Working Paper Series
Department of Economics
University of Verona

Inequality of Opportunity in Sweden:A Spatial Perspective

Umut Turk, John Östh

WP Number: 9 May 2017

ISSN: 2036-2919 (paper), 2036-4679 (online)

Inequality of Opportunity in Sweden:A Spatial Perspective *

Umut TÜRK[†]& John ÖSTH[‡]

Abstract

This paper investigates the spatial dimensions of inequality of opportunity by integrating parental neighbourhood characteristics as exogenous factors influencing the life chances of individuals. We construct egocentric neighbourhoods, where contextual variables are quantified by an approach based on k nearest neighbours. The analyses are carried out with multilevel models departing significantly from previous studies where solely OLS regressions were employed. Using Swedish longitudinal register data, we show that the parental neighbourhood is highly influential in educational inequality of opportunity and remains so for earnings inequality of opportunity even years after exposure.

Keywords neighbourhood effects, Equality of Opportunity

*This paper benefited from the insightful comments and suggestions of the participants of the Winter School on Space, Mobility and Opportunities in Canazei, Italy and Seminar Series in Verona, Italy

[†]Economics Department, University of Verona, Verona Italy. Email address: umut.turk@univr.it

[‡]Department of Social and Economic Geography, Uppsala University, Uppsala Sweden

1 Introduction

In recent years, the equality of opportunity (hereafter EOp) concept has often been mentioned in discussions concerning distributional disparities (Cities and Social Equity, 2009; Urban Equity of Life, 2014; Handbook of Income Distribution, 2015). It is almost a universally accepted principle developed recently by Roemer (1998) and already with numerous empirical applications (see Ramos & Van de Gaer, 2012; Roemer & Trannoy, 2013; Brunori et al., 2013, for reviews). In this framework the overall inequality observed in different spheres of social life is decomposed into ethically acceptable (fair) and offensive (unfair) components. So-called *circumstances* representing unfair sources of inequality, are predetermined and beyond people's control such as gender, race and family background (Roemer, 1998; Roemer & Trannoy, 2013). On the other hand, so-called *effort* comprises fair (acceptable) sources of inequality for which individuals are held responsible. Inequality that is due to circumstances is defined as inequality of opportunity (hereafter IOp).

The IOp is a product of several underlying unfair inequalities, such as inequality due to differences in social treatment, inequality of access to basic opportunities, inequality due to exogenous genetic factors and inequality due to parental resources and location (de Barros, 2009). A number of empirical studies seek to disentangle these underlying inequalities in opportunities through a set of circumstance variables. To assess IOp accurately, analyses thus need to take a comprehensive approach on the circumstance variables employed. Most existing studies have been limited to accounting for inequalities due to parental resources and social treatment, with gender, race, parental income and education as typical predictors of circumstances. So far, however, there has been little discussion of the spatial sources of IOp and none for inequalities due to parental location.

It has already been suggested that residential location may generate unfair inequalities especially for children (Ross et al., 2002; de Barros et al., 2008). However, the empirical studies that have sought to evaluate IOp, include geography either as a reference to a general Urban/Rural division of birthplace (Ferreira et al., 2010), or as large administrative units, for instance regions (Peragine & Serlenga, 2008; Checchi et al., 2010; Singh, 2012). Leaving aside the problems of robustness due to spatial effects in place such as the modifiable areal unit problem (MAUP) (Openshaw, 1984), on such a scale geography does not represent the residential characteristics to which individuals are exposed. Therefore, a more specific characterization of spatial patterns that communicates the current and past residential environment of individuals and interaction among them must be included in analyses.

This paper focuses on the role of people's parental neighbourhood characteristics as a source of IOp in education and income and their own neighbourhood characteristics as a source of legitimate inequality in income distribution. There is an extensive body of literature investigating the neighbourhood effects on various outcomes of individuals. The literature offers empirical evidence on the link between neighbourhoods and the several life chances of residents. Most previous findings are relevant to this study, for example, neighbourhood effects on child outcomes (see Leventhal & Brooks-Gunn, 2000, for a review), on labour market and economic outcomes (see Vartanian, 1999) and spatial mismatch implications (for a review, see Kain, 1992), health outcomes such as psychological wellbeing (see Ludwig et al., 2013), behavioural outcomes such as the likelihood of committing a crime and drug-alcohol consumption (Case & Katz, 1991). However, it may be problematic to attribute causal relations to how these effects take place (Buck, 2001). Considerable effort has been put into providing a proper definition of neighbourhoods, and causal relations with their observed effects (see Ellen & Turner, 1997, for a review). Galster (2001) argues that the neighbourhood is a multidimensional phenomenon, in which four actors (households, businesses, property holders and local government) act both as consumers and producers in shaping the structural, demographic and social-interactive characteristics of neighbourhoods (Galster, 2001). As consumers, residents are exposed to institutional mechanisms, to peers and networks, to environmental aspects (i.e. polluted air) and to offered accessibility of opportunities (Sharkey & Faber, 2014). In the light of these suggestions, we treat neighbourhoods as the environment surrounding residents where its scale is based on the interaction possibilities between individuals.

The aim of the present work is twofold. The first part of the paper is devoted to the analysis of neighbourhood effects. We use a multilevel modelling strategy to disentangle the influence of circumstances in relation to parental neighbourhood on educational attainment when living with parents and additionally parental and own-neighbourhood characteristics influencing disposable income when living independently of parents. Using the longitudinal register database from Sweden, we focus on the whole population of 1985 cohort. The database provides family background variables such as parental education, employment, marital status and national origin, and provides information on individuals disposable incomes and compulsory exam marks that are taken as dependent variables. Geographical information on residents is used to construct bespoke neighbourhoods with the aim of channeling several characteristics of residential locations based on a k-nearest neighbours approach (Östh et al., 2015). In addition, we include a measure of negative environment surrounding the parental neighbourhood derived from the Corine (coordination of information on the environment) database and a measure of job/housing balance in own-neighbourhood. In the second part of the paper, we construct a model to perform a comprehensive investigation of IOp with particular attention to the spatial sources of inequality. As in Bourguignon et al. (2007) and Ferreira & Gignoux (2011) we formulate a situation where the within-group inequalities are eliminated, so that the overall inequality both in educational attainment and disposable income is decomposed into circumstance (IOp) and effort components. As it is the foremost aim of this study, we also decompose circumstances and effort into spatial and aspatial components.

