Stockholm Research Reports in Demography | no 2022:19

COVID-19 in The Neighbourhood

The Socio-Spatial Selectivity of Severe COVID-19 Cases in Sweden, March 2020-June 2021

Juta Kawalerowicz, Agneta Cederström, Eva Andersson and Bo Malmberg

ISSN 2002-617X | Department of Sociology

Stockholm University
Demography Unit

Covid in the neighbourhood

The Socio-Spatial Selectivity of Severe COVID-19 Cases in Sweden, March 2020-June 2021

Juta Kawalerowicz

Department of Human Geography, Stockholm University

Agneta Cederström

Department of Public Health Sciences, Stockholm University

Eva Andersson

Department of Human Geography, Stockholm University

Bo Malmberg

Department of Human Geography, Stockholm University

Abstract

In this paper we analyse spatial and temporal inequalities in the risk of intensive care unit (ICU) admission for COVID-19 in Sweden between March 2020 and June 2021. The analysis is based on geocoded and time-stamped data from the Swedish Intensive Care Registry. We merge this data with a classification of Swedish neighbourhoods developed with multi-scalar measures of education, income, poverty rates, employment, social allowances, and migration. We examine 1) if residence in more socio-economically deprived or diverse types of neighbourhoods was associated with higher risk of ICU admission for COVID-19, net of known individual and neighbourhood level epidemiological factors 2) if residence in more affluent neighbourhoods was associated with lower risk of ICU admission for COVID-19 3) how have these patterns changed overtime during the three waves of the pandemic. The highest risk was associated with living in neighbourhoods characterised by rural town disadvantage coupled with diversity under wave 3. In the third wave residence in such neighbourhood types was associated with four times higher risk of ICU admission, compared to the reference category of living in homogenous rural neighbourhoods with average levels of deprivation under wave 1. Looking at disparities within each wave we found that residence in most affluent urban areas was at first associated with higher risk and then with lower risk of ICU admission for COVID-19. In contrast to earlier studies, we find that the largest inequalities between different neighbourhood types could be seen in the first wave and not the second. Overtime, the risks converged between different types of neighbourhoods.

Keywords: Covid-19, spatial inequalities in health, stages of disease, multi-scalar, k-neighbours

Stockholm Research Reports in Demography 2022:19

ISSN 2002-617X

Introduction

Since the outbreak of COVID-19 pandemic in Sweden in February 2020, the Swedish Health Agency has confirmed over 21,000 fatal cases of the disease. The pandemic left its mark on life expectancy in Sweden, which declined by more than half a year due to the excess mortality related to COVID-19 (Aburto et al. 2022). Since the early days of the pandemic, the extent and severity of the disease was not evenly distributed across space, with urban, socio-economically deprived, and ethnic minority communities hit particularly hard (Adhikari et al., 2020; Clouston et al. 2021; Sandhu et al. 2021; Kamis et al. 2021). In this paper we focus on geographic variation in the extend and severity to which different types of communities have been affected by COVID-19 in Sweden (Brandén et al., 2020; Rostila et al., 2020; Florida and Mellander 2021; Fonseca-Rodríguez et al. 2021). Some researchers suggested, in line with the *stages of disease* model, that with the emergence of new transmittable diseases spatial inequalities first increase because affluent areas have more resources to apply mitigation strategies. Then, as mitigation strategies and treatments become more widely accessible, incidence rates converge and inequalities between different communities' decline (Clouston et al., 2016). The *stages of disease* model has been tested with respect to COVID-19, but to our knowledge this is the first study which looks at overtime community-level disparities, while controlling for some of the known individual level epidemiological factors. By using data on intensive care admissions for three waves of the pandemic in Sweden (until the first week of June 2021), we examine whether disparities between different types of communities were growing and if outcomes have become more spatially polarized overtime.

This study adds to our knowledge of how COVID-19 pandemic unravelled in Sweden. Swedish studies on community-level disparities in exposure to COVID-19 have used large scale administrative units such as municipalities, regions, metropolitan areas or concentrated on urban areas such as Stockholm County or neighbourhoods in largest cities (Brandén et al., 2020; Florida and Mellandar 2021; Calderón-Larrañaga et al., 2020). Omission of rural areas leaves questions about how these areas were affected in comparison to cities, and once accounting for differences in population density. Additionally, Stockholm municipality encompasses some of most affluent, as well as least affluent areas in Sweden and using large and using aggregated geographies such as municipalities does not allow to distinguish between different types of neighbourhoods within a given municipality. To overcome this problem, we use multiscalar neighbourhood typology that was developed using 2016 grid cell data on income, education, unemployment, social assistance, and migration. By using cluster analysis on multiscalar measures tapping in different domains, we seek to capture neighbourhood boundaries that reflect daily experiences of its residents and what these residents perceive as their neighbourhood, i.e., living next to people of similar characteristics. We then model COVID-19 hospitalizations, distinguishing between poor and affluent or diverse and homogenous neighbourhoods in urban and rural settings.

Literature review

Researchers have established several factors associated with hospitalization risk and mortality from COVID-19. Characteristics related to adverse outcomes in other diseases, for instance cardiovascular diseases, have been identified as risk factors for severe COVID-19 infection. These include older age, being male, ethnic minority background, lower socioeconomic status, being unmarried (de Lusignan et al., 2020; Drefahl et al., 2020; Aradhya et al., 2020). What is known about community level factors associated with COVID-19? Early studies reported higher incidence rates of COVID-19 in areas with high population density, diverse population and overcrowded housing (Abedi et al., 2020; Chen and Krieger, 2021; Arbel et al., 2022). One emerging pattern is that COVID-19 exacerbates existing inequalities by hitting socio-economically disadvantaged communities particularity hard (Adhikari et al., 2020; Clouston et al., 2021; Sandhu et al., 2021; Meurisse et al., 2022). In particular, the role of ethnicity and economic deprivation or a possible interaction between the two has been discussed with relation to COVID-19. For example, in one of the first large American studies on community level disparities in COVID-19, Adhikari and colleagues report that poor American counties with large share of ethnic minority population had 8 times higher infection rate and 9 times higher death rate when compared to poor counties which were mostly white. For less deprived counties the gradient for diversity could be observed but it was less pronounced and non-linear. In contrast, Sandhu et al. (2021) reports in their analysis of COVID-19 in Detroit that race was not an important predictor of hospitalisation, unlike residence in economically distressed neighbourhoods. These findings suggest that the relation between community-level economic deprivation, ethnicity and severity of COVID-19 are complex, and that the role of these factors may change as the pandemic unfolds.

As data became available researchers started to look at how community and neighbourhood level factors influenced COVID-19 transmission on top of individual level predictors. For example, in a study of COVID-19 mortality among those aged 70 and older in the Stockholm County, Brandén et al. (2020) report that during the first wave both household and neighbourhood level characteristics were associated with mortality on top of individual level socio-economic predictors. The authors show that mortality was higher in districts with higher number of confirmed cases as well as more densely populated neighbourhoods. Similarly, Drefahl and colleagues observe that COVID-19 mortality mirrors in many respects overall mortality except for its spatial character, where higher mortality is observed in more urban areas (2020). For instance, in a study of COVID-19 mortality across US counties, Carozzi et al. estimate causal effects of population density. The authors report that while urban areas had higher mortality in the early onset of the disease, regional population density was not associated with COVID-19 mortality or confirmed cases in later time periods (2020). Additionally, it is possible that the association between population density and COVID-19 varies by context. In a study of municipal-level variation in incidences in Belgium during the three waves between 2020-2021, Meurisse et al (2022) find that urbanization remains a significant factor, even after controlling for median age and area deprivation, with lower predicted incidence rates in more urbanized areas.

Another question relates to how COVID-19 affected different types of communities at different points in time. Clouston et al. examine mortality trends for historical epidemics and propose the "stages of disease" model to analyse how novel disease spreads in the population (2016). In the early days of the outbreak the disease is in the *natural mortality* stage, a period when little is known about the disease. With scarce information about the disease the applicability of mitigation strategies is limited. The second stage is *producing inequalities*, it is characterised by inequal diffusion of preventive innovations. As more knowledge about the disease becomes available, groups with more resources (economic, educational, political or social) are better equipped to avoid exposure and access preventive strategies or treatments. What follows is *reducing inequalities* stage. As advantaged groups reach a point of saturation in their uptake of preventive innovations, these innovations become more evenly distributed and widely accessible. Finally, novel diseases enter the *disease reduction/elimination* stage, when transmissions are low, and we observe limited differences in mortality by SES[1]. How well does this model explain the dynamics of COVID-19 in different types of communities? Kamis et al. (2021) examine COVID-19 mortality rates across US counties and report results in line with the stages of diseases framework. The authors found that spatial disparities in mortality from COVID-19 were lower in the initial period (April-May 2020). During the second period of the pandemic (June-July 2020), county-level percentage of overcrowded households, a marker of deprivation that they adopted, was a stronger predictor of mortality compared to first and third period (August-October 2020). Similar results were reported by Clouston et al. (2021) who used survival analysis to show that in the early stages of the pandemic US county-level incidences and mortality were associated with higher SES, while later they were associated with lower SES. The authors argue that this shift started when states started implementing lockdowns. In another study Meurisse and colleagues (2022) examine the association between area level deprivation and COVID-19 incidence during three pandemic waves in Belgium. The authors find the effect of deprivation differed significantly between periods. The largest differences between deprivation quintiles were observed in wave 2, between August 2020 and December 2020, while in remaining periods differences were less pronounced.

The Swedish context

In Sweden, the initial outbreak of COVID-19 coincided with winter break at the beginning of March. Although some uncertainty remains with regards to the channels by which the virus was brought to Sweden (the point of entry hypothesis), Dyrak and Albert show that early strains of the virus came from Italy and Austria and were most likely introduced by holiday makers returning home from ski holidays. Once in Sweden, the disease started to spread quickly through community transmission. Genetic material from viral samples collected in the early phase of the pandemic shows that most people who became ill with COVID-19 in the spring of 2020 were infected in Sweden, rather than abroad (Dyrak and Albert, 2021). Another channel of transmission, which quickly reached groups that were most susceptible to the virus, was nursing homes employees, whose exposure in their own social networks was correlated with mortality in the nursing homes where they worked (Nilsson, 2021). After the initial weeks, the pandemic entered producing inequality stage in the "stages of disease" model. This started with introduction of mitigation strategies, some of which such as a recommendation to work from home were less inclusive of lower SES groups. One proposed explanation as to why low

[1] With our data ending in June 2021, it is not possible to study this final stage.

SES groups had higher mortality rate from COVID-19 is the frontline workers hypothesis. According to this hypothesis, service and care workers (taxi drivers, transportation personnel and retail workers) who are overrepresented in more disadvantaged neighbourhoods were less able to follow the recommendations issued by the authorities. These workers tend to live in multigenerational households or overcrowded housing, hence once the virus was introduced to their communities and social networks, it spread quickly (Brandén et al. 2020; Billingsley et al., 2020; Andersson et al. 2021). Additionally, some researchers draw attention to the SES gradient in knowledge of and adherence to recommendations issued by the Swedish authorities. Karim suggests that health literacy and knowledge about available treatments and interventions offers some explanation for differences in COVID-19 outcomes between different socio-economic groups. The author shows that having a medial professional in the family was associated with lower probability of not being tested prior to admission to the ICU (2021).

The question of how COVID-19 affected different communities was examined using Swedish data. Researchers reported that excess mortality at the peak of the first wave in April 2020 in the Stockholm County was higher in municipalities with lower education, income and higher share of foreign-born residents (Calderón- Larrañaga et al., 2020). Further studies have shown that municipalities with higher shares of first- and second-generation migrants experienced more hospitalizations and higher mortality in 2020 (Florida and Mellander 2021; Fonseca-Rodríguez et al. 2021). One study which looks at geographical variation for all of Sweden, albeit at the municipality level. Florida and Mellander (2021) analyse weekly infection rates in 290 municipalities during the first wave (until August 2020) and find that in comparison to spatial diffusion factors, geographic variation is only modestly associated with factors like density, population size and the socioeconomic characteristics. This leads the authors to conclude that “when it comes to place-based characteristics, there appears to be a high degree of randomness in the geographic variation of COVID-19 across Sweden”. In another study Calderón-Larrañaga et al. (2021) take a closer look at a lower geographic resolution (DeSO) which, arguably, may be a better operationalization of a neighbourhood. The study is limited to urban areas and looks only at neighbourhoods in the Stockholm County. The authors analyse excess mortality and its variation by tertiles of income, education, share of foreign-born and unemployment. They find that throughout the first wave there was a clear pattern where neighbourhoods with lowest income, lowest educational attainment, higher share of foreign born and unemployed had increased excess mortality (excess of 171%, 162%, 178% and 174%

respectively for each variable). Although they look at the period from Mach 2020 till mid-May 2020, the plots of excess mortality according to neighbourhood characteristics suggest that the gap between affluent neighbourhoods and neighbourhoods with higher share of foreign-born residents had been widening overtime. The authors conclude that “rather than being socially neutral as claimed in the early days of the pandemic, COVID-19 exacerbates existing social inequalities in health and disease.” Commenting on the overtime shifts, Sigurjónsdóttir et al. (2020) noted that districts of Stockholm with higher share of residents with foreign background experienced a rapid increase of the number of confirmed cases during the first wave but not during the second wave: “by early November (week 45) of 2020, the cumulative incidence of Covid-19 was for the first time lower in Rinkeby-Kista than in Stockholm Region as a whole, and the district, together with Spånga-Tensta, remained one of the least affected districts in Stockholm municipality for the remainder of 2020". In contrast to Stockholm, districts with higher share of population with foreign background in Malmö (Rosengård) continued to be heavily affected during the second wave.

The aim of this paper is to examine disparities in COVID-19 hospitalizations between different types of communities and at different periods of the pandemic. Following on the literature, we apply the “stages of disease” model and expect that most disparities would be observed during the second wave of the disease, while the first and the third waves would show less pronounced inequalities. Socio-economic deprivation is expected to be associated with worse outcomes. Additionally, we expect to see that disadvantaged neighbourhoods would be more exposed to the disease and that in the subgroup of disadvantaged communities those that are more diverse would be especially vulnerable. In our analysis we can distinguish between socio-economically disadvantaged neighbourhood types with high and low levels of diversity to examine differences in COVID-19 hospitalizations risks for residents living in them. Finally, the question of whether rural areas were less affected, once controlling for population density, remains debated. In this analysis we can study differences in risk of ICU admission by different neighbourhood types in urban and in rural settings.

Data and Method

Previous studies often focused on COVID-19 mortality or confirmed cases as dependent variables. In this study we use data on intensive care admissions. The risk of death from COVID-19 increases steeply with older age while testing is more likely to be accessed by individuals with higher SES, access to information about medical care options and those who live closer to testing facilities (Leung 2020; Karim 2021; Green et al., 2021, Karim, 2021). Intensive care admissions are not free of biases, but they are a better way to capture the extent and severity with which different types of communities have been affected by COVID-19. Data on COVID-19 intensive care unit (ICU) admissions comes from SmiNet database from National Board of Health and Welfare and was sourced by the Swedish Public Health Agency. This data was then combined with Swedish population register data maintained by Statistics Sweden. Our sample consisted of 10310465 individuals, 7397 of whom were admitted to ICU between March 2020 and the first week of June 2021. Data on ICU admissions was time-stamped and by merging it with registers we could establish places of residence for individuals in the sample. Admission to ICU could be classified into waves as follows: first wave (March 2020-August 2020), second wave (September 2020-January 2021) and third wave (February 2021- first week of June 2021). Figure 1 shows the distribution of ICU admissions in the period under study.

Knowing places of residence, we can assign individuals in the sample to their residential neighbourhood cluster types. This way we get a descriptive overview of the type of neighbourhoods where individuals live. We use Swedish register data to create a fine-grained neighbourhood typology based on multi-scalar measures of population composition computed for individualized neighbourhoods with equal population size. Our indicators measure the extent of socio-economic and ethnic segregation at nine scales ranging from 200 to 51200 closest neighbours. We use hierarchical clustering methods to develop a typology where we can assign each of 213663 residential areas (i.e., inhabited grid cells) in Sweden into one of ten cluster types which very in terms of socio-economic affluence, attachment to the labour market and migration at both micro and macro scales. We control for individual level variables known to affect risk of sever COVID-19 infection and exposure: sex, age, region of origin as well as neighbourhood population density (DeSo). Table 1 provides information about descriptive statistics on the variables we use. To show how neighbourhood cluster type contributed to risk of ICU admission we fit Poisson models with robust standard errors. These will be first run for the sample, next we fit models where we interact neighbourhood cluster type with wave. Results are shown in Tables A1-A2. In the text we use Figures 9-11 to discuss the results.

