, F. M. (2007). Localised social networks, socialisation and social mobility in a low-income neighbourhood in the Netherlands. *Urban Studies*, *44*(13), 2587–2603.

Sampson, R. J., Morenoff, J. D., & Gannon-Rowley, T. (2002). Assessing "Neighborhood effects": Social processes and new directions in research. *Annual Review of Sociology*, *23*(1), 443–478.

Sanders, J. M., & Nee, V. (1987). Limits of ethnic solidarity in the enclave economy. *American Sociological Review*, *52*(6), 745-773.

Sleutjes, B., & De Valk, H. (2015). Residential segregation patterns of migrant groups in Dutch cities: The role of scale. *NIDI Working paper*(2).

Snel, E. (2011). Wonen en wijken [Living and neighbourhoods]. In J. Dagevos (Ed.), *Poolse migranten. De positie van Polen die vanaf 2004 in Neder-*

*land zijn komen wonen* (pp. 81–89). The Hague: Sociaal en Cultureel Planbureau.

Statistics Netherlands. (2015). *Statline.* Retrieved from `http://statline.cbs.nl`

Stevens, G. (1992). The social and demographic context of language use in the United States. *American Sociological Review*, *57*(2), 171-185.

Van den Brakel, M., Bosch, M., Hartgers, M., Kessels, R., Kessels, W., Lautenbach, H., ... Van der Vliet, R. (2014). Overzicht van integratie [Overview of integration]. In R. Van der Vliet, O. J, & E. Wobma (Eds.), *Jaarrapport integratie 2014* (p. 24-74). The Hague: Statistics Netherlands.

Van der Heijden, P. G., Cruyff, M., & Van Gils, G. (2013). *Aantallen geregistreerde en nietgeregistreerde burgers uit MOE-landen die in Nederland verblijven, Rapportage schattingen 2009 en 2010* [Number of registered and non-registered citizens from CEE countries in the Netherlands 2009 and 2010]. Utrecht: Universiteit Utrecht.

Van Ham, M., Hedman, L., Manley, D., Coulter, R., & Östh, J. (2014). Intergenerational transmission of neighbourhood poverty: An analysis of neighbourhood histories of individuals. *Transactions of the Institute of British Geographers*, *39*(3), 402–417.

Van Ham, M., Manley, D., Bailey, N., Simpson, L., & MacLennan, D. (2011). Neighbourhood effects research: New perspectives. In M. Van Ham, D. Manley, N. Bailey, L. Simpson, & D. MacLennan (Eds.), *Neighbourhood effects research: New perspectives* (p. 1-21). Dordrecht: Springer.

Van Tubergen, F., & Kalmijn, M. (2005). Destination-language proficiency in cross-national perspective: A study of immigrant groups in nine western countries. *American Journal of Sociology*, *110*(5), 1412–1457.

Van Tubergen, F., & Kalmijn, M. (2009). Language proficiency and usage among immigrants in the Netherlands: Incentives or opportunities? *European Sociological Review*, *25*(2), 169–182.

Weltevrede, A., De Boom, J., Rezai, S., Zuijderwijk, L., & Engbersen, G. (2009). *Arbeidsmigranten uit Midden-en Oost-Europa. een profielschets van recente arbeidsmigranten uit de MOE-landen* [Labour migrants from Middle and Eastern Europe. a profile of recent labour migrants from CEE-countries]. Rotterdam: Erasmus University Rotterdam.

Wilson, K. L., & Portes, A. (1980). Immigrant enclaves: An analysis of the labor market experiences of Cubans in Miami. *American Journal of Sociology*, *86*(2), 295-319.

Wilson, W. J. (1987). *The truly disadvantaged: The inner city, the underclass, and public policy.* Chicago, IL: University of Chicago Press.

Wong, D. W. (1993). Spatial indices of segregation. *Urban studies*, *30*(3), 559–572.

Wong, D. W., Lasus, H., & Falk, R. (1999). Exploring the variability of segregation index D with scale and zonal systems: An analysis of thirty US cities. *Environment and Planning A*, *31*(3), 507–522.