The results of the model show that the educational inequality of opportunity is 0.1137 as measured by mean logarithmic deviation (MLD) and 0.1528 as measured by Gini index where 36.94% of IOp (MLD measure) is due to spatial circumstances (i.e. parental neighbourhood). While income IOp is 0.0106 as measured by MLD and 0.0695 as measured by the Gini, where 16.66% of IOp (MLD measure) is attributable to spatial circumstances (i.e. parental neighbourhood) and 1.95% of fair inequality is attributable to spatial effort (i.e. own-neighbourhood).

This study makes several contributions to the existing research: i) via the comprehensive register data we cover a large set of circumstance variables, often not possible due to data restrictions; ii) adapting several bespoke neighbourhood characteristics provides robust, externally valid and policy-oriented identification of the spatial factors as affecting opportunities; iii) the measure of inequality of opportunity is associated with parental neighbourhood characteristics for the first time in this study, and through a multilevel modeling strategy the results are robust with respect to previous studies where problems such as spatial autocorrelation have been ignored.

2 Previous Work

Despite the increasing number of empirical papers assessing the degree and nature of inequality of opportunity in different contexts none, to the best of our knowledge, includes parental neighbourhood into typically used parental background attributes. Some associate geography-of-birthplace as a circumstance, but this is often limited to Urban/Rural classifications (see Ferreira et al., 2011) or to very large administrative units such as regions (see Cogneau & Mesplé-Somps, 2008; Singh, 2012). Others partition the study area into fewer but bigger macro regions and evaluate inequality of opportunity separately (see Peragine & Serlenga, 2008; Checchi & Peragine, 2010). Among empirical EOp studies in which geography is taken into account, the work of Checchi et al. (2010) comes closest to what might be perceived as neighbourhoods. It provides a comparison between the inequality of opportunity levels in 25 European countries where the degree of population density in residential areas is included as circumstance variables.

The lack of spatial considerations in the current literature is probably related to a general lack of spatially coded data available to researchers. For many countries even a very limited amount of data in relation to parental background may not be available. Another reason might be the fact that a spatial approach requires the recognition of the link between temporal geography of residence and any opportunity distribution. As is often underlined by the scholars proposing variants of the equality of opportunity approach, children should not be held responsible for their choices in any way (de Barros, 2009; Björklund et al., 2012). Such a view is readily extendable to include the residential decisions of parents on behalf of their children. Therefore, given the context of equal opportunities literature there should be no objections to defining parental neighbourhood as a circumstance.

Several empirical studies investigate the effects of neighbourhoods on educational attainment (Garner & Raudenbush, 1991), drop-outs (Crane, 1991) and outcomes such as reading, maths achievement (see Ludwig et al., 2013) and higher education participation (Andersson & Malmberg, 2015). However, the extensive body of neighbourhood literature shows no consensus on the durability of neighbourhood effects. At least two distinctive empirical strategies seek to investigate whether neighbourhood characteristics continue to be effective years after the exposure ends.

The first strategy studies the correlation between siblings and neighbouring children in adult outcomes, a way of quantifying the variation/correlation in earnings that can be attributed to neighbourhood histories. For instance Page & Solon (2003) study the correlation between adult earnings of once neighbouring children and between brothers, where the defined neighbourhood contains 20-30 contiguous dwelling units. Their findings demonstrate a positive correlation in earnings among formerly neighbouring children (half of the correlation observed for brothers), and is interpreted as residential immobility i.e. the children who grew up in urban areas end up in urban areas where the earnings are higher and those who grew up in rural areas remain in rural areas with lower earnings. Using a similar approach Raaum et al. (2006) find declining neighbourhood effects on earnings and educational attainment as years go by.

The second empirical strategy makes use of so-called moving to opportunity (MTO) experiments, which are randomized social experiments on housing mobility conducted by the U.S. Department of Housing and Urban Development (HUD). Ludwig et al. (2013) show that moving to less disadvantaged neighbourhoods (census tracts) improves both mental and physical health conditions. However such an impact is not observed for economic conditions and educational attainment even for the children who were exposed to a better environment at an early age. In a recent study Chetty et al. (2015) conclude that for children, each additional year spent in less deprived neighbourhoods (U.S. counties) increases the likelihood of college attendance and of higher earnings in adult life [1].

Although the above papers differ in important respects in how they study neighbourhood effects, they all identify a neighbourhood as an area comprising a predefined administrative unit. A few studies make use of bespoke, individualized neighbourhood units. For example Van Ham et al. (2012) adopt this method for the Stockholm metropolitan area and show that the negative effects of neighbourhood history are both inherited and persistent over time. Similarly when investigating a population of parental home-leavers in Stockholm, Hedman et al. (2015) observe negative effects of exposure to a poverty-concentration parental neighbourhood even after 17 years of living away from parents.

[1] (see also Chetty & Hendren, 2015)

Throughout this study we refer to a reference year in which the statistics for neighbourhood histories are linked to individuals. The above and other papers (Quillian, 2003; Clark & Ledwith, 2006) use the same empirical technique. If there is high residential mobility where both within and between neighbourhood shifts occur, a one-year reference might lead to measurement errors. However, we believe that our reference to a single year for parental neighbourhood does not bring large measurement error bias given the high similarity in peoples neighbourhoods overtime not only in Sweden (see Van Ham et al., 2012) but also in many other countries (see Kunz et al., 2003; Quillian, 2003; Sharkey & Faber, 2014).

This study seeks to bridge the gap between the literature dedicated to the theory and methods of equality of opportunities and to neighbourhood effects. For the educational IOp investigation we start from previous empirical studies on neighbourhood effects, but given the lack of consensus on the durability of such effects, for the income IOp investigation we first show how the neighbourhood histories of individuals exert persisting effects on life chances, therefore their contribution to inequality should be quantified and accommodated in a matrix devoted to circumstances.

3 Data

This study uses the PLACE longitudinal database (located at the Department of Social and Economic Geography, Uppsala University) which contains socio-economic, demographic and geographical information for all Swedish residents since 1990. Following the same individuals over time, we investigate the distribution of compulsory examination marks in 2001 and the distribution of disposable incomes in 2010 for the whole 1985 cohort as our variables of interest. Two sets of independent variables are considered in the model: circumstances as measured by parental background and parental neighbourhood characteristics and effort as measured by educational level and own neighbourhood characteristics.