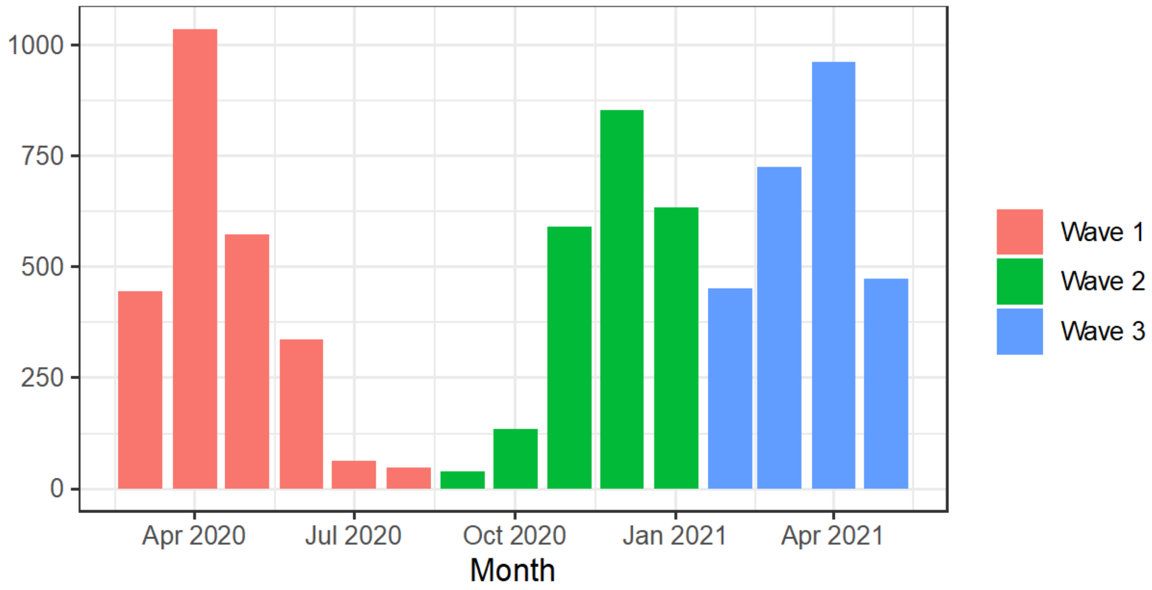

Figure 1 COVID-19 ICU admissions by wave

	count	percent	count in ICU	percent in ICU
Sex				
Men	5183788	50.3	5173	70.1
Women	5126677	49.7	2206	29.9
Age				
[0,20)	2288660	22.2	66	0.9
[20,30)	1281292	12.4	176	2.4
[30,35)	733135	7.1	150	2
[35,40)	654307	6.3	165	2.2
[40,45)	631144	6.1	270	3.7
[45,50)	665194	6.5	474	6.4
[50,55)	667762	6.5	693	9.4
[55,60)	642307	6.2	894	12.1
[60,65)	571678	5.5	1071	14.5
[65,70)	541531	5.3	1084	14.7
[70,75)	559451	5.4	1090	14.8
[75,80)	470114	4.6	808	10.9
[80,85)	296697	2.9	346	4.7
[85,90)	181623	1.8	82	1.1
[90,100)	121517	1.2	10	0.1
[100,115]	4053	0	0	0
Region of Origin				
Sverige	8299298	80.5	4540	61.5
EU/EEA	578518	5.6	579	7.8
non EU/EEA	1432649	13.9	2260	30.6
Population density in DeSO (persons/sqkm)				
[0,139]	2062985	20	1153	15.6
(139,487]	2062062	20	1207	16.4
(487,1690]	2061622	20	1481	20.1
(1690,5170]	2061309	20	1597	21.6
(5170,57700]	2061736	20	1938	26.3

Table 1 Descriptive statistics for sample

Neighbourhood cluster types

We use neighbourhood cluster typology to overcome the problems mentioned earlier in this paper. This neighbourhood cluster types allows to draw crisp boundaries between neighbourhoods with similar characteristics, not according to administrative units but characteristics of the residents and their local surroundings. Researchers often use fixed geographical sub-divisions such as Census tracts even though results of such analyses are strongly influences by how boundaries of such sub-divisions are drawn (Openshaw 1984). Second, our classification uses individualized scalable neighbourhoods (Östh et al., 2014; Hennerdal, 2019) which means that it looks at both smaller and larger special context. Individualized neighbourhoods can be constructed by expanding a buffer around different residential locations until the population encircled by the buffer corresponds to a selected population threshold (the so-called k-neighbour approach). When this threshold is reached, one computes aggregate statistics on selected socio-economic variables for the encircled population. By varying the population threshold, contextual measures computed in this way can be designed to focus only on the closest neighbours or on a larger number of neighbours. Such multiscalar approach is useful for epidemiological studies where proximity to contrasting types of neighbourhoods (a situation where a neighbourhood is rather affluent but borders another type of neighbourhood with very different characteristics) or the scale of segregation (our typology distinguished between small and larger scale deprived diverse neighbourhood types) could be an additional factor.

For developing the neighbourhood cluster typology, we allow the scale (i.e., the number of the closest neighbours) to vary from 200 to 51200 closest neighbours in successive doublings of the population thresholds. Seven individual level indicators were extracted to use as input: (1) Having a tertiary education, (2) Having taxable income in the highest decile, (3) Being in employment, (4) Having received social allowance during the year, (5) Being at risk of poverty, (6) Country of birth outside of Sweden in EU/EFTA country, (7) Country of birth outside of Sweden in non-EU/EFTA country. The individual level data was aggregated to 250-meter grid cell squares or residential areas based on their geo-coordinates[2].

Using 7 indicators at 9 scales (200, 400, 800, 1600, 3200, 6400, 12800, 25600 and 51200) gave a total of 63 measures of neighbourhood context that can be used to classify residential areas

[2] In a small number of cases (mostly in sparsely populated locations) 250-meter squares were not available, then data was aggregated at 1,000-meter squares.

using cluster analysis. However, since many of these variables will be highly correlated, we used factor analysis that compresses the 63 original indicators to 8 orthogonal factors before proceeding with the cluster analysis (for a similar method see Clark et al. 2015). More details about the procedure can be found in the Appendix. Table 2 presents an overview of the 10 neighbourhood cluster types that were identified. More details on the distribution of variables for different scales in each neighbourhood cluster type are shown in Figures 2-8. The x axis refers to k-scale, or the threashold for k nearest neighbours that are considered, while y axis refers to proportion among the k nearest neighbours with a given characteristic. Our classification offers a data-driven insight into what are the most typical residential contexts in contemporary Sweden and how they are distinguished in terms of socio-economic affluence, labour market attachment and immigration at different scales. Five neighbourhood cluster types are identified as predominantly rural and five as predominantly urban. In the remaining analysis we break down one urban cluster, Urban diverse (U_DIV) into its core and a buffer zone. The making of neighbourhood cluster types is explained in detail in the Appendix.

Name	**Cluster**	**Key characteristics**
Rural town diversity	RT_DIV	Small scale migration
Rural town adjacent	RT_ADJ	Adjacent to social assistance
Rural town working-class	RT_WC	Employed with low income, EU migrants
Rural homogenous	R_HOM	Few migrants
Rural border area	R_BOR	Low registered income, EU migrants
Urban diverse	U_DIV	Large scale migration
Urban adjacent	U_ADJ	High contrast over scales
Urban homogenous	U_HOM	Medium academic with high income
Urban academic	U_ACA	Academic with medium income
Urban elite	U_ELI	Academic with high income

Table 2 Neighbourhood cluster types and descriptions

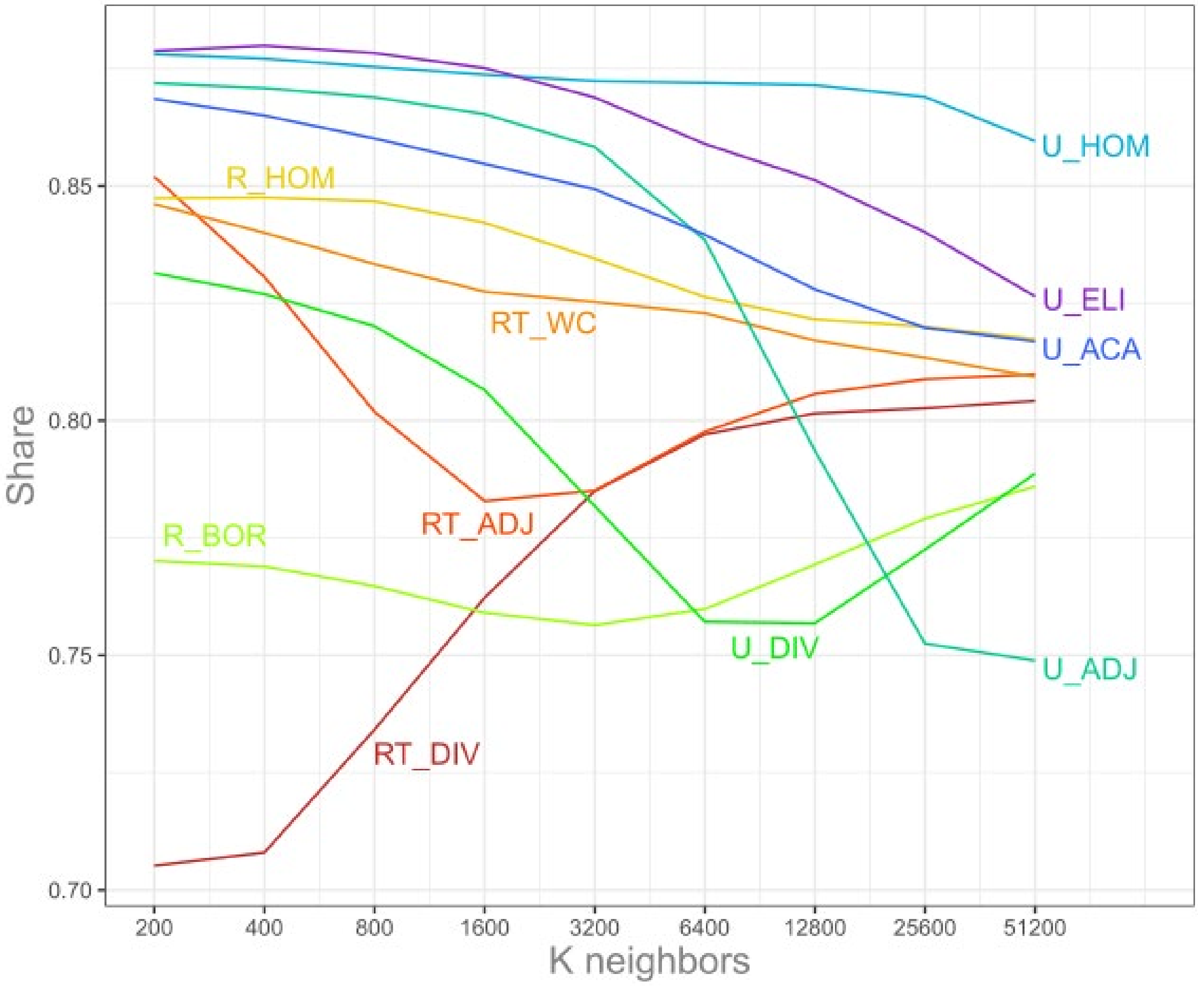

Figure 2 Share in employment by neighbourhood cluster type

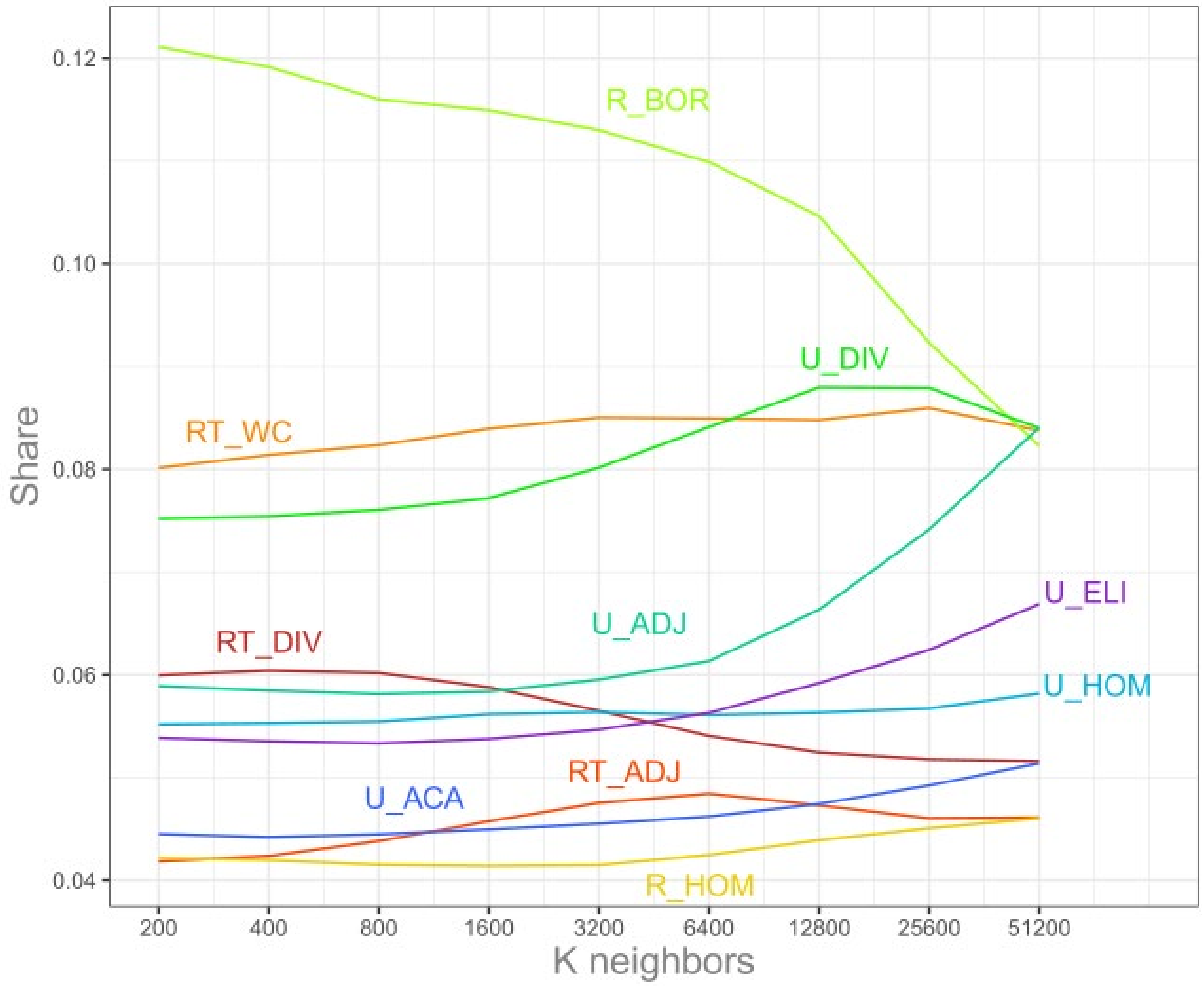

Figure 3 Share of European born by neighbourhood cluster type

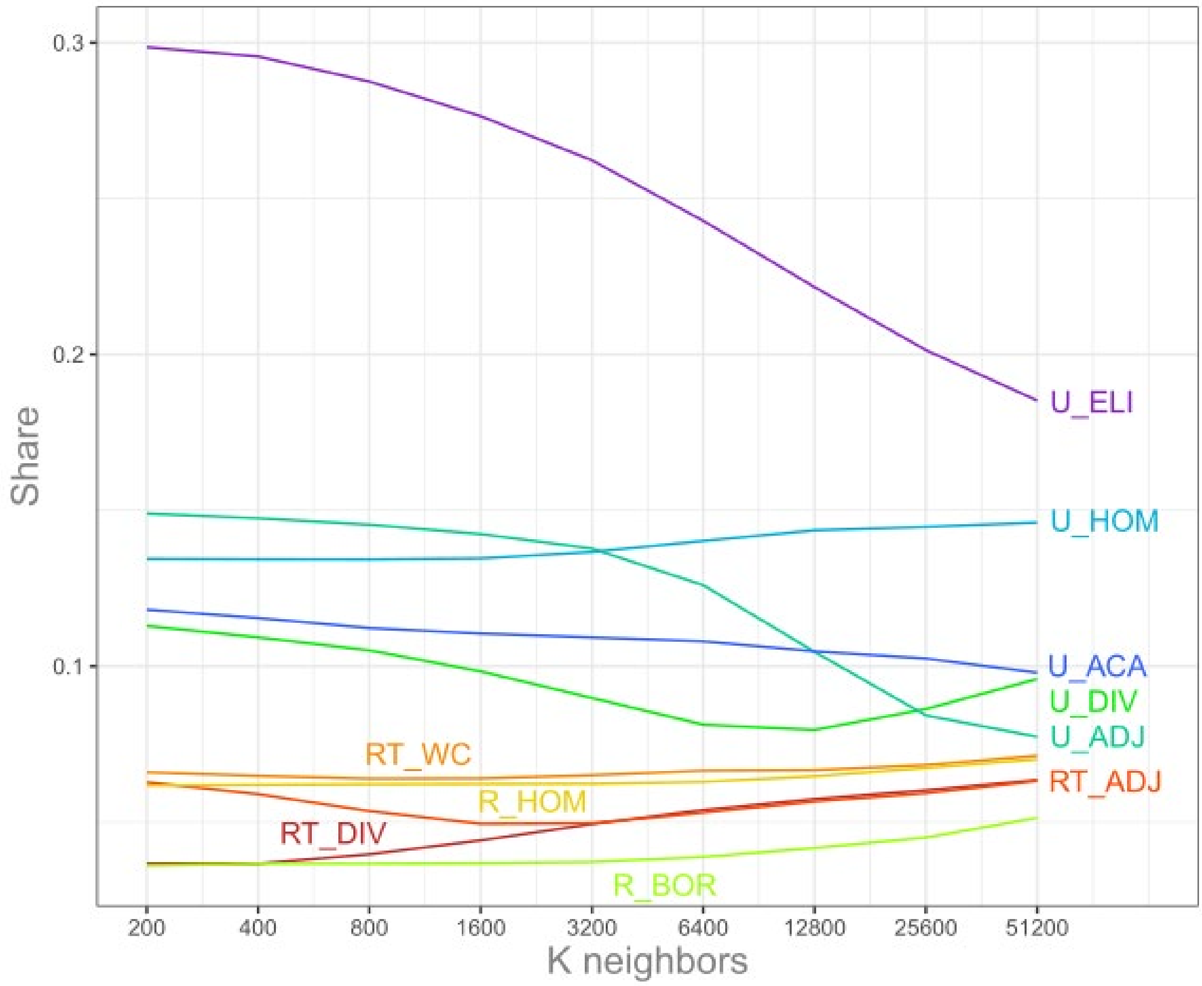

Figure 4 Share in top income decile by neighbourhood cluster type

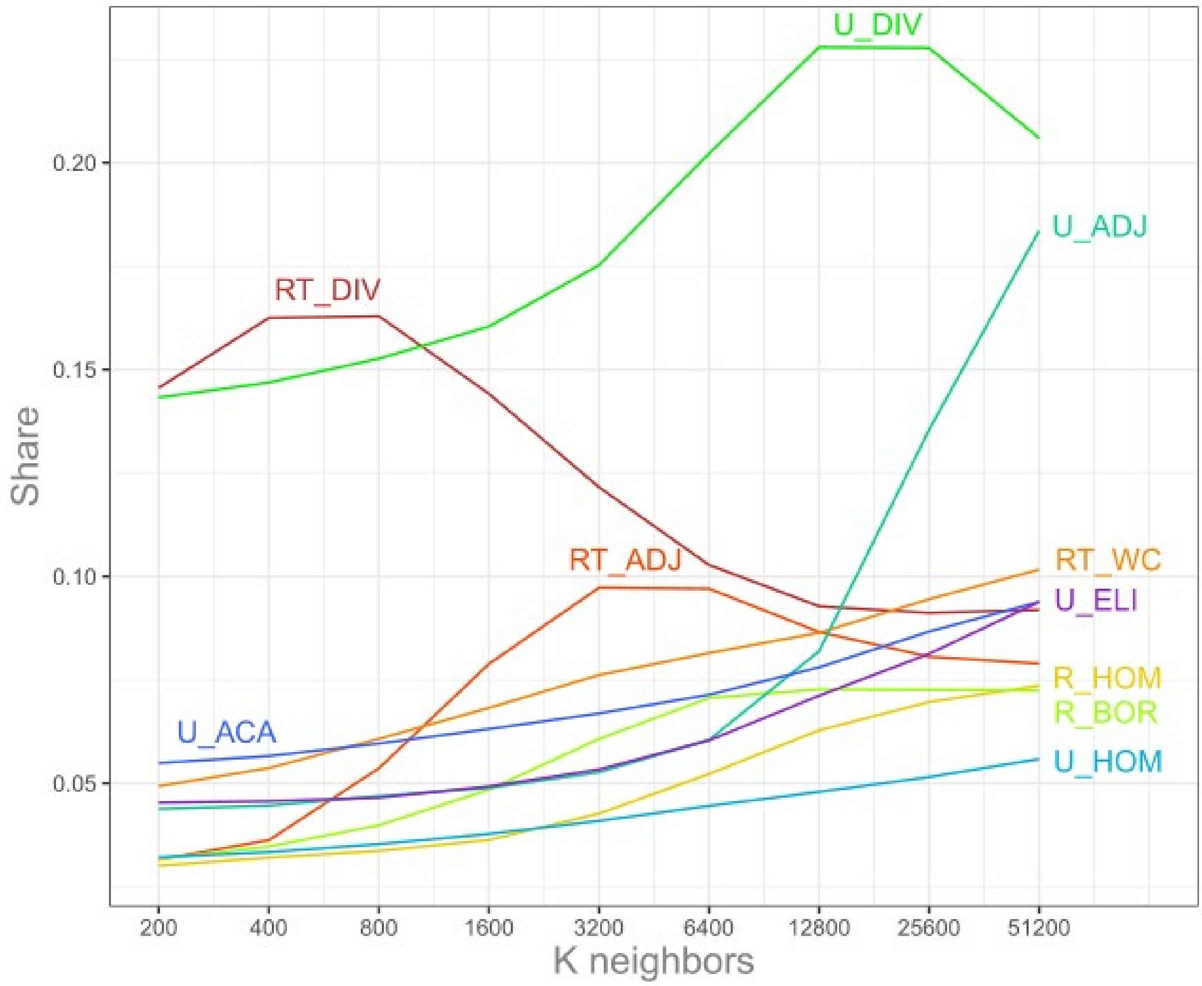

Figure 5 Share non-European born by neighbourhood cluster type

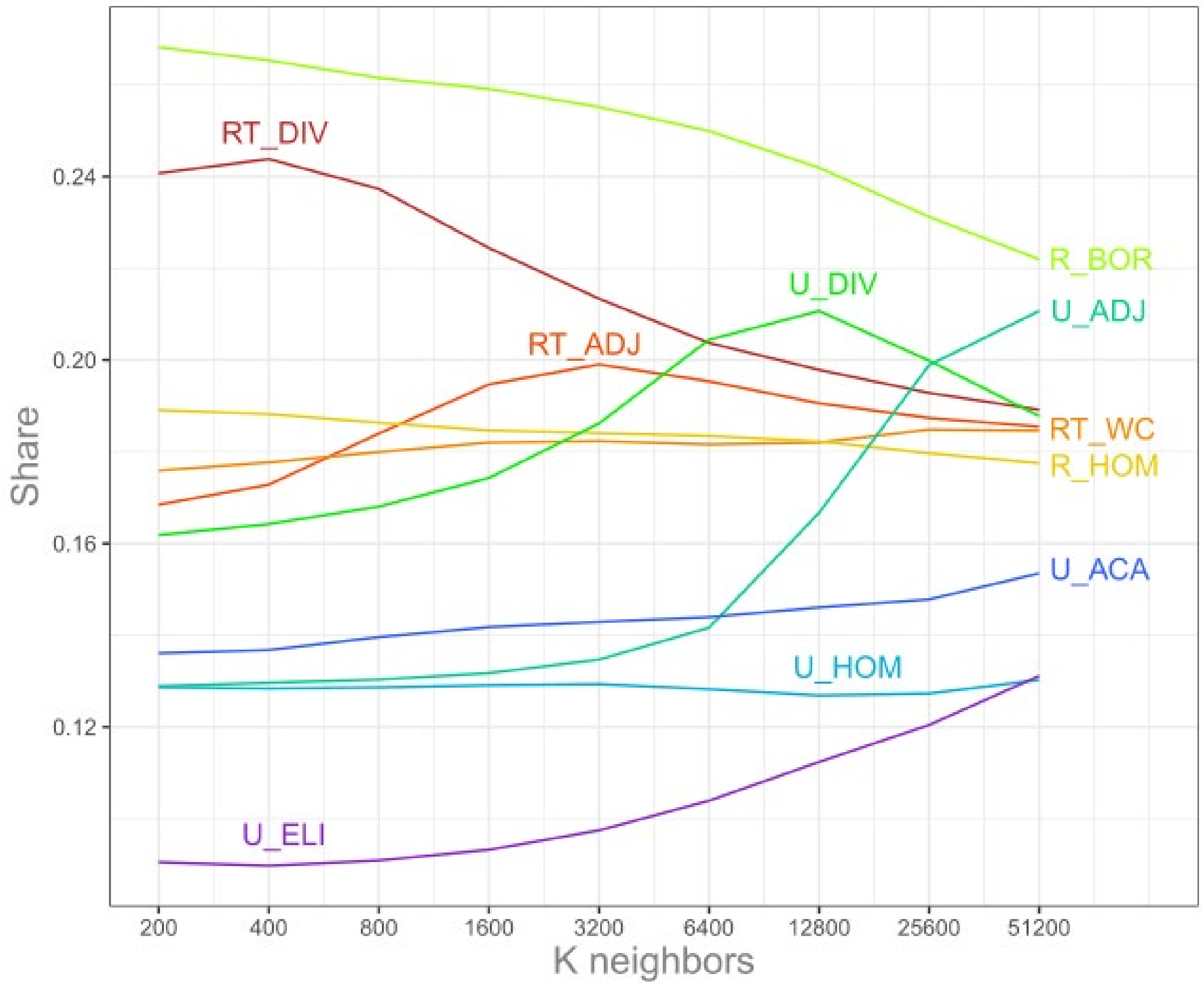

Figure 6 Share in poverty by neighbourhood cluster type

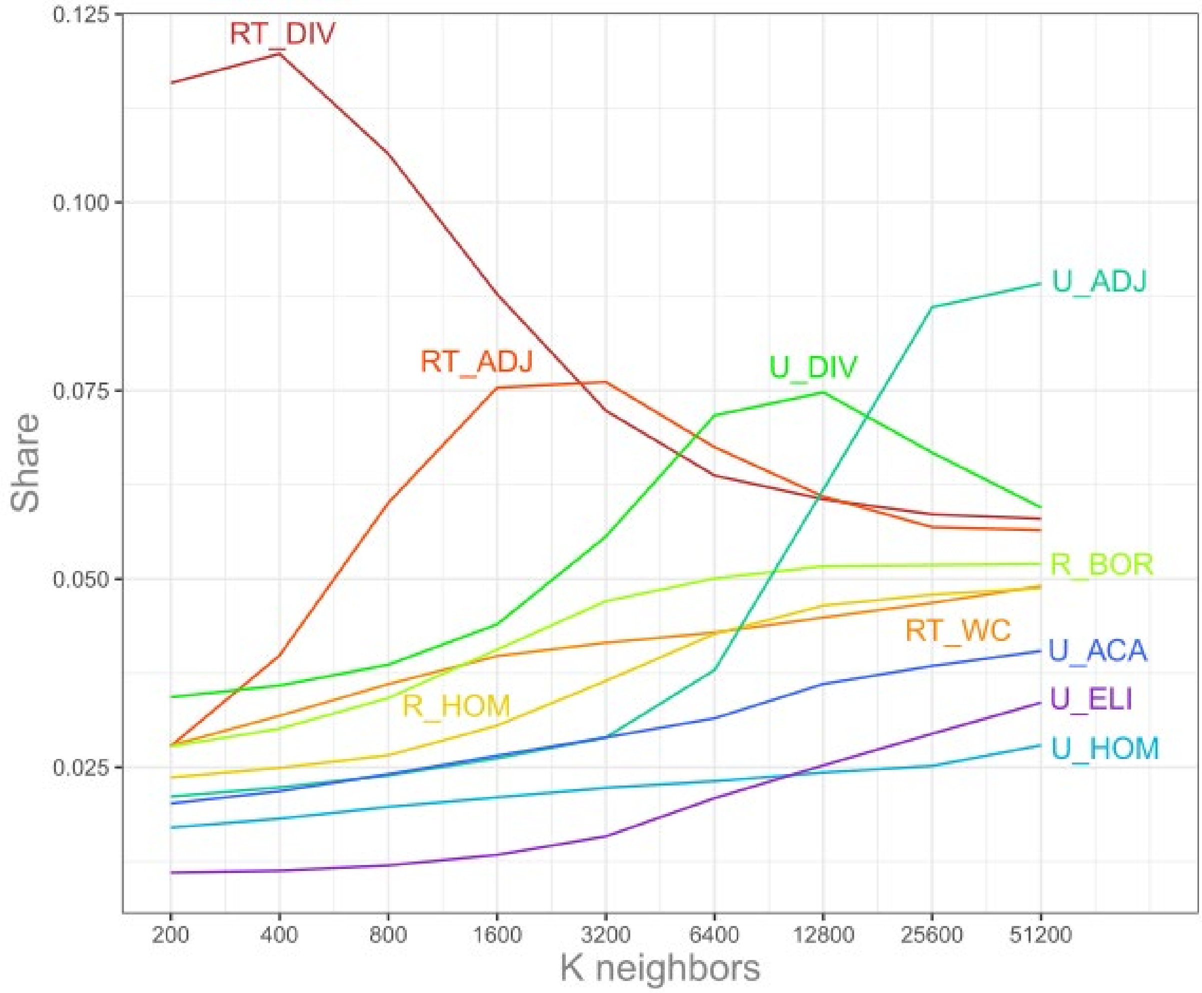

Figure 7 Share with social assistance by neighbourhood cluster type

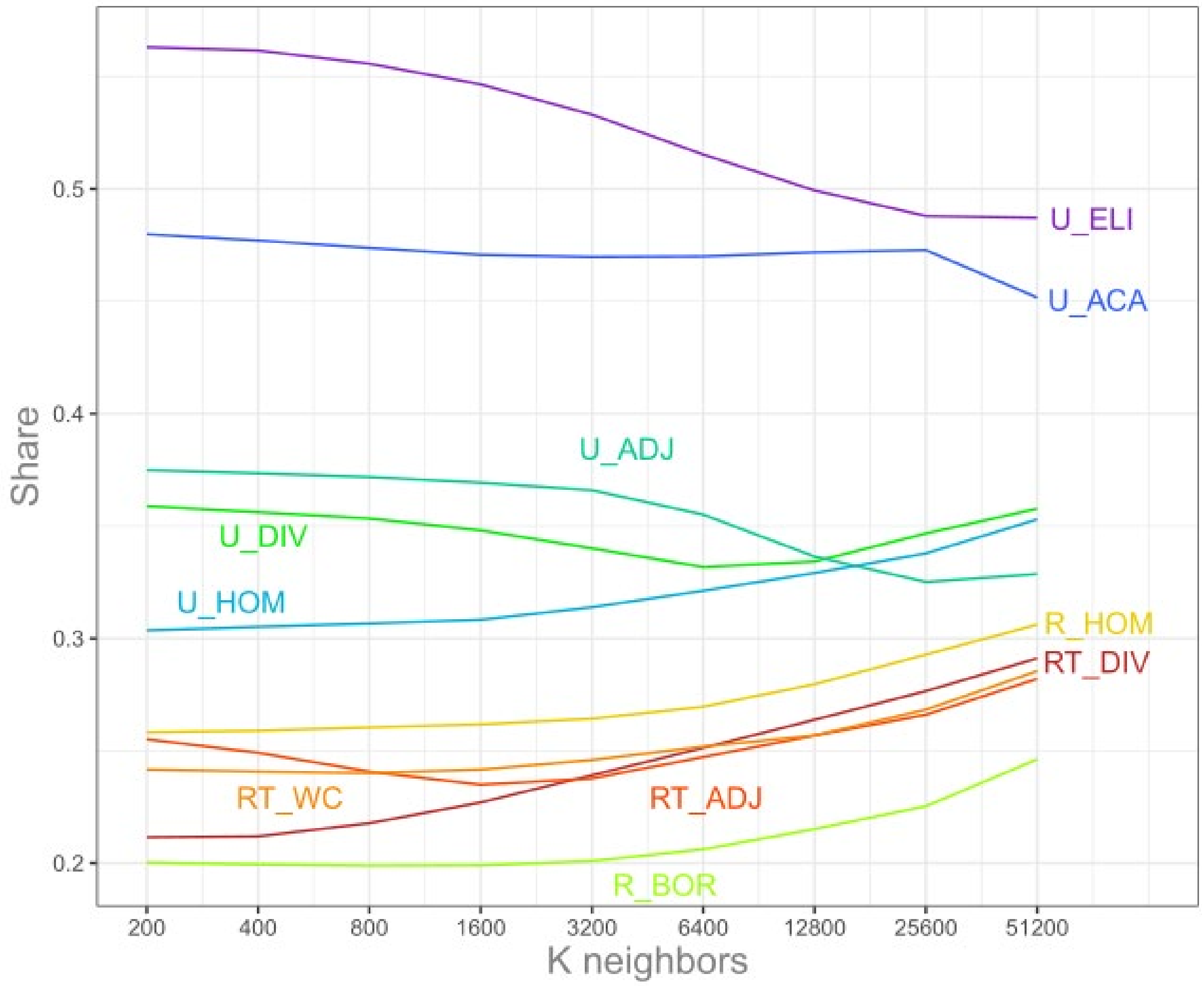

Figure 8 Share tertiary education by neighbourhood cluster type

Results

ICU admissions for COVID-19 and neighbourhood cluster types

Figure 9 show incidence rate ratio (RR) for ICU admission with COVID-19 for all three waves together. Figures 10 shows incidence rate ratio from an interaction model with reference category set as one of the neighbourhood cluster types (R_HOM) in wave 1 while Figure 11 shows results from a model when the reference category is set as the same neighbourhood cluster types (R_HOM) in each wave. We present both models, since Figure 10 is better for interpreting the overall relative risks and to establish residence in which neighbourhood cluster types and at which time were associated with highest relative risk of ICU admission, while Figure 11 is more suitable for showing what differences in relative risk looked like for subsequent waves.

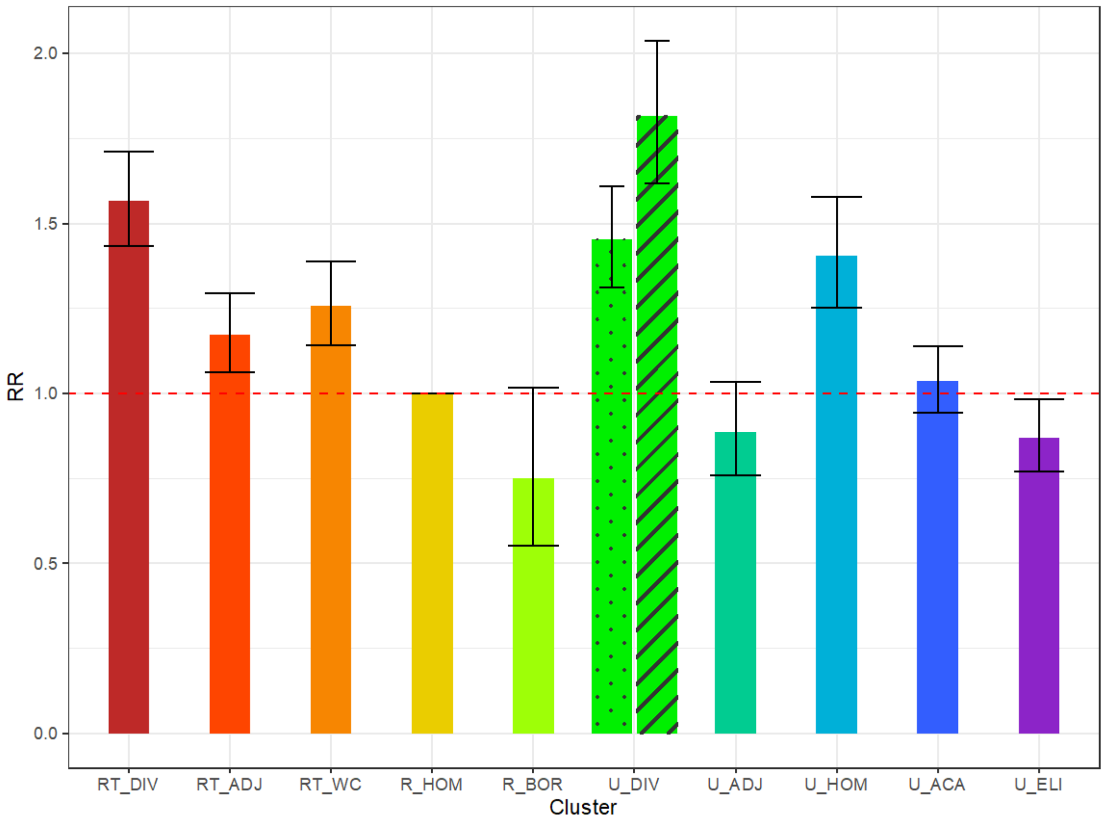

Figure 9 Predicted relative risk by neighbourhood cluster type. Poisson model with controls for sex, age, region of birth and neighbourhood population density.

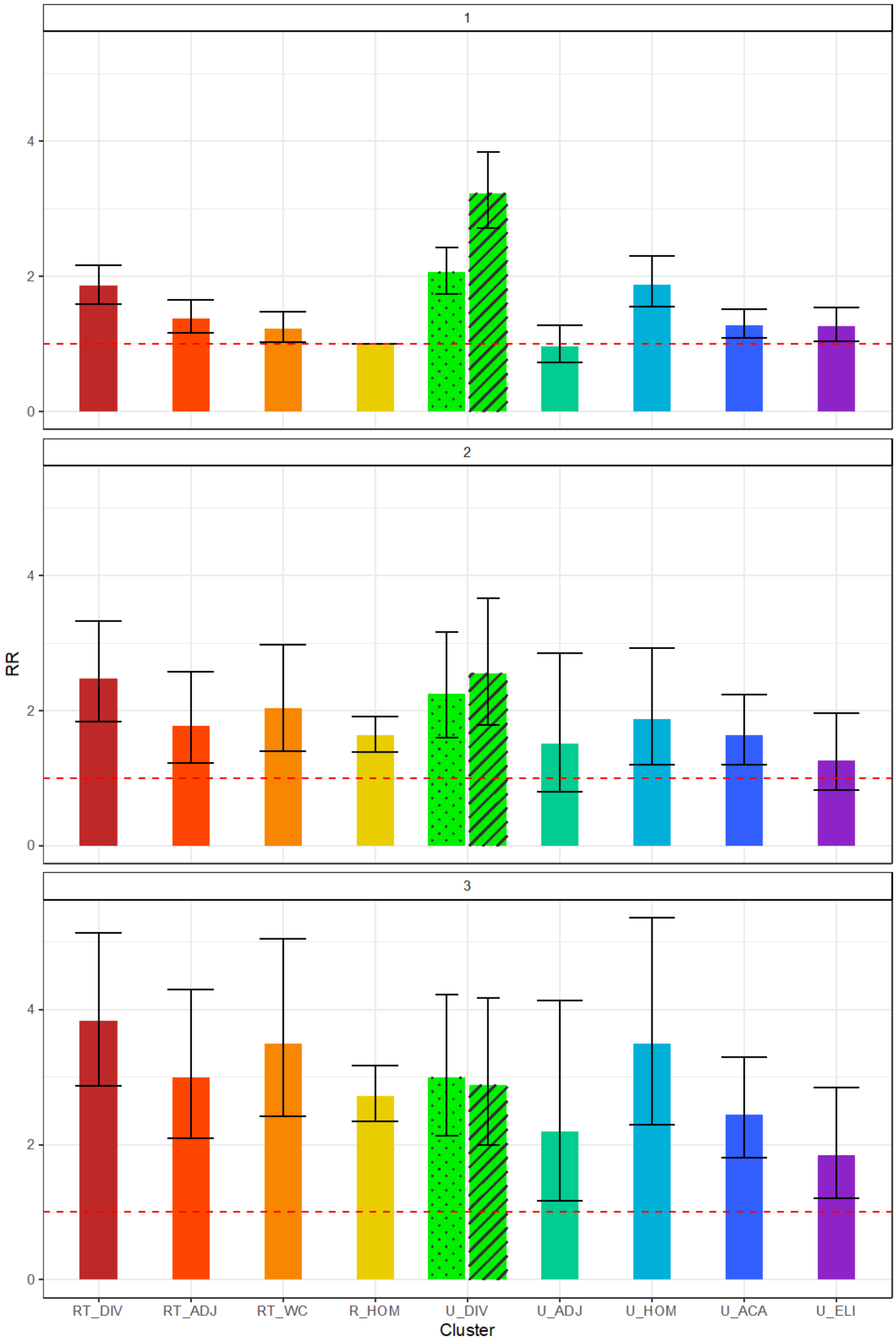

Figure 10 Predicted relative risk by neighbourhood cluster type and wave. Poisson model with controls. Reference category set as R_HOM in wave 1. Note that the bar for R_BOR was ommited due to high confidence intervals in wave 2 and 3.