For each individual, we use spatial and aspatial information from the dataset (see Table 2). The aspatial set of variables includes several covariates typically used in EOp studies that are informative of the family background and other inherited circumstances. Eight such variables are used: gender, educational level, disposable income for 16-year olds residing in the household of upbringing (measured as part of household disposable income), whether or not a visible minority (VM, here understood as all individuals born outside Europe, USA, Canada or Australia), parent's marital status (single parent or dual parent households), parental education and employment status. The parental educational level is measured as the highest educational level reached by either of the parents. Employment status is measured as each parent working or not working in 2001.

The spatial set of variables includes parental neighbourhood characteristics in 2001 and own neighbourhood in 2010. These are quantified using a k nearest neighbour (knn) algorithm. Generating a form of scalable egocentric neighbourhood, this technique departs from each residential location and begins counting in every direction until a threshold (k) is reached. It then relates the population involved to the total counted population. The method does not require the use of predetermined administrative units and thus provides an efficient, comparable and robust definition of place (Östh et al., 2015). The computations were carried out using EquiPop software (Östh, 2014), which sorts people (in this case) according to a georeferenced grid and generates contextual variables quantifying the share of a given attribute within their k nearest neighbours, including for large data sets such as ours.

Table 2 [About Here]

We channel the following parental neighbourhood characteristics from 2001: the share of similar-age peers among the nearest 40 neighbours that accounts for socialization and network patterns, the share of visible minority

(VM) neighbours among the nearest 400 neighbours show the degree of segregation and deprivation, the share of single parents and families with 3 or more children (large families) among the nearest 40 neighbours account for household and housing characteristics.

In addition to these bespoke neighbourhoods, the negative environment surrounding the parental neighbourhood is constructed based on Corine (coordination of information on the environment) data, which is available as 100-meter pixel raster images. ArcGIS software is used to match the land cover data to the coordinates of individuals (both available as 100x100 geo-coordinates) and after a classification of good/bad elements of Corine, the exposure to negative surroundings within a 500m radios is imported into the data as a column vector. We use smaller k-levels for the year 2001 as the potential interaction with the neighbourhoods might be limited compared to 2010. see Table 3 for an interpretation of different k values.

The own neighbourhood in the year 2010 is defined as the share of VMs among the nearest 1600 neighbours and a measure of job/housing balance is computed for 2010 as follows: we first classify individuals according to their level of education and the jobs available to them under three categories: low, intermediate and high. Then for each residential location, the nearest 10,000 jobs and the longest distance to reach the workplace are computed by EquiPop. The assumption is that individuals seek jobs that correspond to their level of education. Thus for a lower educated person this method looks for available jobs in the low category alone. The observed Cartesian distance between home and work can be used as a crude measure of job accessibility at any location i. However, since some of the studied individuals were not in employment in 2010 and others may have travelled distances that are considerably different from others residing in close proximity, data from near neighbours need to be interpolated. In order to depict a local commuting distance that renders potential commuting distances for non-commuters and renders commuting distances that reduce outlier effects for individuals with very short or long distances we employ a Kriging strategy where the 12 nearest neighbours constitute the search radius for the commuting distance interpolation surrounding any location where a population member resided in 2010. The smoothed interpolation expresses a commuting distance used as a representation of the potential commuting distance at any location i. [2].

4 Analytic Framework

As in Ferreira & Gignoux (2011), we model compulsory examination exam as a function of circumstances (reduced form equation) as follows:

$$g_i = f(C_i, u_i,) \tag{1}$$

and the disposable income as a function of circumstances and effort as follows:

$$y_i = f(C_i, E_i(C_i, v_i), u_i,) \tag{2}$$

$$E_i = BC_i + v_i \tag{3}$$

where g_i is compulsory examination grade of i, y_i represents disposable income, C_i a vector of circumstances and E_i is of effort, finally u_i is unobserved determinants of disposable income such as luck. We recognize the correlation between effort on circumstances and other unobserved determinants with equation (3).

In general (1) and (2) are estimated by OLS regressions (see Bourguignon et al., 2007; Ferreira & Gignoux, 2011). In this study we employ a multilevel model with linear specification. It is obvious that spatial factors play a key role in this study. For this reason we need to specify a model that caters for the spatial patterns of variation that may lead to erroneous inferences. By employing the Morans I test on the regression residuals we can test if there are any spatial dependencies not catered for in the specified models

[2] Kriging was conducted in ARCInfo using the ordinary spherical semivariogram method, k=12

(Moran, 1950). Four models were tested: (1) full model OLS, (2) empty model MLM, (3) contemporary model MLM and (4) full model MLM. Model results reveal that the OLS and empty models fail to take the spatial autocorrelation into account. That the empty model fails to explain variation is expected since no parameters are included, but that the full OLS model lies comparatively close to the empty model and far from remaining models clearly indicates that using OLS does not cater for the spatial variation present in the dataset. Of the remaining two models, the contemporary model is the one with no spatial autocorrelation, whilst the full model displays a weak but significant spatial autocorrelation. The chief difference between models explains why the full models show spatial autocorrelation. In the contemporary model, individual level parameters as well as contemporary contextual variables are introduced. Variables and the multilevel approach account for the spatial variation in regression residuals. However, in the full model, contextual variables from the year 2001 are also included. The variables introduced improve the overall model fit (see Table 1) but also introduce a spatial bias related to the sorting of individuals during the years of upbringing.

Table 1 [About Here]

We specify the empirical models as:

$$g_{ij} = a_0 + a_{ij}C_{ij} + a_j x_j + t_j + q_{ij} \tag{4}$$

$$y_{ij} = \beta_0 + \beta_{ij}C_{ij} + \alpha_{ij}E_{ij} + \beta_j x_j + u_j + z_{ij} \tag{5}$$

$$E_{ij} = b_0 + b_{ij}C_{ij} + b_j x_j + v_j + e_{ij} \tag{6}$$

for individual i living in municipality j, g_{ij} represents the log of compulsory examination marks, y_{ij} is the log of disposable income, β_0 and a_0 are the intercepts, x_j represents municipality-level covariates, t_j and u_j are municipality-specific random effects. In order to measure income IOp, the correlation between circumstances C_{ij} and E_{ij} effort needs to be examined. Again we follow Roemer (1998) in treating the effort variables, because a fundamental aspect in this setting is the fact that the distribution of effort within each circumstance group is itself a characteristic of that type; since this is beyond individual control, it constitutes a circumstance. [3]. Therefore only genuine effort $\hat{e}_{ij}$ must be derived. Finally we estimate the following model:

$$y_{ij} = \beta_0 + \beta_{ij}C_{ij} + \alpha_{ij}\hat{e}_{ij} + \beta_j x_j + u_j + z_{ij} \tag{7}$$

where $\hat{e}_{ij}$ is the estimate obtained in (6).