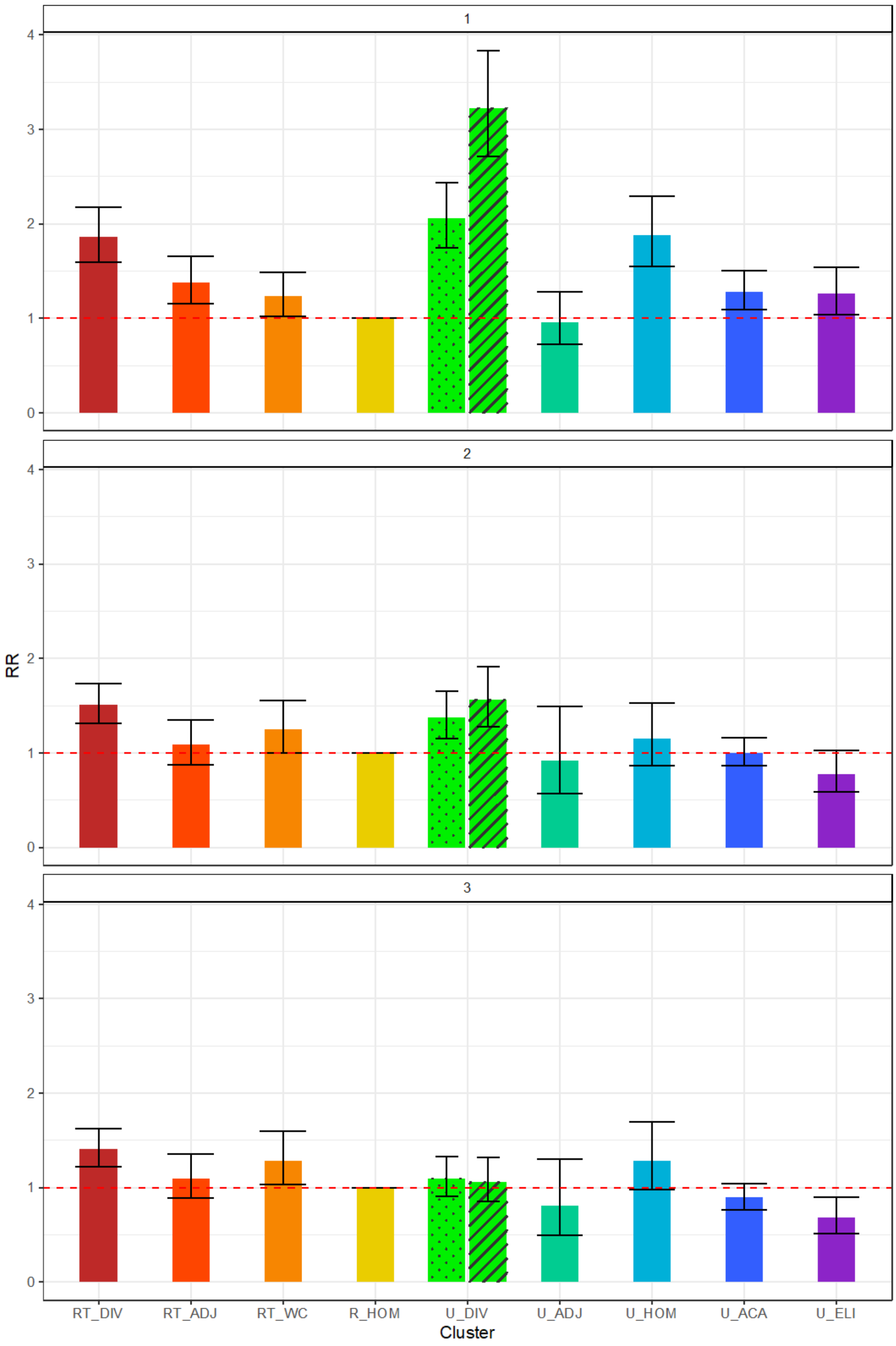

Figure 11 Predicted relative risk by neighbourhood cluster type and wave. Poisson model with controls. Reference category set as R_HOM in each wave. Note that the bar for R_BOR was ommited due to high confidence intervals in wave 2 and 3.

ICU admissions for COVID-19 by neighbourhood cluster type

Figure 9 shows differences in how residence in different neighbourhood cluster types was associated with the risks of ICU admission for COVID-19. Residence in Urban diverse core

(U_DIV) neighbourhood cluster type was associated with nearly doubling (RR 1.82, 95% CI 1.62-2.04) of the risk of ICU admission with COVID-19, as compared to residence in rural homogenous (R_HOM) neighbourhood cluster type. Importantly, this is not an artefact of higher population density in Urban diverse area (and hence transmissibility of the disease) because the model controls for population density at the neighbourhood level. Relative risk for ICU admission is also high for Rural town diverse (RT_DIV) neighbourhood cluster type (RR 1.57, 95% CI 1.43-1.71) and Urban diverse buffer (U_DIV) which is a residual category for Urban diverse neighbourhood cluster type (RR 1.62, 95% CI 1.82-2.04). Somewhat surprisingly, the fourth highest relative risk was observed for Urban homogenous (U_HOM) neighbourhood cluster type (RR 1.41, 95% CI 1.25-1.58). This neighbourhood cluster type can be described as places where high labour market attachment and relative affluence are not coupled with educational attainment and which is also, as for the urban context, relatively isolated from diversity. The neighbourhood cluster types associated with lowest relative risks are Rural border (R_BOR) and Urban elite (U_ELI), with relative risk 0.75 (95% CI 0.55-1.02) and 0.87 (95% CI 0.77-0.98) respectively. Although we do see some overall urban-rural gradient, high relative risk associated with living in Rural town diverse (RT_DIV), a neighbourhood cluster type commonly observed in centres of small and middle-sized towns, does not allow to make strong conclusions in this regard (RR 1.57, 95% CI 1.43-1.71).

ICU admissions for COVID-19 by neighbourhood cluster type and wave

Figures 10-11 further divide relative risks for neighbourhood cluster types by waves. To make these figures clearer we omitted Rural Border (R_BOR), a neighbourhood cluster type which had wide CI in wave 2 and 3. Looking at Figure 10, the highest relative risk for ICU admission for COVID-19 was associated with residence in Rural town diverse (RT_DIV) neighbourhood cluster type under wave 3 (RR 3.83, 95% CI 2.88-5.14). This cluster refers to smaller scale diversity and disadvantage and it is often observed in smaller cities and towns. For the majority of neighbourhood cluster types we observed that the relative risk increases with time. A clear example is Rural homogenous (R_HOM) neighbourhood cluster type where 95% CI do not overlap between consecutive waves, but we also see a similar trend for Rural town diverse (RT_DIV) or Rural town working-class (RT_WC). This means that with time residence in these neighbourhood cluster types became associated with increasing risk of ICU admission for COVID-19, relative to wave 1. The pattern for decreasing relative risk is not that clear-cut. For example, residence in Urban diverse (U_DIV) core neighbourhood cluster type was associated with lower relative risk during wave 2, compared to wave 1 (although the difference is not statistically significant because 95% CI overlap). Next, we look at Figure 11 to see how differences between neighbourhood cluster types evolved within waves. Here, differences between neighbourhood cluster types became less pronounced overtime. During wave 1 we see the largest differences, which was driven by elevated relative risk for residence in Urban diverse (U_DIV) core neighbourhood cluster type. Additionally, with exception of Urban adjacent (U_ADJ) neighbourhoods cluster type, living in urban clusters was associated with higher risk for ICU admission, compared to reference category of residence in Rural homogenous (R_HOM) neighbourhood cluster type. During wave 2 we see that residence in some of urban neighbourhood cluster types that was associated with increased risk of ICU admissions was no longer a significant factor. Such neighbourhood cluster types are: Urban elites (U_ELI), Urban homogenous (U_HOM) and Urban academic (U_ACA). By wave 3 the differences between relative risks associated with residence in different neighbourhood cluster types are not significant. There are three exceptions. First there is residence in Urban elite (U_ELI) neighbourhood cluster type which is associated with significantly lower relative risk of ICU admission for COVID-19 than other clusters (RR 0.68, 95% CI 0.51-0.90). Second, residence in Rural town diversity (RT_DIV) and Rural town working-class neighbourhood cluster type is associated with higher risk of ICU admissions (respectively RR 1.41, 95% CI 1.23-1.62 and RR 1.28, 95% CI 1.03-1.59). This is interesting, especially given than by wave 3 we see that living in Urban diverse core (U_DIV) neighbourhood cluster type is no longer associated with increased risk of ICU admission. Instead, there are two other neighbourhood cluster types which replace this neighbourhood cluster type as a place with the highest risk of ICU admissions – Rural town diverse (RT_DIV) and Rural town working-class (RT_WC).

Discussion

In this study we investigate individual level risk of ICU admissions for COVID-19 in different types of neighbourhoods in Sweden. The study contributes to a broader literature looking at spatial inequalities in health. We use high quality data on ICU admissions for COVID-19 from the Swedish Intensive Care Registry and cover a period of 15 months. Our study is unique in this regard, as many previous studies of COVID-18 focused on shorter time periods or looked at the initial wave. As suggested by the "stages of disease" model, the stage at which data is collected and analysed is likely to play a role in what conclusions are drawn. For example, Abedi et al., 2020 who looked at cases in seven stages (including Louisiana) until April 2020 reported higher of COVID in areas with higher median income in US while for instance

Madhav et al. 2020, who included a longer period and focused on Louisiana, reported that area deprivation is associated with higher risk of COVID-19. Hence, in future studies of how context affects COVID-19 it is important to look at how these associations unfold in time.

Another novelty of this approach is that we use grid cell data to create a neighbourhood classification which allows distinguishing between different types of neighbourhoods in a “crispier” way. We construct it by using both characteristics of immediate and more remote neighbourhood and we use clustering algorithms to arrive at a classification distinguishing 10 predominant neighbourhood cluster types in Sweden. Our analysis reveals stark differences between risk of ICU admissions for residents of different types of neighbourhoods. We found that the highest risk was associated with residence in Rural town diverse (RT_DIV) neighbourhood cluster type, especially during the third wave. In consecutive waves the inequalities in risk between different neighbourhood cluster types decline, mostly due to decline in relative risk for Urban diverse (U_DIV) cluster. One hypothesis is that these areas are home to more mobile international populations and are hence more exposed to a global pandemic. Another line of inquiry which deserved further examination is the occupational structure of such neighbourhoods and their lower ability to self-isolate or for residents to work remotely. Both would put these areas at risk of being exposed early to the virus. Why do we see most pronounced overtime decline in Urban diverse neighbourhood cluster type? It could be because behavioural adaptation, large sections of susceptible populations developing immunity or because of improving intervention from local authorities. As demonstrated in Kwon et al. (2021) higher rates of COVID-19 infections are associated with weaker social distancing at the community level and differences in individual protective behaviour (in the Kwon et al. study, measured by self-reported wearing of face masks). Thus, it could be that such factors have played a role in determining relative differences in ICU admissions for COVID-19 across neighbourhood types. If this is the case, convergence in admittance rates could be due to a convergence in behaviour.

Our study corroborates earlier findings on regional differences in Sweden. The convergence that is present in the Swedish data can also be observed in the data on the weekly subnational 14-day notification rate of new COVID-19 cases that is released by European Centre for Disease Prevention and Control (ECDC, 2021, 19 aug) or in the study of Covid cases in different parts of Stockholm and Malmö by Sigurjónsdóttir et al. (2020). The ECDC study showed that during the first 8 months of the epidemic the coefficient of variation in the cumulative rate of infection across NUTS regions were at 100% or above but then declined to below 40%. The pattern of convergence that we observed for neighbourhood level factors has previously been reported by other researchers for individual level factors. For instance, Andersson et al. 2021 observed a decline in the coefficient size for country of birth in COVID-19 mortality overtime, here we also see that residence in urban neighbourhoods with highest diversity was associated with increased risk during the first wave, but this risk became insignificant by the third wave. Importantly, there may be a different dynamic for diversity and disadvantage in urban and rural areas, since we observe that living in rural town diversity (RT_DIV) neighbourhood cluster types is a risk factor for COVID-19 ICU admission during all waves.

Moreover, the overtime convergence in relative risk is in line with prediction form the "stages of disease" model, however here we observe that the inequalities in risks are largest already during the first wave, and not in the second wave as some other researchers report for other countries. This may be due to different criteria used for constructing waves (especially for Kamis et al. 2021). However, the periods used in the study by Meurisse and colleagues corresponds closer with periods used for the current study and the authors find that in Belgium inequalities were at its peak during the second wave. It seems that in Sweden producing inequalities stage, when different access to resources produces inequalities in exposure to disease and risk of transmission, might have been reached earlier. Sweden is an interesting outlier with respect to the level of restrictions that were put in place during the pandemic (Born et al. 2021), and it is possible that these policies, together with spatial inequalities in access to healthcare, contributed to the production of spatially unequal outcomes.

Acknowledgements

We acknowledge financial support from the Swedish Foundation for Humanities and Social Sciences (Riksbankens Jubileumsfond, RJ), grant registration number M18-0214:1 and Swedish Research Council for Health, Working Life and Welfare (Forte), grant registration number 2016-07128.

References

Abedi, V., Olulana, O., Avula, V., Chaudhary, D., Khan, A., Shahjouei, S., Li, J., & Zand, R. (2021). Racial, economic, and health inequality and COVID-19 infection in the United States. *Journal of racial and ethnic health disparities*, *8*(3), 732-742.

Aburto, J. M., Schöley, J., Kashnitsky, I., Zhang, L., Rahal, C., Missov, T. I., Mills, M., & Kashyap, R. (2022). Quantifying impacts of the COVID-19 pandemic through life-expectancy losses: a population-level study of 29 countries. *International journal of epidemiology*, *51*(1), 63-74.

Adhikari, S., Pantaleo, N. P., Feldman, J. M., Ogedegbe, O., Thorpe, L., & Troxel, A. B. (2020). Assessment of community-level disparities in coronavirus disease 2019 (COVID-19) infections and deaths in large US metropolitan areas. *JAMA network open*, *3*(7), e2016938-e2016938.

Andersson, G., Drefahl, S., Mussino, E., Modig, K., & Mayer, A. (2021) Dödlighet och sjuklighet i covid-19 i Sverige under mars 2020-februari 2021. Underlagsrapport till SOU 2021:89 Sverige under pandemin Stockholm 2021. Submission to Swedish Coronakommissionen, available from https://coronakommissionen.com/wp-content/uploads/2021/10/underlagsrapport-andersson-m-fl-dodlighet-och-sjuklighet-i-covid-19-i-sverige.pdf

Aradhya, S., Brandén, M., Drefahl, S., Obućina, O., Andersson, G., Rostila, M., Mussino, E., & Juárez, S. P. (2020). Lack of acculturation does not explain excess COVID-19 mortality among immigrants. A population-based cohort study. Stockholm Research Reports in Demography. Preprint. https://doi.org/10.17045/sthlmuni.13110365.v2

Arbel, Y., Fialkoff, C., Kerner, A., & Kerner, M. (2022). Do population density, socio-economic ranking and Gini Index of cities influence infection rates from coronavirus? Israel as a case study. *The Annals of regional science*, *68*(1), 181-206.

Born, B., Dietrich, A. M., & Müller, G. J. (2021). The lockdown effect: A counterfactual for Sweden. *Plos one*, *16*(4), e0249732.

Brandén, M., Aradhya, S., Kolk, M., Härkönen, J., Drefahl, S., Malmberg, B., & Mussino, E. (2020). Residential context and COVID-19 mortality among adults aged 70 years and older in Stockholm: a population-based, observational study using individual-level data. *The Lancet healthy longevity*, *1*(2), e80-e88.

Chen, J. T., & Krieger, N. (2021). Revealing the unequal burden of COVID-19 by income, race/ethnicity, and household crowding: US county versus zip code analyses. *Journal of Public Health Management and Practice*, *27*(1), S43-S56.

Calderón-Larrañaga, A., Vetrano, D. L., Rizzuto, D., Bellander, T., Fratiglioni, L., & Dekhtyar, S. (2020). High excess mortality in areas with young and socially vulnerable populations during the COVID-19 outbreak in Stockholm Region, Sweden. *BMJ global health*, *5*(10), e003595.

Dekhtyar, S. (2020). High excess mortality in areas with young and socially vulnerable populations during the COVID-19 outbreak in Stockholm Region, Sweden. *BMJ global health*, *5*(10), e003595.

Carozzi F, Provenzano S, Roth S (2020) Urban density and COVID-19, discussion paper series, IZA Institute of Labor economics, IZA DP No. 13440

Clark, W. A. V. (1991) Residential preferences and neighborhood racial segregation - a test of the schelling segregation model. *Demography*, 28, 1-19.

Clouston, S. A., Rubin, M. S., Phelan, J. C., & Link, B. G. (2016). A social history of disease: Contextualizing the rise and fall of social inequalities in cause-specific mortality. *Demography*, *53*(5), 1631-1656.

Clouston, S. A., Natale, G., & Link, B. G. (2021). Socioeconomic inequalities in the spread of coronavirus-19 in the United States: A examination of the emergence of social inequalities. *Social science & medicine*, *268*, 113554.

Drefahl, S., Wallace, M., Mussino, E., Aradhya, S., Kolk, M., Brandén, M., Malmberg, B., & Andersson, G. (2020). A population-based cohort study of socio-demographic risk factors for COVID-19 deaths in Sweden. *Nature communications*, *11*(1), 1-7.

Dyrdak, R., & Albert, J. 2021. Utredning av den initiala smittspridningen av SARS-CoV-2 till och inom Sverige. Underlagsrapport till SOU 2021:89 Sverige under pandemin Stockholm 2021. Submission to Swedish Coronakommissionen, available from https://coronakommissionen.com/wp-content/uploads/2021/10/underlagsrapport-dyrdak-albert-utredning-av-den-initiala-smittspridningen-av-sars-cov-2-till-och-inom-sverige.pdf

Florida, R., & Mellander, C. (2021). The geography of COVID-19 in Sweden. *The Annals of Regional Science*, 1-26.

Fonseca-Rodríguez, O., Gustafsson, P. E., San Sebastián, M., & Connolly, A. M. F. (2021). Spatial clustering and contextual factors associated with hospitalisation and deaths due to COVID-19 in Sweden: a geospatial nationwide ecological study. *BMJ Global Health*, *6*(7), e006247.

Green, M. A., García-Fiñana, M., Barr, B., Burnside, G., Cheyne, C. P., Hughes, D., Ashton, M., Sheard, S. & Buchan, I. E. (2021). Evaluating social and spatial inequalities of large scale rapid lateral flow SARS-CoV-2 antigen testing in COVID-19 management: An observational study of Liverpool, UK (November 2020 to January 2021). *The Lancet Regional Health-Europe*, *6*, 100107.

Hennerdal, P. (2019). Geocontext [Python]. https://github.com/PonHen/geocontext

Kamis, C., Stolte, A., West, J. S., Fishman, S. H., Brown, T., Brown, T., & Farmer, H. R. (2021). Overcrowding and COVID-19 mortality across US counties: Are disparities growing over time? *SSM-population health*, *15*, 100845.

Karimi, A. 2021. Heterogeneity in mortality, morbidity, and health behaviors: Covid-19 and beyond. Underlagsrapport till SOU 2021:89 Sverige under pandemin Stockholm 2021. Submission to Swedish Coronakommissionen, available from https://coronakommissionen.com/wp-content/uploads/2021/10/underlagsrapport-karimi-heterogeneity-in-mortality-morbidity-and-health-behaviors.pdf

Kwon, S., Joshi, A. D., Lo, C. H., Drew, D. A., Nguyen, L. H., Guo, C. G., Ma, W., Mehta, R.S., Shebl, F.M., Warner, E.T. and Astley, C.M (2021). Association of social distancing and face mask use with risk of COVID-19. *Nature Communications*, *12*(1), 1-10.

Leung, C. (2020). Risk factors for predicting mortality in elderly patients with COVID-19: a review of clinical data in China. *Mechanisms of ageing and development*, *188*, 111255.

de Lusignan, S., J. Dorward, A. Correa, N. Jones, O. Akinyemi, G. Amirthalingam, N. Andrews, R. By- ford, G. Dabrera, A. Elliot, et al. (2020). Risk factors for sars-cov-2 among patients in the oxford royal college of general practitioners' research and surveillance centre primary care network: a cross- sectional study. *The Lancet Infectious Diseases 20*(9), 1034–1042.

Meurisse, M., Lajot, A., Devleesschauwer, B., Van Cauteren, D., Van Oyen, H., Van den Borre, L., & Brondeel, R. (2022). The association between area deprivation and COVID-19 incidence: a municipality-level spatio-temporal study in Belgium, 2020–2021. *Archives of Public Health*, *80*(1), 1-10.

Madhav, KC., Oral, E., Straif-Bourgeois, S., Rung, A. L., & Peters, E. S. (2020). The effect of area deprivation on COVID-19 risk in Louisiana. *PLoS One*, *15*(12), e0243028.

Nilsson, P. 2021. Betydelsen av anställda och anhörigas sociala nätverk för smittspridning av COVID-19 på äldreboenden under 2020. Underlagsrapport till SOU 2021:89 Sverige under pandemin Stockholm 2021.

Openshaw, S. (1984). The modifiable areal unit problem, CATMOG 38. *Geo Abstracts*, Norwich, UK.

Östh, J., B. Malmberg, & E. K. Andersson (2014). Analysing segregation using individualised neighbourhoods. In *Social-spatial segregation: Concepts, processes, and outcomes* (135–162). Policy Press.

Rostila, M., Cederström, A., Wallace, M., Brandén, M., Malmberg, B., & Andersson, G. (2021). Disparities in COVID-19 deaths by country of birth in Stockholm, Sweden: A total population-based cohort study. *American Journal of Epidemiology*.

Sandhu, A., Korzeniewski, S. J., Polistico, J., Pinnamaneni, H., Reddy, S. N., Oudeif, A., Meyers, J., Sidhu, N., Levy, P., Samavati, L., Badr, MS., Sobel, JD., Sherwin, R., & Chopra, T. (2021). Elevated COVID19 mortality risk in Detroit area hospitals among patients from census tracts with extreme socioeconomic vulnerability. *EClinicalMedicine*, *34*, 100814.

Sigurjónsdóttir, H. R., Sigvardsson, D., & Oliveira e Costa, S. (2021). *Who is left behind? The impact of place on the possibility to follow Covid-19 restrictions*. Nordic Council of Ministers.

Appendix A – regression results

Table A1 Poisson regression with robust standard errors. Dependent variable is admission to ICU

VARIABLE	RR	(LCL-UCL)
(INTERCEPT)	0	(0,00-0,00)
SEX		
MEN	1	ref
WOMEN	0,42	(0,40-0,44)
AGE		
[0-29)	1	ref
[20,30)	3,72	(2,79-2,79)
[30,35)	5,37	(4,00-4,00)
[35,40)	6,24	(4,67-4,67)
[40,45)	11,1	(8,45-8,45)
[45,50)	20,54	(15,81-15,81)
[50,55)	30,92	(23,92-23,92)
[55,60)	43,25	(33,57-33,57)
[60,65)	60,09	(46,70-46,70)
[65,70)	70,45	(54,76-54,76)
[70,75)	73,28	(56,95-56,95)
[75,80)	66,78	(51,74-51,74)
[80,85)	45,41	(34,75-34,75)
[85,90)	18,66	(13,43-13,43)
[90,100)	3,74	(1,92-1,92)
[100,115]	0	(0,00-0,00)
REGION OF ORIGIN		
SWEDEN	1	ref
EU/EEA	1,37	(1,26-1,50)
NON-EU/EEA	2,95	(2,78-3,13)
CLUSTER		
RT_DIV	1,57	(1,43-1,71)
RT_ADJ	1,17	(1,06-1,30)
RT_WC	1,26	(1,14-1,39)
R_HOM	1	ref
R_BOR	0,75	(0,55-1,02)
U_DIV (BUFFER)	1,45	(1,31-1,61)
U_DIV (CORE)	1,82	(1,62-2,04)
U_ADJ	0,89	(0,76-1,03)
U_HOM	1,41	(1,25-1,58)
U_ACA	1,04	(0,94-1,14)
U_ELI	0,87	(0,77-0,98)

<u>DESO POPULATION DENSITY QUINTILES</u>		
(0-139]	1	ref
(139,487]	1,06	(0,97-1,15)
(487,1690]	1,26	(1,16-1,37)
(1690,5170]	1,29	(1,19-1,41)
(5170,57700]	1,51	(1,37-1,65)

Table A2 Poisson regression with robust standard errors and interaction. Dependent variable is admission to ICU for COVID-19

VARIABLE	RR	LCL-UCL)
(INTERCEPT)	0	(0,00-0,00)
SEX		
MEN	1	ref
WOMEN	0,42	(0,40-0,44)
AGE		
[0-29)	1	ref
[20,30)	3,72	(2,79-4,96)
[30,35)	5,37	(4,00-7,20)
[35,40)	6,24	(4,67-8,33)
[40,45)	11,1	(8,44-14,59)
[45,50)	20,54	(15,81-26,67)
[50,55)	30,92	(23,92-39,97)
[55,60)	43,25	(33,56-55,75)
[60,65)	60,1	(46,70-77,35)
[65,70)	70,46	(54,76-90,68)
[70,75)	73,3	(56,95-94,34)
[75,80)	66,79	(51,75-86,21)
[80,85)	45,42	(34,75-59,37)
[85,90)	18,66	(13,42-25,95)
[90,100)	3,74	(1,92-7,29)
[100,115]	0	(0,00-0,00)
REGION OF ORIGIN		
SWEDEN	1	ref
EU/EEA	1,37	(1,26-1,50)
NON-EU/EEA	2,95	(2,78-3,13)
DESO POPULATION DENSITY QUINTILES		
(0-139]	1	ref
(139,487]	1,06	(0,97-1,15)
(487,1690]	1,26	(1,16-1,37)
(1690,5170]	1,29	(1,19-1,41)
(5170,57700]	1,51	(1,37-1,65)
CLUSTER_WAVES		
RT_DIV # WAVE 1	1,86	(1,59-2,17)
RT_ADJ # WAVE 1	1,38	(1,16-1,65)
RT_WC # WAVE 1	1,23	(1,02-1,48)
R_HOM # WAVE 1	1	ref
R_BOR # WAVE 1	0,5	(0,25-1,00)

U_DIV (BUFFER) # WAVE 1	2,06	(1,74-2,43)
U_DIV (CORE) # WAVE 1	3,23	(2,72-3,84)
U_ADJ # WAVE 1	0,96	(0,72-1,28)
U_HOM # WAVE 1	1,88	(1,55-2,30)
U_ACA # WAVE 1	1,28	(1,04-1,54)
U_ELI # WAVE 1	1,26	(1,04-1,54)
RT_DIV # WAVE 2	1,33	(1,15-1,53)
RT_ADJ # WAVE 2	1,28	(1,06-1,56)
RT_WC # WAVE 2	1,65	(1,36-2,01)
R_HOM # WAVE 2	1,63	(1,39-1,91)
R_BOR # WAVE 2	3,5	(1,54-7,92)
U_DIV (BUFFER) # WAVE 2	1,09	(0,92-1,30)
U_DIV (CORE) # WAVE 2	0,79	(0,65-0,96)
U_ADJ # WAVE 2	1,57	(1,10-2,23)
U_HOM # WAVE 2	0,99	(0,78-1,27)
U_ACA # WAVE 2	1,28	(1,10-1,48)
U_ELI # WAVE 2	1	(0,79-1,27)
RT_DIV # WAVE 3	2,07	(1,80-2,37)
RT_ADJ # WAVE 3	2,16	(1,80-2,60)
RT_WC # WAVE 3	2,83	(2,36-3,40)
R_HOM # WAVE 3	2,72	(2,34-3,17)
R_BOR # WAVE 3	3,69	(1,56-8,76)
U_DIV (BUFFER) # WAVE 3	1,46	(1,22-1,73)
U_DIV (CORE) # WAVE 3	0,89	(0,73-1,09)
U_ADJ # WAVE 3	2,28	(1,62-3,23)
U_HOM # WAVE 3	1,86	(1,48-2,34)
U_ACA # WAVE 3	1,9	(1,65-2,19)
U_ELI # WAVE 3	1,47	(1,16-1,85)

Appendix B – geographical distribution of neighbourhood cluster types

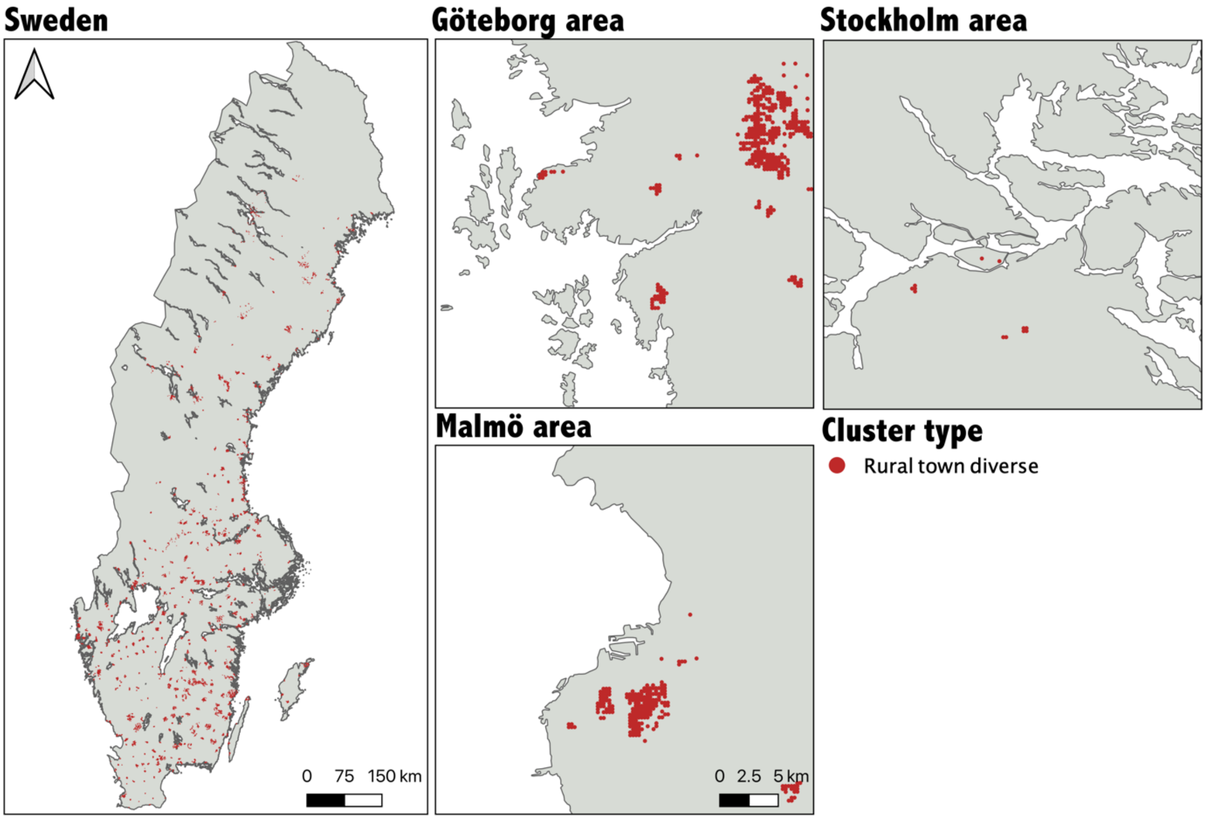

Figure 12 Rural town diverse neighbourhood cluster type (RT_DIV).

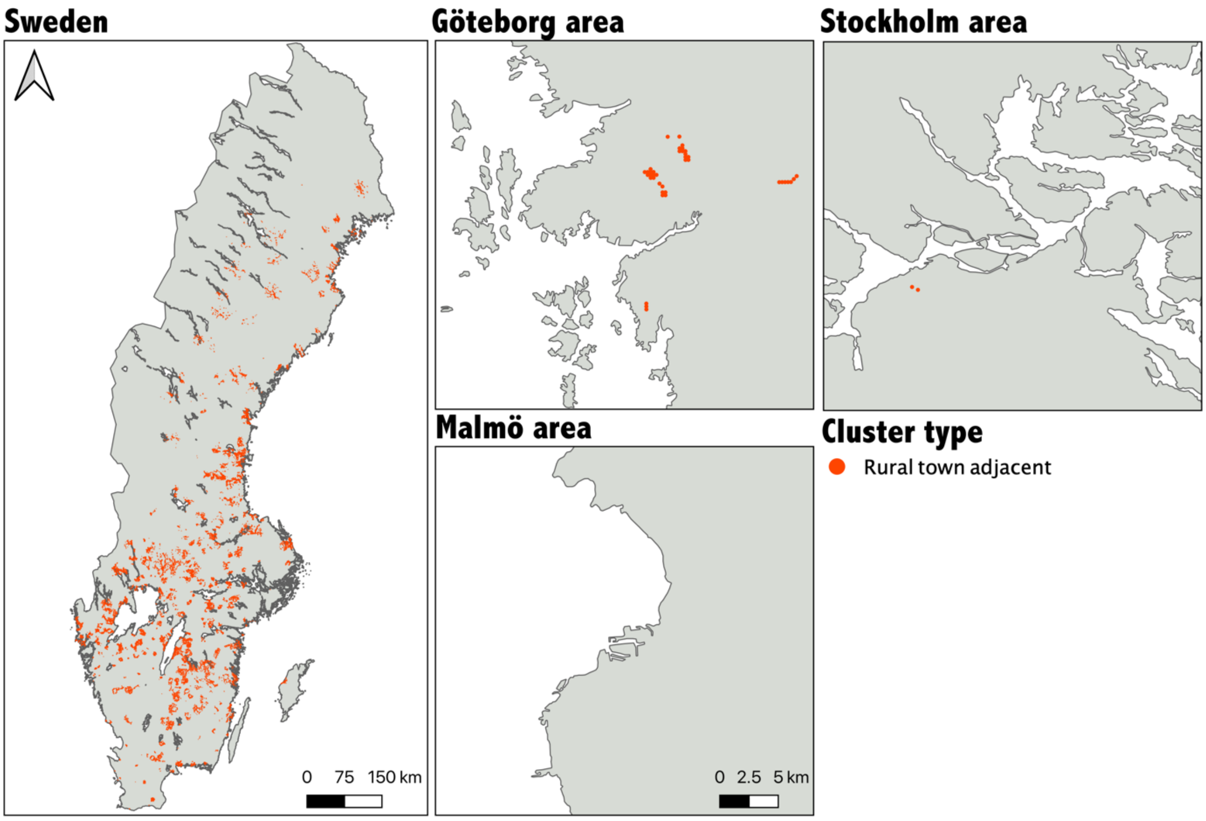

Figure 13 Rural adjacent neighbourhood cluster type (RT_ADJ).

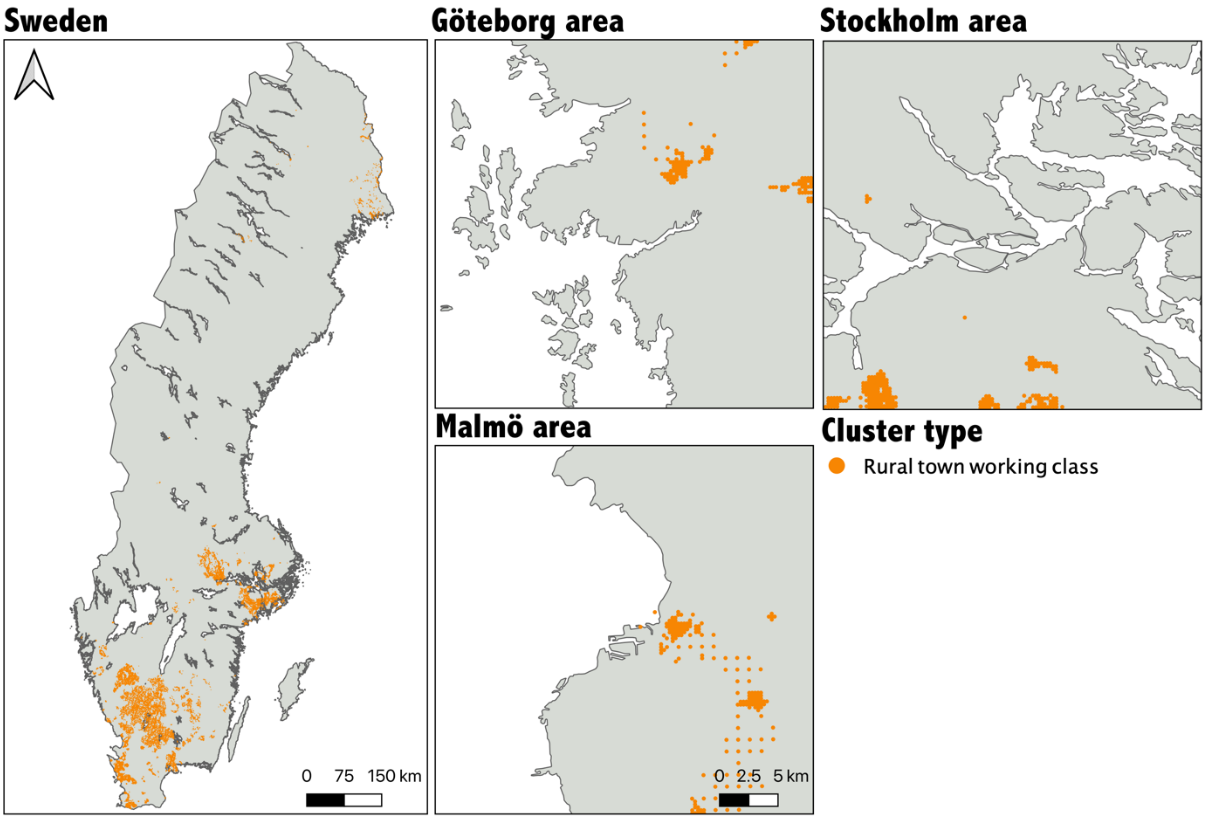

Figure 14 Rural town working class neighbourhood cluster type (RT_WC).