Using the estimates from reduced form equation (4) and from the full model (7), we construct a counterfactual distribution of compulsory examination marks g_{ij} and of disposable income y_{ij} where all within inequalities are eliminated as follows:

$$g_i^c = exp[C_i\hat{a}_{ij}] \tag{8}$$

and

$$y_i^c = exp[C_i\hat{\beta}_{ij}] \tag{9}$$

Subsequently the absolute and relative inequality of opportunity measures are calculated both with a path-independent decomposable inequality index, namely the mean logarithmic deviation (MLD) and with the Gini index as $IO = I(g_i)$ and $IO = I(y_i)$. Following this procedure, we can see how much of the inequality is due to inequality in opportunities and the share attributable to effort.

$$EIOp = \frac{I(g_i^c)}{I(g_i)} \tag{10}$$

and

$$IOp = \frac{I(y_i^c)}{I(y_i)} \tag{11}$$

Using the same techniques we further decompose the relative contributions of spatial and aspatial factors to both circumstances and effort partitions of inequality. This practice is able to pinpoint the extent to which neighbourhoods influence given outcomes.

[3]see Jusot et al. (2013) for other approaches

$$EIOp_{spatial} = \frac{I(g_{ij}^{sc})}{I(g_{ij}^{c})} \tag{12}$$

In a similar manner for earnings inequality:

$$IOp_{spatial} = \frac{I(y_{ij}^{sc})}{I(y_{ij}^{c})} \quad \text{and} \quad IO_{spatial} = \frac{I(y_{ij}^{se})}{I(y_{ij}^{e})} \tag{13}$$

Therefore, $IOp_{spatial}$ quantifies the relative contribution of parental neighbourhoods to overall inequality due to circumstances and $IO_{spatial}$ indicates the relative contribution of own-neighbourhood to overall inequality due to effort.

5 Findings

The main goal of this study is to investigate the spatial sources of inequality in relation to neighbourhood characteristics to which individuals are exposed. In this section we first briefly report the regression results of two models on educational attainment and disposable income respectively, then we show the outcomes of inequality decomposition into circumstances/effort and spatial/aspatial components. Before we proceed with the inequality decomposition, we verify the temporal extent of parental neighbourhood histories.

Table 4 shows the marginal effects of circumstance variables on compulsory examination marks. We employed the following 10 circumstance variables including the spatial covariates: gender, minority status (VM or not), the highest level of education attained by parents and employment and the marital status of parents, disposable income in 2001 and the share of the following attributes in the neighbourhood (k levels in parenthesis): single headed families (k=40), same age children (k=40), families with at least 3 children (k=40) and visible minorities (k=400). All coefficients have expected signs and are statistically significant at the 0.001 level (p values=0.00), except the share of same-age peers in the neighbourhood that is also significant but at the 0.05 level with a negative coefficient sign. This result may be interpreted as the distraction impact of having same-age peers in the neighbourhood due to the longer time spent on non-school activities. However we acknowledge the importance of socialization for the development of children. Besides, as mentioned in the following paragraphs, having same-age peers in the parental neighbourhood is positively associated with the subsequent disposable income of adults.

Of the remaining variables, living in a VM-concentrated area shows the strongest effect on marks. A likely explanation for the strong effect is that the share of VMs in a neighbourhood might coincide with the poverty concentration and other possible adverse characteristics of the locality. This can be seen from the maps in Fig.1 where on the left hand side the VM population in the 400 nearest neighbours for the whole population is shown and on the right hand side the poverty concentration(OECD criteria) among the 400 nearest neighbours is mapped for the whole Stockholm metropolitan area. These maps show how the two aspects of the locality are statistically entangled, so that almost the same pattern of segregation is observed for both attributes of neighbourhoods.

Furthermore, positive effects are observed for students with employed parents (slightly higher if the mother is employed), with the presence of at least one highly educated parent at home and for students with high disposable incomes. Negative effects are found for students with a single parent and those who belong to VMs. The estimates for the single-parent and large families in the neighbourhood as well as negative surroundings show a negative association with educational attainment although the multilevel model controls the variability both at the individual and municipality level. The negative effect of single-mother concentration in neighbourhoods is a well explored phenomenon especially in the US, the single-parent specification in our model shows a similar pattern in Sweden. Due to lower household income, these families reside in worse neighbourhoods, therefore the variable performs as a proxy of residential environment and housing conditions to which the study population is exposed. A similar interpretation can be given to large-family concentration. Since large apartments are not found in the central districts, these families reside in rural areas or areas with rural character far from amenities. Furthermore, with three or more children, mothers stop studying at an early age, therefore large-family concentrated areas might be characterized by lower human capital accumulation.

Table 4 [About Here]

To analyze disposable income, we added gender to the circumstance variables from the previous model with the following effort variables: compulsory examination marks, job/housing balance(k=10000) and observed commuting distance between job opportunities and individual residences and the share of VMs (k=1600) in 2010. The variable for the highest level of educational attainment among parents interacts with ten city classes to account for different degrees of industrialization in cities [4]. We examine the correlation between effort and circumstances by equation (3) for all effort variables. This procedure guarantees that the effort variables reflect only pure effort, without the influence of observed circumstances. Then we substitute the resulting residuals terms in equation (2). Now, the circumstances in (5) are expected to reflect both their direct impacts on the response variable and indirect effects on 5 effort variables. An important result to note is that as we regress the VM concentration of own-neighbourhood in 2010 on the circumstance variables, most of the variation is explained by the VM share of the parental neighbourhood. It is apparent from this result that the study population ended up in similar neighbourhoods as their parents. This residential immobility, or similarity in neighbourhood characteristics over time, points to long-term exposure to whatever effects neighbourhoods produce and the likelihood that

[4]Multilevel regression showed a negative association with the disposable income of individuals and their parental education. To correct this we used a classification that sorted municipalities by the degree of industrialization

these effects are transmitted to offspring. People may sort into neighbourhoods because of hedonistic motivations (quality services etc.) or because they might prefer to live with similar people. On the other hand, some neighbourhoods might be well (or less well)-endowed in terms of public goods and services because certain income/education groups and taxpayers live in those areas. We do not attribute any causal links between the two. However, from the equal opportunities perspective, being locked-in parental neighbourhoods (or to those with similar characteristics to parental neighbourhoods) is clearly a factor influencing life chances. For this reason, even though we deem individuals responsible for their choice of neighbourhood, it seems reasonable to derive pure effort purged of the influence of parental neighbourhood and other circumstances through the procedure explained above.

Table 5 shows regression results of the multilevel model for the log of disposable income. All the variables are statistically significant and those in common with the previous model show the same association with disposable income, except the share of same-age peers, which now has a positive coefficient sign. This is an interesting result that shows how the effects of residential contexts may differ over time. Growing up in an area with the strong likelihood of interaction with similar-age peers probably increases the chances of finding a better job and of being successful when employed through childhood ties.

Of the aspatial circumstance variables, gender, belonging to VMs and having a single parent are negatively associated and disposable income in 2001, parental education and employment positively associated with disposable income. As for the spatial circumstances (i.e. parental neighbourhood attributes), the strongest effects are found for the share of VMs in the parental neighbourhood in 2001. Furthermore, growing up in areas with a high proportion of single-parent and large families and having negative environmental surroundings are negatively associated with subsequent earnings. Turning now to the discussion of the long-term effects of neighbourhoods, we are in position to conclude that historical neighbourhood characteristics influence adult earnings even though a range of individual and household characteristics are present in the model.

The only aspatial effort variable employed was compulsory examination marks. It shows a relatively lower effect on the response variable, which is in part attributable to the fact that all effort variables were purged of any influence of circumstances, as explained in the previous section. Regarding the spatial effort variables (i.e. own-neighbourhood characteristics) living in a neighbourhood with a high proportion of VMs and the observed commuting-distance between home and job are negatively associated and the degree of job/housing balance positively associated with disposable income. Since the spatial mismatch hypothesis first advanced by Kain (1968), there has been great interest in understanding differences in unemployment and job search success rates, job accessibility and job/housing mismatch (see for example Kain, 1992; Van der Klaauw & Van Ours, 2003; Houston, 2005). We are unaware of any application of job/housing (mis)match using the k nearest neighbour algorithm, which in return accounts for both residential location-driven and skill-based job accessibility.

Table 5 [About Here]

In terms of variation, the multilevel model indicates that the variance in disposable income is largely attributable to individuals. However it is important to remember that the fixed part of the model includes not only individual-level effects but also contextual variables that are defined individually both in parental neighbourhoods and when living independently of parents. Moreover, the 1% variation is explained by the municipality level covariates. If people were forced to live in certain municipalities, we would conclude that this value is the inequality of opportunity produced by Swedish municipalities. But since individuals are entitled to choose freely where to live, we consider them responsible for their choice of municipality in 2010 but not in 2001 while living with parents.

Fig.2 shows the variables over which the decomposition is undertaken and the second and third columns of Table 6 illustrate the magnitude of income and educational inequality of the entire Swedish population born in 1985. Based on the estimated coefficients from equations (4) and (7), the overall opportunity share in total inequality in income is computed as 8.05% and the overall opportunity share in educational inequality is 42.63% as measured by MLD. Note that since the effort partition contains both the effort and the unexplained part of disposable income variation and only the unexplained part of variation in educational inequality, the IOp estimates are lower bound.

As far as disposable income inequality is concerned, the relative decomposition of circumstances shows that 16.66% of the total share of circumstances is attributable to spatial circumstances (parental neighbourhood) and 83.33% is to aspatial circumstances. While the fair inequality decomposition indicates that the 01.95% of the total effort is due to spatial effort (own-neighbourhood) and the remaining part is caused by aspatial effort (compulsary examination marks) and the residual of the model. The corresponding decomposition for educational inequality shows that the spatial circumstances (parental neighbourhood) represent 36.44% of total circumstances. Even though earnings and educational attainment are two different outcomes, we can conclude that the neighbourhoods are more influential while the exposure is ongoing.

Table 6 [About Here]

We also conducted separate analyses for gender. The estimates of total inequality in educational attainment and income and related decomposition results are shown in Table 7 for females and Table 8 for the male population. A higher income IOp is observed among men than women. However, the relative income IOp is almost identical for both. The latter result seems to be related to the spatial sources of IOp. For women parental neighbourhood is more influential than for men (27.27% for women compared to 16.66% for men). On the other hand, spatial effort counts more for men than for women with 5.21% and 4.06% respectively. Regarding the IOp in educational attainment, overall circumstances explain a higher degree of variation for male students. Again this result seems to be related to parental neighbourhoods. For male students 30.15 percent of the total circumstance pertains to spatial circumstances, it is only 13.75 percent for female students.

Comparing the results in Table 7 concerning the effects of spatial circumstances, parental neighbourhood is more influential for the educational attainment of male students, hence during exposure. Once in employment, male students seem to be more successful in overcoming these effects through spatial effort than the female population. That is to say the male population uses *mobility* as an instrument at their disposal to generate additional income and to decrease the gap with higher income groups. In line with the findings for the whole population, for men the influence of parental neighbourhood proportionally decreases from 2001 to 2010. However, it is interesting to note that for the female population parental neighbourhoods cause a lower variation in marks during exposure and a higher variation in subsequent earnings than males. One interpretation of this pattern is that while being exposed to characteristics explained above for parental neighbourhoods, female students might manage to focus on their studies and reflect the adverse effects of parental neighbourhoods to a relatively lower degree to their marks. Yet, as females grow up, they might be building personal identities similar to that of the residents of their parental neighbours. This is a relevant interpretation especially given the fact that neighbourhood statistics for single-parent and large families mostly relate to women. Another view relates to mobility patterns. For the female population, parental neighbourhoods potentially become own-neighbourhoods since they seem to be immobile.

Table 7 [About Here]

Table 8 [About Here]

6 Concluding Remarks

A society is said to be equal in opportunities if the life chances of individuals do not depend on the factors beyond their choice and effort and the systematic differences in any outcome that are explained by so-called circumstances is considered as inequality of opportunity. It has been shown repeatedly that parental background plays an important part in the life chances of individuals. So far, however, there has been no discussion of the role of parental neighbourhood as a source of illegitimate inequality.

In this paper investigating the inequality in educational and earnings opportunities in Sweden for the whole 1985 cohort, we included parental neighbourhood statistics in a matrix of circumstance variables and own neighbourhood characteristics in a matrix devoted to effort variables. We constructed egocentric neighbourhoods where a count of k-nearest population forms the neighbourhood and the overall share of individuals who carry given characteristics identifies the likelihood of interactions. In addition, the share of negative components surrounding parental residence and a measure of housing/job market balance and the observed commuting distance between own-neighbourhoods and job opportunities were added to the analyses. Instead of the standard OLS approach, we utilized a multilevel model, which overcame most of the spatial autocorrelation problem.