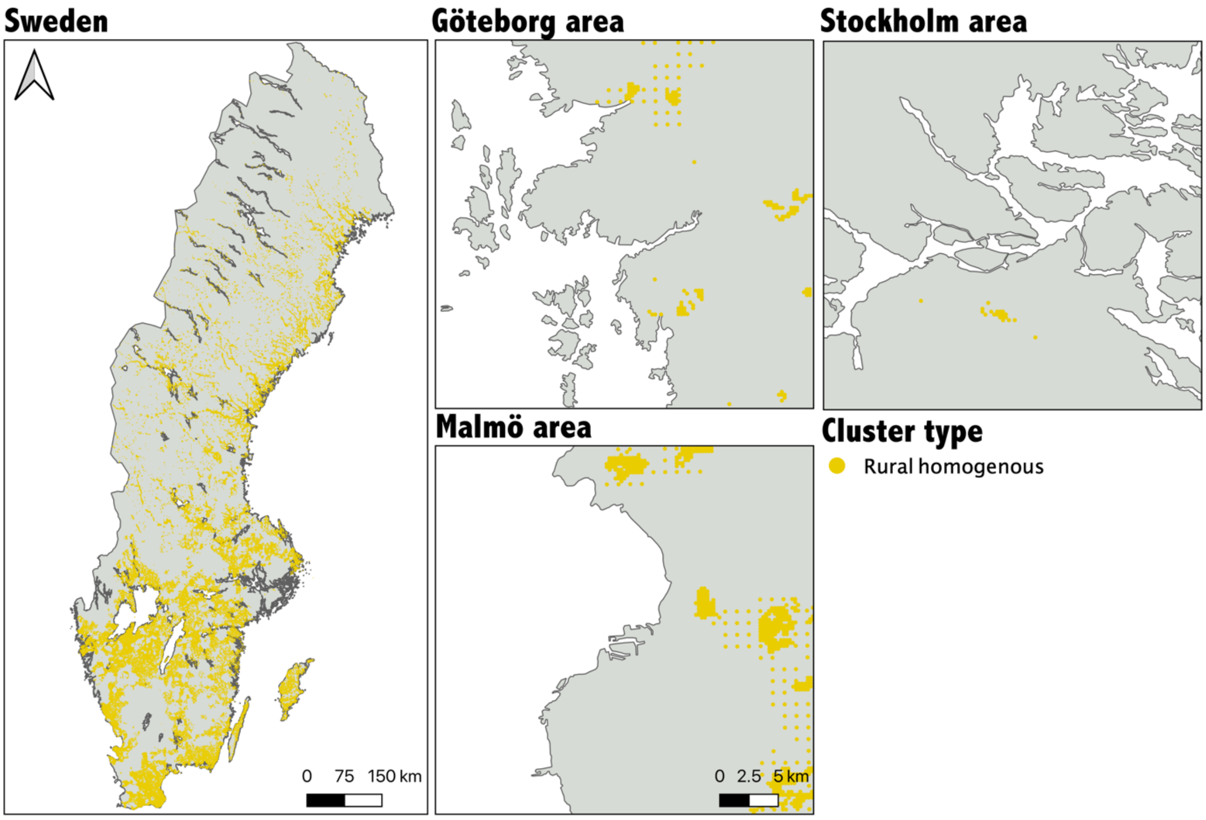

Figure 15 Rural homogenous neighbourhood cluster type (R_HOM).

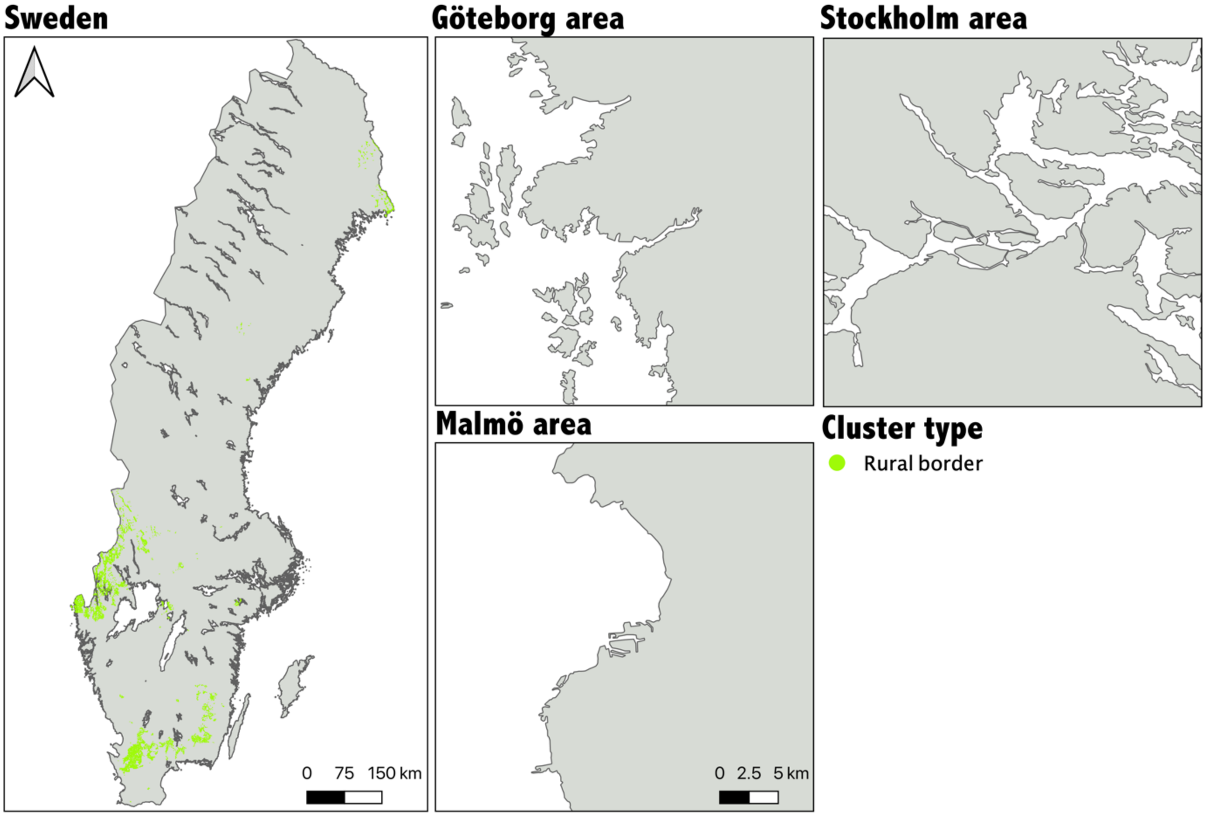

Figure 16 Rural border neighbourhood cluster type (R_BOR).

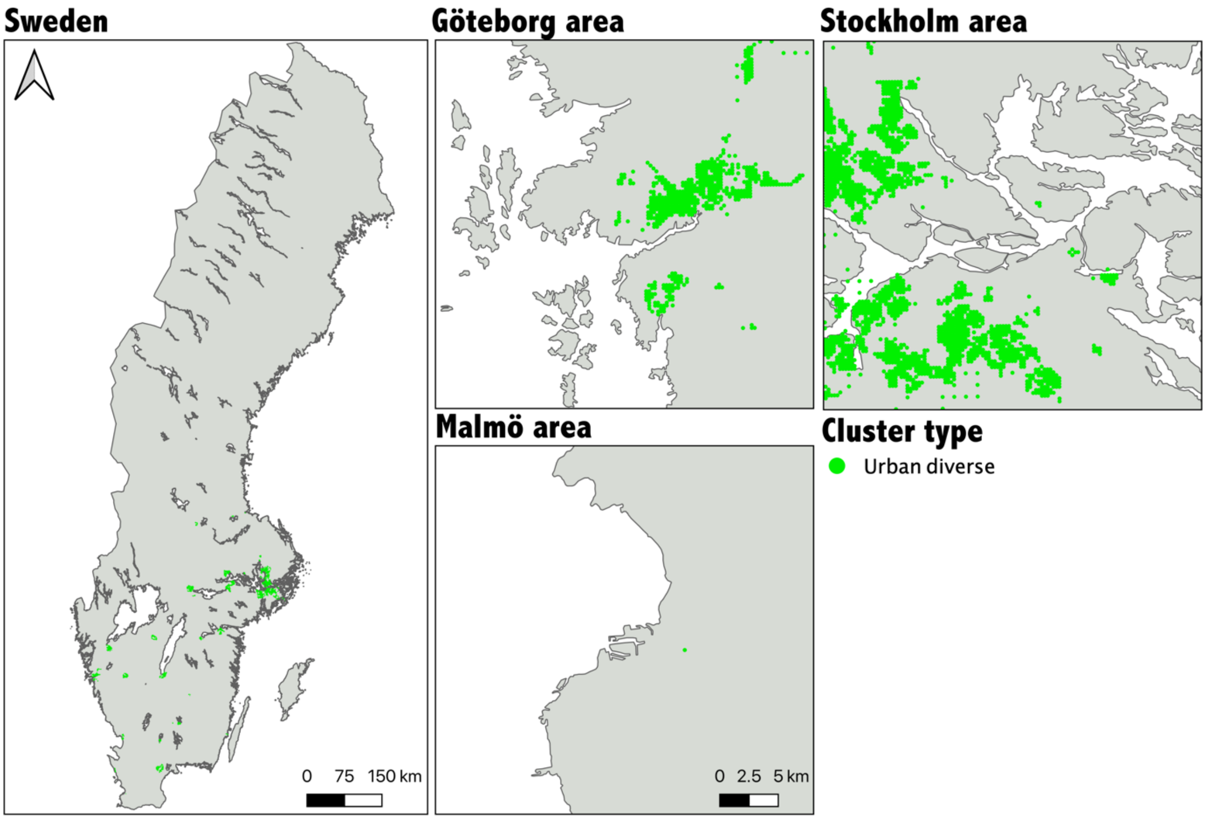

Figure 17 Urban diverse (core and buffer) neighbourhood cluster type (U_DIV).

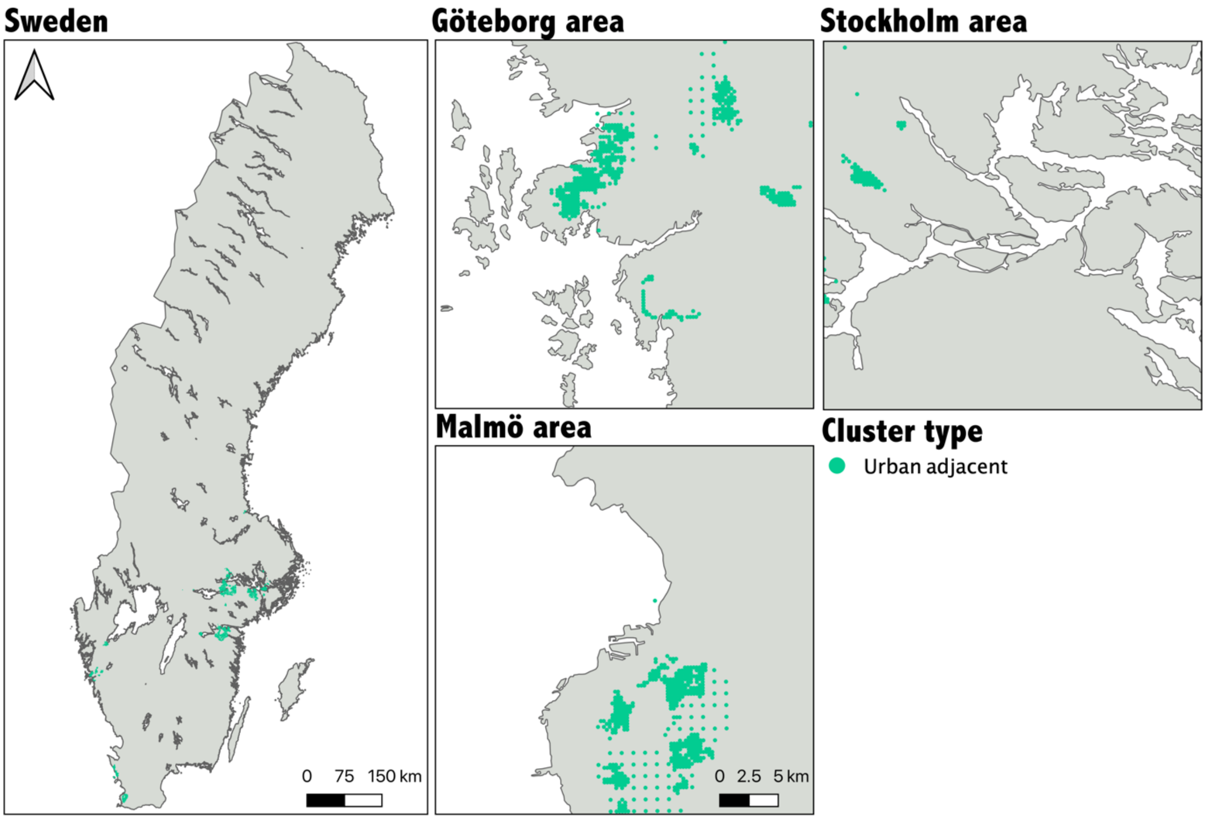

Figure 18 Urban adjacent neighbourhood cluster type (U_ADJ).

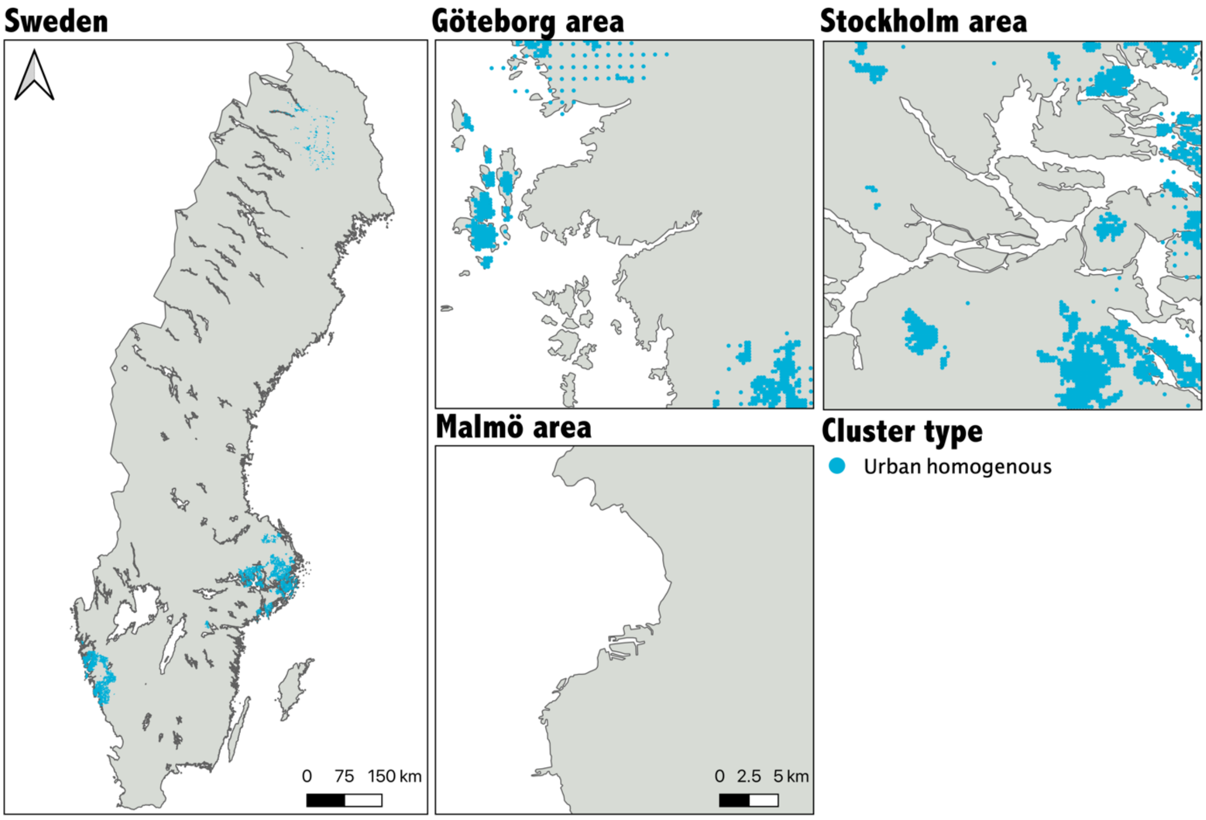

Figure 19 Urban homogenous neighbourhood cluster type (U_HOM).

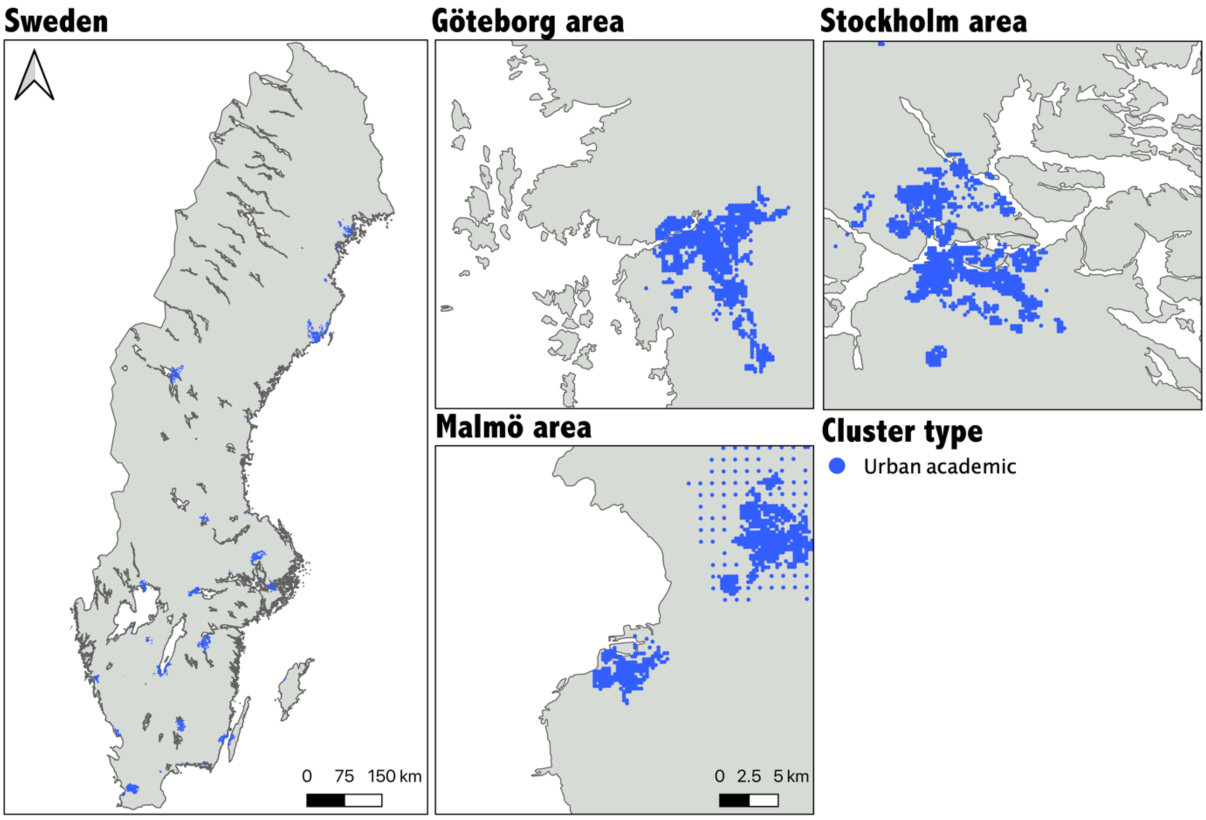

Figure 20 Urban academic neighbourhood cluster type (U_ACA).

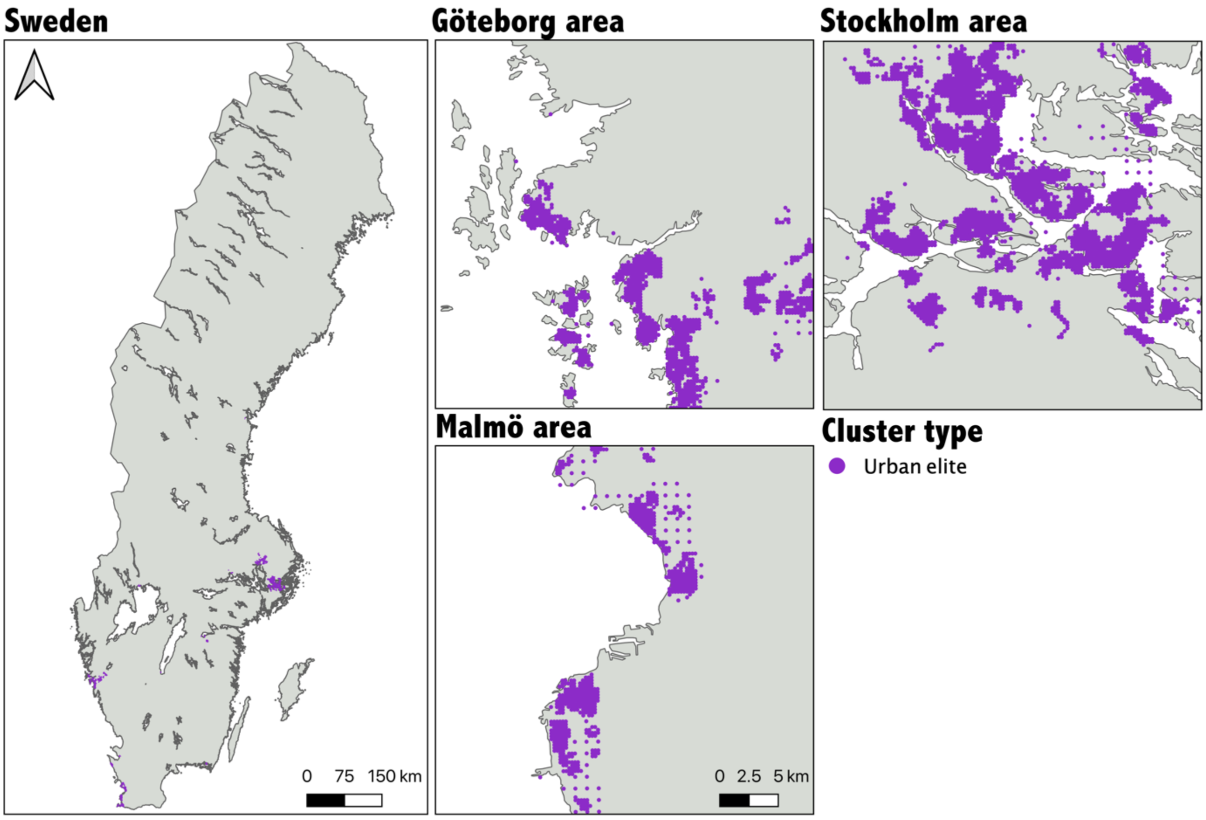

Figure 21 Urban elite neighbourhood cluster type (U_ELI).

Appendix C – details on construction the neighbourhood type classification

To create neighbourhood typology, we selected 7 variables and for each inhabited grid cell we calculated shares among the closest neighbours at 9 scales ranging from 200 to 51200 closest neighbours. Using 7 indicators at 9 scales (k = 200, 400, 800, 1600, 3200, 6400, 12800, 25600 and 51200) gave a total of 63 measures of neighborhood context that can be used to classify residential areas using cluster analysis. However, since many of these variables will be highly correlated, we used factor analysis that compresses the 63 original indicators to 8 orthogonal factors before proceed- ing with the cluster analysis (for a similar method see Clark et al. 2015). The reason for initial factor analysis is that given the strong correlation between indicators across different scales it is possible to capture most of the variation in neighborhood composition using a small number of factors. Moreover, given the number of our observation, reducing the number of measures of neighborhood context from 63 to 8 makes clustering algorithms more computationally manageable. The factor analysis was based on correlations and the number of factors was selected based on them having eigenvalues higher than one. The factors were rotated using the varimax method. Figure 22 to Figure 29 show the results of the factor analysis. The panels in these figures show the loading of the different factors for each indicator. The first factor represents the *large-scale elite context* with high factor loadings for tertiary education and high income. The second factor repre- sents *small scale disadvantage and diversity* because of high values for social assistance and non-EU/EFTA immigrants at lower k-levels (closer neighbors). The third factor represents *large-scale diversity* (high factor loadings at all k-levels). It operates at larger scale than *small scale disadvantage and diversity* and tends to have low values in mid-Sweden. The fourth factor is a mirror reflection of factor two but at larger k-scales, suggesting that is represents areas *adjacent to disadvantage and diversity*. Factors five and six signal *adjacent to small-scale disadvantage* and *adjacent to small- scale diversity* with high values for disadvantage and diversity for medium range k-values. The main difference is that for factor five we also observe high factor loadings for non-EU/EFTA immigration, while for factor six it is less pronounced. The seventh factor represents *large- scale poverty;* we also note that for this factor disadvantage appears to be un- related to immigration. The eighth factor represent *large-scale high income* with high levels of top income earners at all k-scales that, unlike for *large- scale elite context*, are not coupled with high educational attainment.

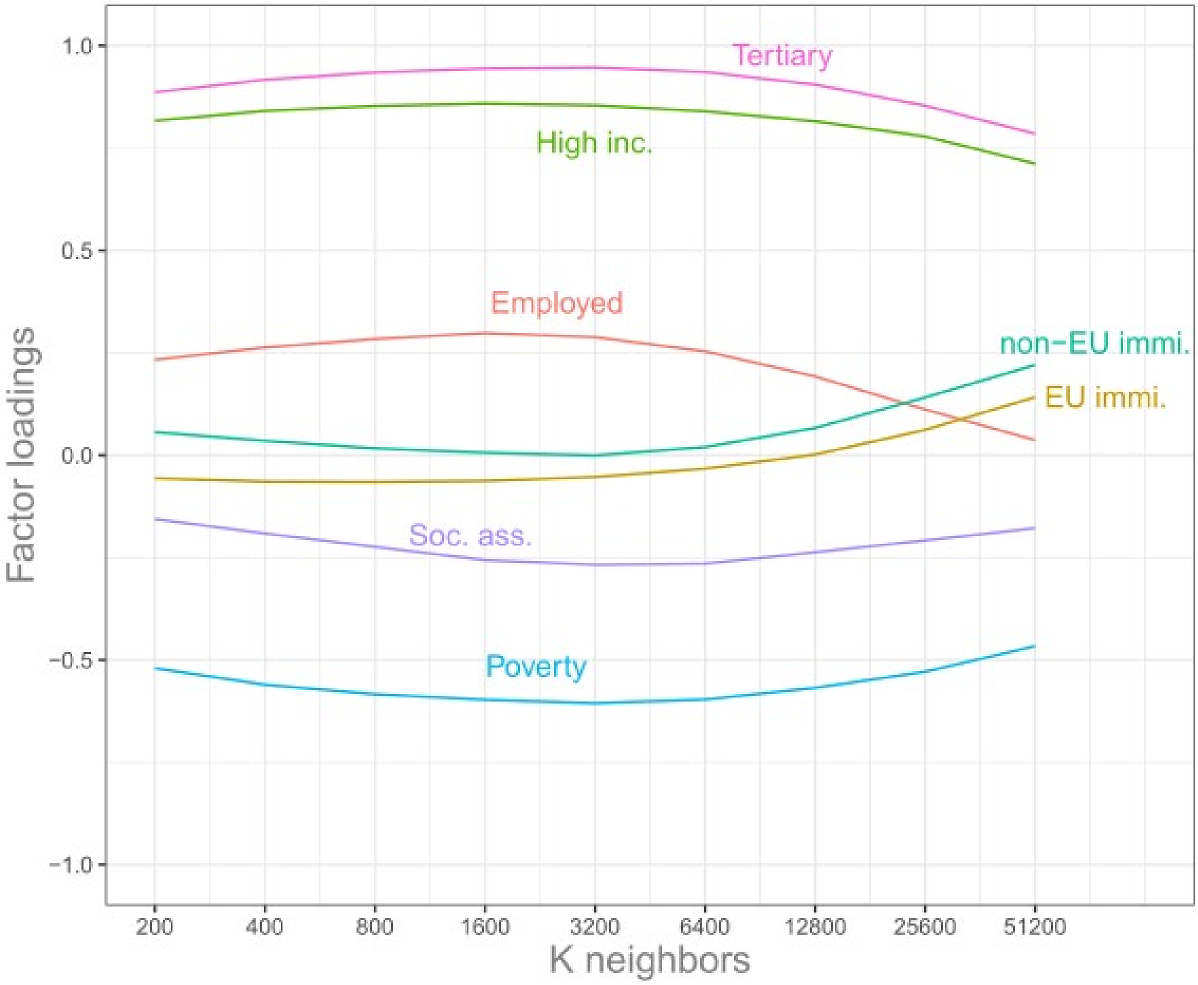

Figure 22 Large scale elite

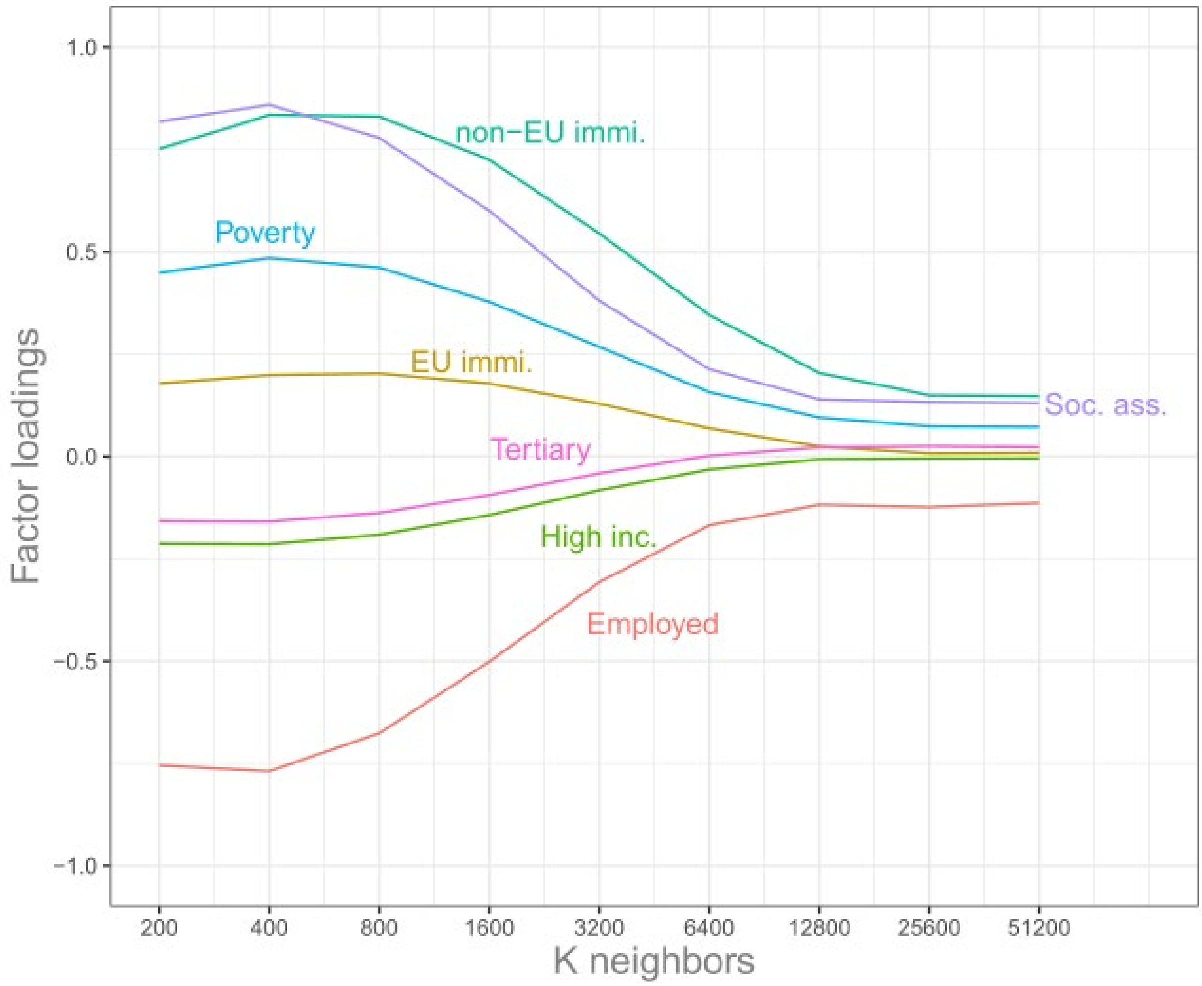

Figure 23 Small scale disadvantage and diversity

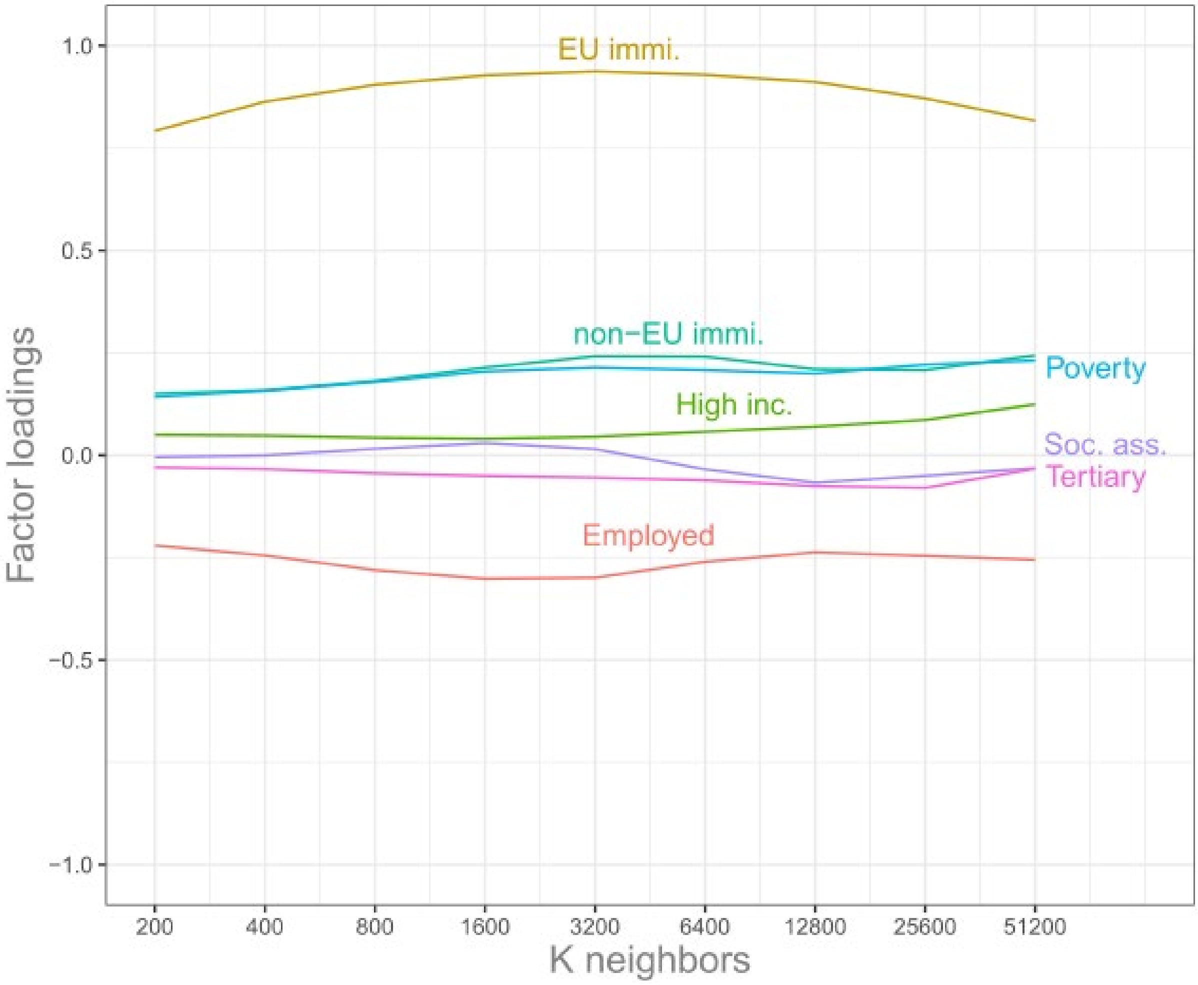

Figure 24 Large scale diversity

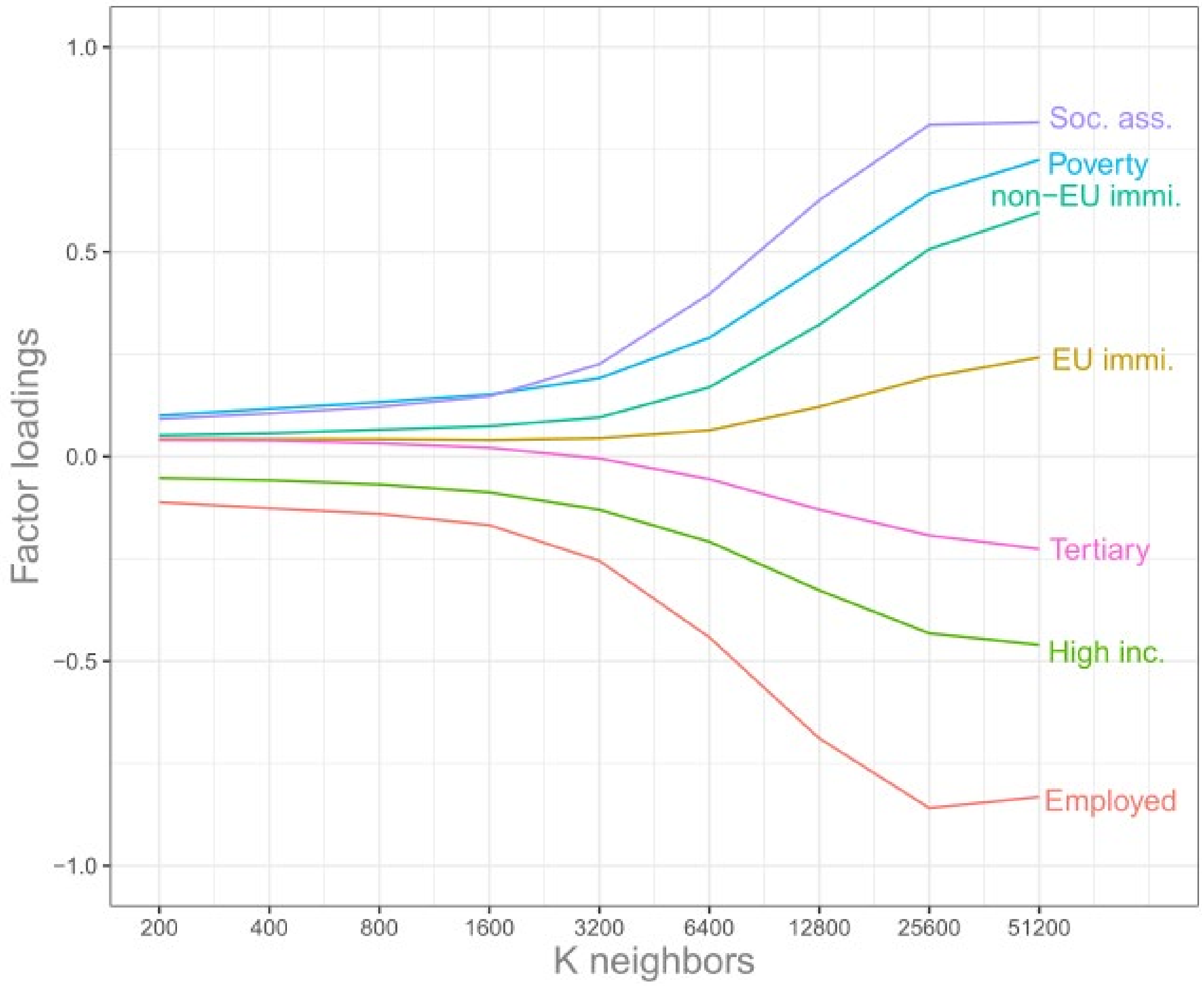

Figure 25 Adjacent to disadvantage and diversity

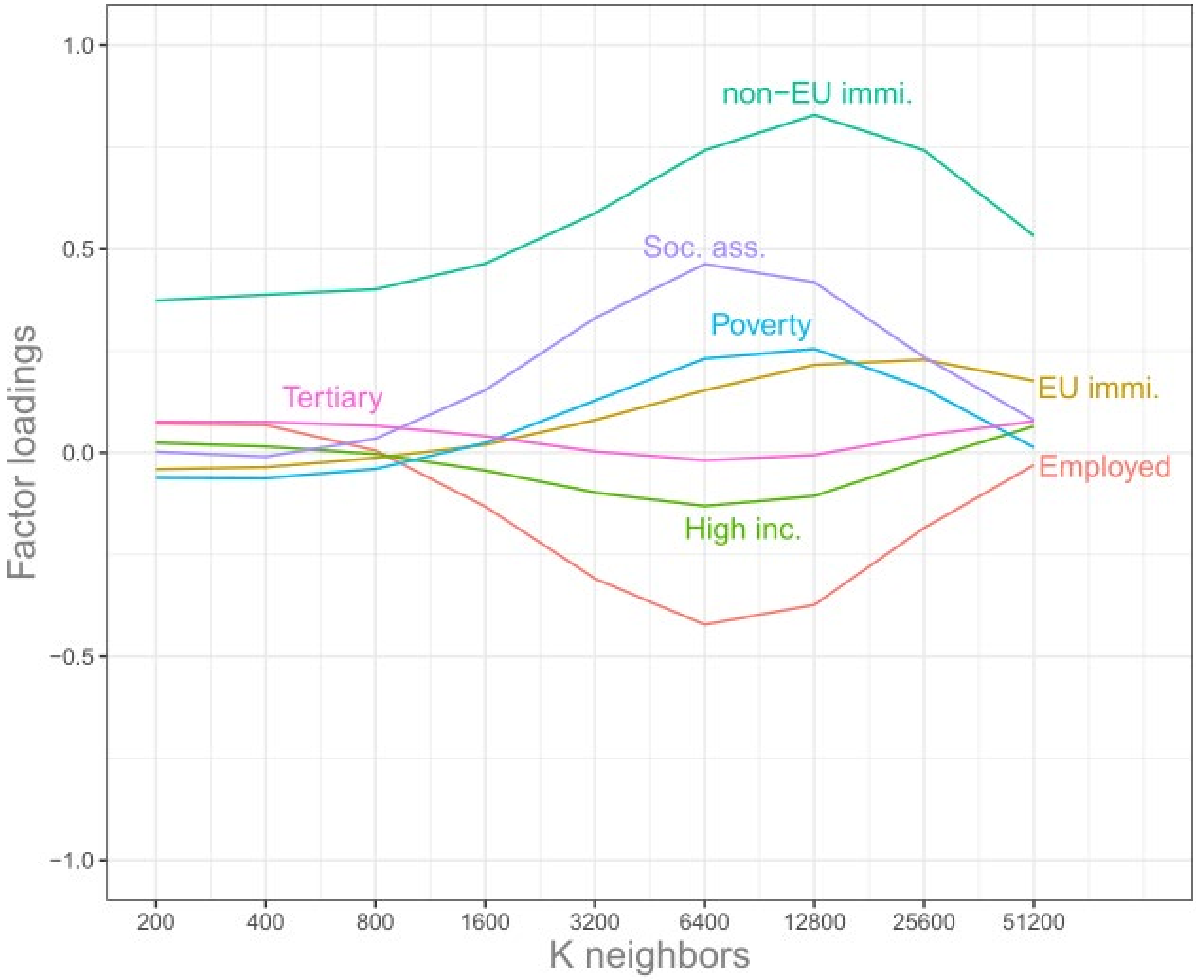

Figure 26 Adjacent to small scale diversity

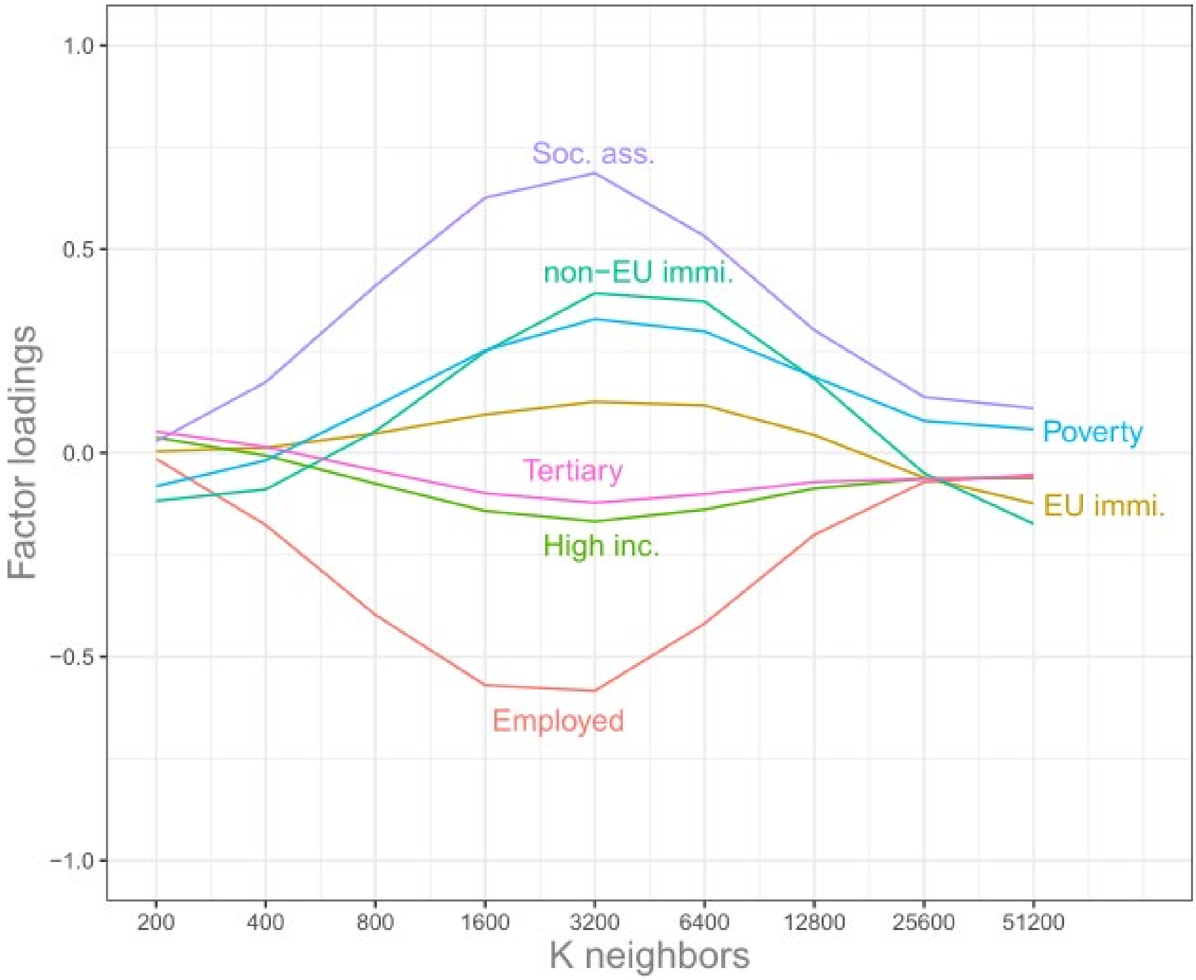

Figure 27 Adjacent to small scale disadvantage

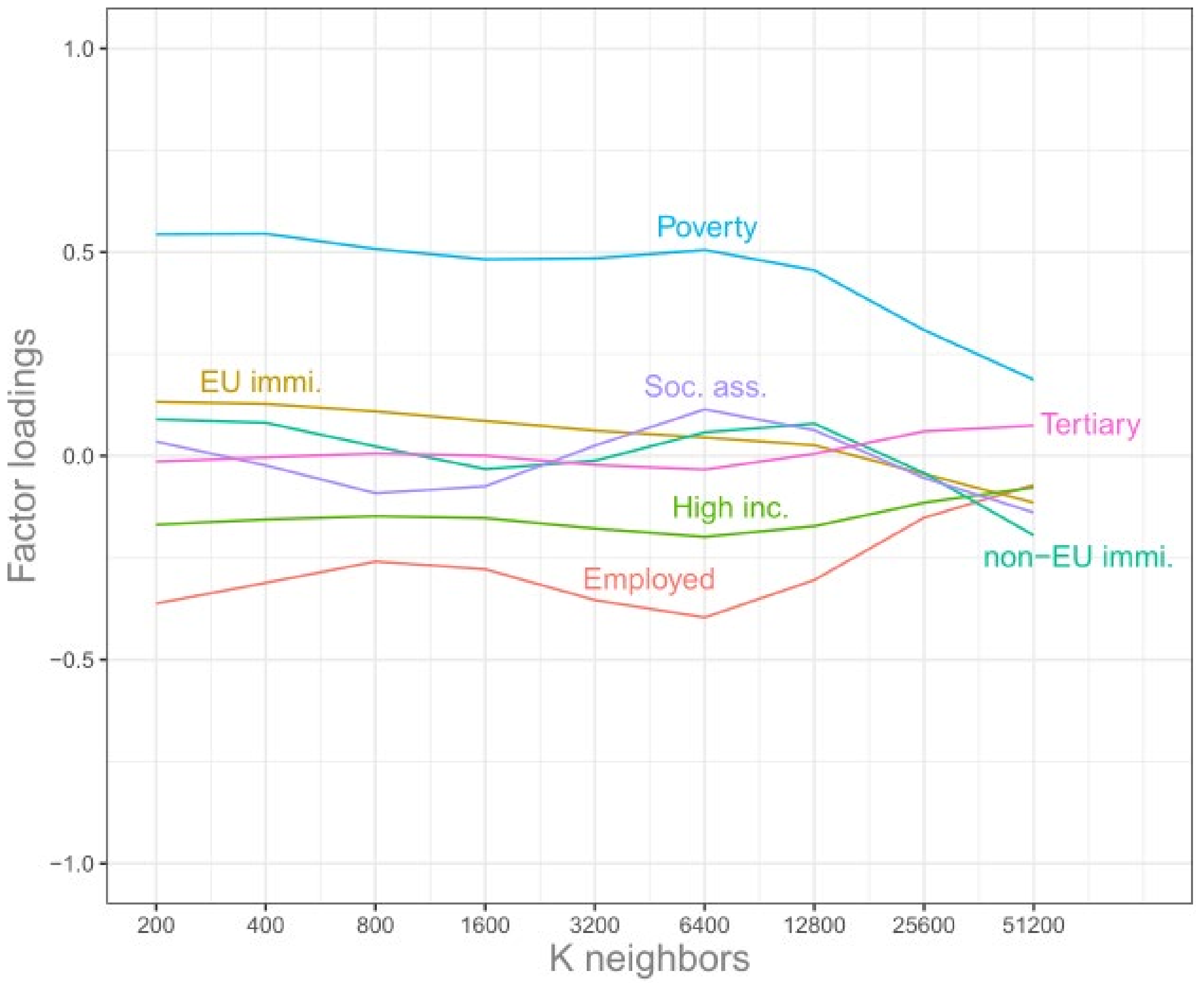

Figure 28 Large scale poverty

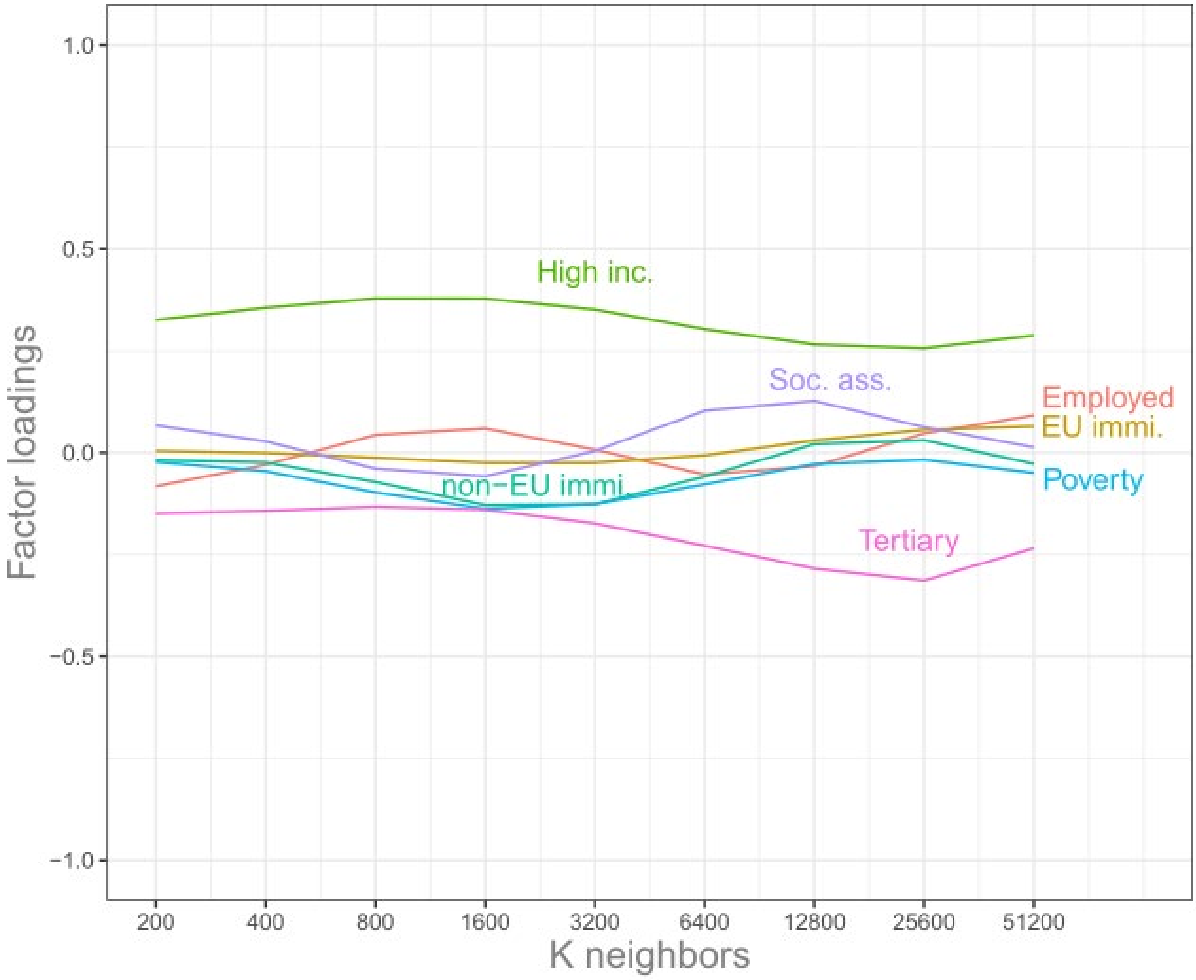

Figure 29 Large scale high income

Stockholm Research Reports in Demography
Stockholm University,
106 91 Stockholm,
Sweden
www.su.se | info@su.se | ISSN 2002-617X

Stockholm University
Demography Unit