The findings indicate that as well as the conventional aspatial circumstance variables, parental neighbourhoods strongly impact educational attainment and even years after exposure they remain influential for earnings distribution. Therefore, based on the evidence from Swedish data, we can conclude that in order to obtain accurate measures of IOP, there is a definite need for a multidisciplinary approach that links individual outcomes to neighbourhoods.

We hope that our findings may influence the way in which IOp analyses are conducted both in terms of methods and techniques to quantify and include characteristics of neighbourhoods. We are aware that the latter requires detailed information on geo-locations. This is the central policy implication of our study, that data collection methods must be designed to contain necessary geographic information on residents. The recent developments in data collection methods associated with "big data" significantly facilitate the collection of contextual variables. For instance there is a vast quantity of information that is made freely available on internet through open maps and the social-media data provides a wealth of information to researchers. Therefore, what is left is the proper handling of geography. The findings of this study recommend using bespoke neighbours to define an individual's residential environment. Creating individualized neighbourhoods based on the k nearest neighbour algorithm enabled us to overcome problems associated with administratively defined areas plagued by indeterminacy. Through this approach, this paper has gone some way towards enhancing our understanding of the temporal effects of neighbourhoods.

Furthermore, our results show that the opportunity gap between individuals widens both for visible minorities and for the residents of visible minority-concentrated neighbourhoods. Therefore, another important implication specific to Swedish data is that in order to decrease inequality in opportunities, an effective policy must target the population belonging to the visible minority population and their residential environment. Observed negative effects with proxies of housing conditions suggest a need for comprehensive analysis of segregation not only by nominal categories but also by income.

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.

Björklund, A., Jäntti, M., & Roemer, J. E. (2012). Equality of opportunity and the distribution of long-run income in sweden. *Social choice and welfare*, *39*(2-3), 675–696.

Bourguignon, F., Ferreira, F. H., & Menendez, M. (2007). Inequality of opportunity in brazil. *Review of income and Wealth*, *53*(4), 585–618.

Brunori, P., Ferreira, F. H., & Peragine, V. (2013). *Inequality of opportunity, income inequality, and economic mobility: Some international comparisons.* Springer.

Buck, N. (2001). Identifying neighbourhood effects on social exclusion. *Urban studies*, *38*(12), 2251–2275.

Case, A. C., & Katz, L. F. (1991). *The company you keep: The effects of family and neighborhood on disadvantaged youths* (Tech. Rep.). National Bureau of Economic Research.

Checchi, D., & Peragine, V. (2010). Inequality of opportunity in italy. *The Journal of Economic Inequality*, *8*(4), 429–450.

Checchi, D., Peragine, V., & Serlenga, L. (2010). Fair and unfair income inequalities in europe.

Chetty, R., & Hendren, N. (2015). The impacts of neighborhoods on intergenerational mobility: Childhood exposure effects and county-level estimates. *Unpublished Manuscript*.

Chetty, R., Hendren, N., & Katz, L. F. (2015). *The effects of exposure to better neighborhoods on children: New evidence from the moving to opportunity experiment* (Tech. Rep.). National Bureau of Economic Research.

Clark, W. A., & Ledwith, V. (2006). Mobility, housing stress, and neighborhood contexts: evidence from los angeles. *Environment and Planning A*, *38*(6), 1077–1093.

Cogneau, D., & Mesplé-Somps, S. (2008). Inequality of opportunity for income in five countries of africa. *Research on Economic Inequality*, *16*, 99–128.

Crane, J. (1991). The epidemic theory of ghettos and neighborhood effects on dropping out and teenage childbearing. *American journal of Sociology*, 1226–1259.

de Barros, R. P. (2009). *Measuring inequality of opportunities in latin america and the caribbean.* World Bank Publications.

de Barros, R. P., Vega, J. R. M., & Saavedra, J. (2008). Measuring inequality of opportunities for children. *Washington, DC: World Bank*.

Ellen, I. G., & Turner, M. A. (1997). Does neighborhood matter? assessing recent evidence. *Housing policy debate*, *8*(4), 833–866.

Ferreira, F. H., & Gignoux, J. (2011). The measurement of inequality of opportunity: Theory and an application to latin america. *Review of Income and Wealth*, *57*(4), 622–657.

Ferreira, F. H., Gignoux, J., & Aran, M. (2010). Inequality of economic opportunity in turkey: An assessment using asset indicators and womens background variables. *State Planning Organization of the Republic of Turkey and World Bank Welfare and Social Policy Analytical Work Program Working Paper*(3).

Ferreira, F. H., Gignoux, J., & Aran, M. (2011). Measuring inequality of opportunity with imperfect data: the case of turkey. *The Journal of Economic Inequality*, *9*(4), 651–680.

Galster, G. (2001). On the nature of neighbourhood. *Urban studies*, *38*(12), 2111–2124.

Garner, C. L., & Raudenbush, S. W. (1991). Neighborhood effects on educational attainment: A multilevel analysis. *Sociology of education*, 251–262.

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.

Houston, D. (2005). Employability, skills mismatch and spatial mismatch in metropolitan labour markets. *Urban Studies*, *42*(2), 221–243.

Jusot, F., Tubeuf, S., & Trannoy, A. (2013). Circumstances and efforts: How important is their correlation for the measurement of inequality of opportunity in health? *Health economics*, *22*(12), 1470–1495.

Kain, J. F. (1968). Housing segregation, negro employment, and metropolitan decentralization. *The Quarterly Journal of Economics*, *82*(2), 175–197.

Kain, J. F. (1992). The spatial mismatch hypothesis: three decades later. *Housing policy debate*, *3*(2), 371–460.

Kunz, J., Page, M. E., & Solon, G. (2003). Are point-in-time measures of neighborhood characteristics useful proxies for children?s long-run neighborhood environment? *Economics Letters*, *79*(2), 231–237.

Leventhal, T., & Brooks-Gunn, J. (2000). The neighborhoods they live in: the effects of neighborhood residence on child and adolescent outcomes. *Psychological bulletin*, *126*(2), 309.

Ludwig, J., Duncan, G. J., Gennetian, L. A., Katz, L. F., Kessler, R. C., Kling, J. R., & Sanbonmatsu, L. (2013). *Long-term neighborhood effects on low-income families: Evidence from moving to opportunity* (Tech. Rep.). National Bureau of Economic Research.

Openshaw, S. (1984). The modifiable areal unit problem..

Östh, J. (2014). *Introducing the equipop software–an application for the calculation of k-nearest neighbour contexts/neighbourhoods.*

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

Page, M. E., & Solon, G. (2003). Correlations between brothers and neighboring boys in their adult earnings: The importance of being urban. *Journal of Labor Economics*, *21*(4), 831–855.

Peragine, V., & Serlenga, L. (2008). Higher education and equality of opportunity in italy. *Research on Economic inequality*, *16*, 67–97.

Quillian, L. (2003). How long are exposures to poor neighborhoods? the long-term dynamics of entry and exit from poor neighborhoods. *Population Research and Policy Review*, *22*(3), 221–249.

Raaum, O., Salvanes, K. G., & Sørensen, E. Ø. (2006). The neighbourhood is not what it used to be*. *The Economic Journal*, *116*(508), 200–222.

Ramos, X., & Van de Gaer, D. (2012). Empirical approaches to inequality of opportunity: Principles, measures, and evidence.

Roemer, J. E. (1998). *Theories of distributive justice*. Harvard University Press.

Roemer, J. E., & Trannoy, A. (2013). Equality of opportunity.

Ross, J. L., Iannotta, J. G., et al. (2002). *Equality of opportunity and the importance of place:: Summary of a workshop*. National Academies Press.

Sharkey, P., & Faber, J. W. (2014). Where, when, why, and for whom do residential contexts matter? moving away from the dichotomous understanding of neighborhood effects. *Annual Review of Sociology*, *40*, 559–579.

Singh, A. (2012). Inequality of opportunity in earnings and consumption expenditure: The case of indian men. *Review of Income and Wealth*, *58*(1), 79–106.

Van der Klaauw, B., & Van Ours, J. C. (2003). From welfare to work: does the neighborhood matter? *Journal of Public Economics*, *87*(5), 957–985.

Van Ham, M., Hedman, L., Manley, D., Coulter, R., & Östh, J. (2012). Intergenerational transmission of neighbourhood poverty in sweden: an innovative analysis of individual neighbourhood histories.

Vartanian, T. P. (1999). Adolescent neighborhood effects on labor market and economic outcomes. *Social Service Review*, *73*(2), 142–167.

Tables and Figures

Table 1: Moran's I Tests

	Full model OLS	Empty model MLM	Contemporary model MLM	Full model MLM
Moran's I	0,005457	0,005947	0,000181	0,00104
Expect I	-0,000025	-0,000025	-0,000025	-0,000025
z-score:	13,437288	15,425446	0,534853	2,98011
p-value:	0,00000	0,00000	0,592751	0,002881

Table 2: Variables

	Variables	Description
1. Individual	Gender	1=Female, 0=Male
	Visible Minority	1=VM, 0=Not a VM
	Compulsory Exam marks	
	Family Background	Parents' Employment Status and Education, single-headed household
2. neighbourhood	Share of Visible Minorities	2001(k=400) and 2010(k=1600)
	Single-Parent Families	2001(k=40)
	Share of families with 3 or more children	2001 (k=40)
	Share of same-age peers	2001(k=40)
	Negative Space	Corine database 2001 (500m radii-based)
	Housing/Job Market Balance	2010 k=10000 and commuting distance

Table 3: K-neighbours

k-neighbours	Possible Interactions
12	Stairs in building
25	Building
50	Bicycle basement, garbage recycling bins etc.
100	Block
200	Bus stop
400	Kiosk, familiar with topology, recognize all neighbours
800	Football field
1600	Small shop
3200	Day care, school
6400	Local square, different retail stores, dentist...
12800	Upper secondary schools, Big stores, communities (sports, religion)
25600	Hospital, place-belonging, municipality

Table 4: Multilevel Model: Log of marks (Compulsory Exam)

Fixed	Coef.	Standard Error	P values
1.Individual			
Employment Father	0.2361	0.0180	0.000
Employment Mother	0.2614	0.0170	0.000
Parental Education	0.1694	0.0123	0.000
Single Parent	-0.2322	0.0143	0.000
Visible Minority	-0.1702	0.0313	0.000
Disposable Income(2001)	0.1310	0.0151	0.000
2.neighbourhood			
Single-Headed Families(2001)_20	- 0.2323	0.0143	0.000
Large Families(2001)_20	-0.2253	0.0468	0.000
Same-Age peers_20	-0.2399	0.1136	0.035
Negative Space	-0.2063	0.0411	0.000
Visible Minority(2001)_200	-1.0354	0.0937	0.000
Random Effects Parameters	Estimate	Standard Error	
Municipality Level var(_cons)	0.0094	0.0020	
Var(Residual)	2.9919	0.0139	
Number of obs = 92674			

Table 5: Multilevel Model: Log of Disposable Income

Fixed	Coef.	Standard Error	P values
1.Individual			
Gender	-0.1644	0.0030	0.000
Compulsory Exam marks	0.0188	0.0008	0.000
Employment Father	0.0714	0.0047	0.000
Employment Mother	0.0784	0.0044	0.000
Parental Education			
x CityClass2	0.0634	0.0108	0.000
x CityClass3	0.0252	0.0081	0.002
x CityClass4	0.1121	0.0223	0.000
x CityClass5	0.0808	0.0151	0.000
x CityClass6	0.0508	0.0220	0.021
x CityClass7	0.1039	0.0144	0.000
x CityClass8	0.0968	0.0303	0.001
x CityClass9	0.0850	0.0135	0.000
x CityClass10	0.0845	0.0200	0.000
Single Parent	-0.0249	0.0037	0.000
Visible Minority	-0.1152	0.0081	0.000
Disposable Income(2001)	0.1211	0.0041	0.000
2.neighbourhood			
Large Families(2001)_40	-0.0382	0.0123	0.020
Single-Headed Families(2001)_40	-0.0789	0.0113	0.000
Same-Age peers_40	0.0790	0.0296	0.002
Negative Space	-0.0291	0.0107	0.007
Visible Minority(2001)_400	-0.3598	0.0248	0.000
Job/Housing Balance(2010)_10000	0.2074	0.0129	0.000
Commuting Distance(2010)	-0.0117	0.0007	0.000
Visible Minority(2010)_1600	-0.6481	0.0205	0.000
Random Effects Parameters	Estimate	Standard Error	
Municipality Level var(_cons)	0.0024	0.0003	
Var(Residual)	0.2045	0.0009	
Number of obs = 91413			

Table 6: Inequality Decomposition

	Income Inequality				Educational Inequality		
Total Inequality (GINI)	0.2674				0.1749		
Total Inequality(MLD)	0.1315				0.2667		
Inequality of Opportunity(GINI)	0.0695				0.1528		
Inequality of Opportunity(MLD)	0.0106				0.1137		
	Effort		Circumstances		Effort	Circumstances	
Contibuiton(%) to Total inequality (MLD)	91.95%		8.05%		57.37%	42.63%	
	Aspatial (residual)	Spatial	Aspatial	Spatial		Aspatial	Spatial
Spatial/Aspatial (MLD)	98.05%	1.95%	83.33%	16.66%	residual	63.06%	36.94%

Table 7: Inequality Decomposition Female Population Only

	Income Inequality(Female)				Educational Inequality (Female)		
Total Inequality (GINI)	0.2495				0.1657		
Total Inequality(MLD)	0.1105				0.2552		
Inequality of Opportunity(GINI)	0.0548				0.1436		
Inequality of Opportunity(MLD)	0.0058				0.1091		
	Effort		Circumstances		Effort	Circumstances	
Contibuiton(%) to Total inequality (MLD)	94.76%		5.24%		57.24%	42.76%	
	Aspatial (residual)	Spatial	Aspatial	Spatial		Aspatial	Spatial
Spatial/Aspatial (MLD)	95.94%	4.06%	72.73 %	27.27%	residual	86.25%	13.75%

Table 8: Inequality Decomposition Male Population Only

	Income Inequality(Male)				Educational Inequality (Male)		
Total Inequality (GINI)	0.2729				0.1784		
Total Inequality(MLD)	0.1423				0.2747		
Inequality of Opportunity(GINI)	0.0623				0.1601		
Inequality of Opportunity(MLD)	0.0082				0.1342		
	Effort		Circumstances		Effort	Circumstances	
Contibuiton(%) to Total inequality (MLD)	94.02%		5.80%		51.12%	48.88%	
	Aspatial (residual)	Spatial	Aspatial	Spatial		Aspatial	Spatial
Spatial/Aspatial (MLD)	94.79%	5.21%	50.00 %	16.66%	residual	69.85%	30.15%

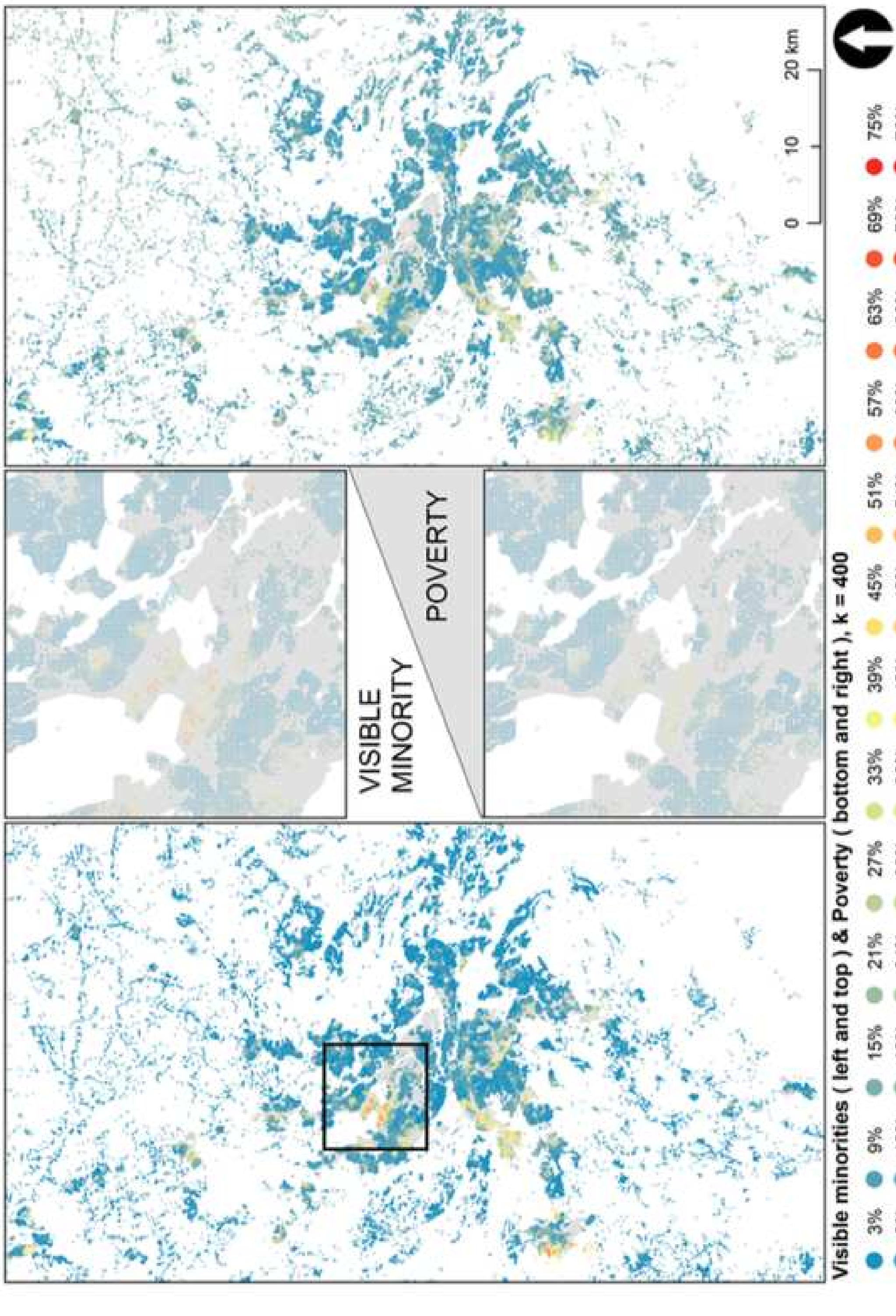
VISIBLE
MINORITY
POVERTY
0
10
20 km
Visible minorities (left and top) & Poverty (bottom and right), k = 400
3% 9% 15% 21% 27% 33% 39% 45% 51% 57% 63% 69% 75%
6% 12% 18% 24% 30% 36% 42% 48% 54% 60% 66% 72% 78%

Figure 1: