

Urban poverty: Measurement theory and evidence from American cities*

Francesco Andreoli[†]

Mauro Mussini[‡]

Vincenzo Prete[§]

Claudio Zoli[¶]

September 4, 2020

Abstract

We characterize axiomatically a new index of urban poverty that i) captures aspects of the incidence and distribution of poverty across neighborhoods of a city, ii) is related to the Gini index and iii) is consistent with empirical evidence that living in a high poverty neighborhood is detrimental for many dimensions of residents' well-being. Widely adopted measures of urban poverty, such as the concentrated poverty index, may violate some of the desirable properties we outline. Furthermore, we show that changes of urban poverty within the same city are additively decomposable into the contribution of demographic, convergence, re-ranking and spatial effects. We collect new evidence of heterogeneous patterns and trends of urban poverty across American metro areas over the last 35 years.

Keywords: Concentrated poverty, axiomatic, Gini, decomposition, census, ACS, spatial.
JEL codes: C34, D31, H24, P25.

*We are grateful to conference participants at RES 2018 meeting (Sussex), LAGV 2018 (Aix en Provence) and ECINEQ 2019 (Paris) and to two anonymous reviewers for valuable comments. The usual disclaimer applies. Replication code for this article is accessible from the authors' web-pages. This article forms part of the research project *The Measurement of Ordinal and Multidimensional Inequalities* (grant ANR-16-CE41-0005-01) of the French National Agency for Research and the NORFACE project *IMCHILD: The impact of childhood circumstances on individual outcomes over the life-course* (grant INTER/NORFACE/16/11333934/IMCHILD) of the Luxembourg National Research Fund (FNR), whose financial support is gratefully acknowledged.

[†]Corresponding author. Department of Economics, University of Verona, Via Cantarane 24, 37129 Verona, Italy and Luxembourg Institute of Socio-Economic Research (LISER), MSH, 11 Porte des Sciences, L-4366 Esch-sur-Alzette/Belval Campus, Luxembourg. E-mail: francesco.andreoli@univr.it.

[‡]DEMS, University of Milan-Bicocca, Via Bicocca degli Arcimboldi, 8 - 20126 Milan, Italy. E-mail: mauro.mussini1@unimib.it.

[§]University of Verona, Via Cantarane 24, 37129 Verona, Italy. E-mail: vincenzo.prete@univr.it.

[¶]University of Verona, Via Cantarane 24, 37129 Verona, Italy. E-mail: claudio.zoli@univr.it.

1 Introduction

Much of the literature studying economic inequality has focused on the distribution of income at national or regional level. Inequality at urban level is also important (Glaeser, Resseger and Tobio 2009). In America, for instance, cities are among the most unequal places in the country (Moretti 2013, Baum-Snow and Pavan 2013). Over the last three decades, urban income inequality has increased substantially in most of American metro areas (Watson 2009), albeit heterogeneously across cities. Inequality within and across neighborhoods is substantial (Wheeler and La Jeunesse 2008) and increasingly related to the trends of citywide income inequality, with low and high income households living in close spatial proximity (Andreoli and Peluso 2018).

Poverty is a key driver of income inequality in American cities. The urban poor population, i.e. the individuals living in households with aggregate income below the federal income poverty line and who reside in cities, has increased from 25.4 mln in 1980 to 31.1 mln in 2000 and up to 43.7 mln in the 2012-2016 period (estimates based on Census and American Community Survey data). These figures correspond to about 11% of the population before 2000, rapidly increasing to 14.9% after the Great Recession. The geography of poverty has also evolved over the same period. The number of census tracts displaying extreme poverty (where at least 40% of the population is poor) has almost doubled since 2000 (rising from 2,510 to 4,412 in 2013), offsetting demographic growth of about 11% during the same period (Jargowsky 1997, Jargowsky 2015).

In this paper, we study the spatial distribution of poverty across neighborhoods of a city. The goal is to establish a measurement apparatus for assessing the extent of *urban poverty* displayed by a city, a concept encompassing concerns for incidence of poverty in the city, as well as for the unequal distribution of poverty across neighborhoods. We outline an axiomatic approach which characterizes a new parametric index, mapping the distribution of poor and non-poor population across neighborhoods of a city into a number, which is the level of urban poverty displayed by that city. Alternative measures of urban poverty widely adopted in the literature and in policy analysis alike may violate some of the axioms we outline.

One of such measures is the *concentrated poverty* index. The index, introduced by

Wilson (1987), is the share of a metro area’s poor population (identified by an exogenous income poverty line) that lives in neighborhoods where poverty is extremely concentrated. According to the American Census Bureau, these neighborhoods qualify as places where more than 20% (or 40% in places where poverty is extremely concentrated) of the resident population is poor. A number of contributions (Jargowsky and Bane 1991, Massey, Gross and Eggers 1991, Jargowsky 1997, Kneebone 2016, Iceland and Hernandez 2017) have documented the patterns and drivers of concentrated poverty across American metro areas. After a decline of concentrated poverty in the 1990s, the 2000s and 2010s have witnessed a re-concentration of poverty, rising from 11% to 14.1% in the largest 100 American metro areas. Patterns are heterogenous across metro areas and depend on differences in the size, geographic location, income inequality alongside the degree of income and ethnic segregation in the city.

In a urban context, an uneven distribution of poverty across neighborhoods is a relevant dimension of collective well-being. In fact, in cities where poverty is highly concentrated in few neighborhoods, poor residents living therein are overexposed to poverty, thus likely suffering the consequences of a *double burden* of poverty related to its geographic distribution. Evidence of negative effects of poverty concentration at neighborhood level has been found on a variety of relevant individual outcomes, such as health (Ludwig et al. 2011, Ludwig et al. 2013), labor market attachment (Conley and Topa 2002), individual well-being (Ludwig et al. 2012) and the economic opportunities of future generations (Chetty, Hendren and Katz 2016, Chetty and Hendren 2018).¹

A way to formalize the spillover effect of concentrated poverty is to assume that individual well-being depends on household characteristics (such as poverty status) alongside the proportion of poor in the neighborhood (as in Bayer and Timmins 2005), and well-being is decreasing in this proportion. The larger the share of population exposed to high-poverty neighborhoods, the stronger urban poverty impacts collective well-being (all else equal). A urban poverty measure which is consistent with this view should not register less urban poverty when the share of poor population living in extreme poverty neighborhoods rises, even if this increment originates from a reduction of the incidence of poverty

¹See Oreopoulos (2003) for a critical assessment.

in neighborhoods where poverty is less extreme. Using simple counterexamples, we show that the concentrated poverty index may violate this intuitive principle (see also Massey and Eggers 1990, Jargowsky 1996).

This paper addresses this measurement concern and introduces a new measure of urban poverty, that is inspired by inequality analysis and is consistent with the intuitive requirement outlined above. Our measure weights three components of urban poverty that positively contribute to it: first, the incidence of poor residents in high poverty neighborhoods; second, the inequality in the distribution of poverty within the cluster of high poverty neighborhoods; third, the extent of inequality in the distribution of poor residents across high poverty and low poverty neighborhoods. The index we characterize shares features in common with prominent members of the class of rank-dependent poverty measures (Bosmans 2014, Ebert 2010, Sen 1976)² and addresses robustness concerns about the way high and low poverty neighborhoods are defined (a related point is raised in Shorrocks 1995, Thon 1979). We organize our results in Section 3, whereas Section 2 provides the setting.

When the focus is on the distribution of poverty across the whole city, the urban poverty index is shown to converge to a Gini-type index. In this case, we demonstrate that the longitudinal variation in urban poverty is additively and non-parametrically decomposable into the contribution of demographic growth, of poverty convergence across neighborhoods and of spatial association in poverty changes (Section 4). The decomposition is relevant for assessing whether urban poverty is mostly driven by neighborhoods that are spatially clustered, unveiling local poverty traps that can potentially reinforce the double burden effects of poverty concentration, or rather urban poverty is idiosyncratic to the neighborhoods characteristics.

In Section 5, we employ our measurement apparatus to assess the dynamics of poverty across all American metro areas over the last 35 years, exploiting rich data from the Census and the American Community Survey (ACS). Our main findings are that: i) American metro areas display strong heterogeneity in urban poverty patterns; ii) Urban poverty has

²Bosmans (2014) suggested a new criterion to compare poverty measures in terms of distribution-sensitivity to income transfers among poor individuals. This distribution-sensitivity criterion is based on “lossy” transfers and is useful to rank many of the best-known rank-dependent poverty measures.

not evolved significantly over the 35 years and has been hardly affected by the Great Recession burst, contrary to the rising trends of concentrated poverty; iii) Both re-ranking and convergence components of urban poverty changes are substantial across metro areas, indicating the role of changes in neighborhood poverty composition; iv) The spatial component of urban poverty is negligible for the large majority of cities, but very significant in largest metro areas where clustering of high-poverty neighborhoods seems to be an issue.

Section 6 concludes with a discussion.

2 Measuring urban poverty

2.1 Setting

For any given city, we consider a partition of the urban space into n neighborhoods. In empirical analysis, neighborhoods can coincide with an administrative division of the territory, such as the partition of American cities into census tracts. We take the partition into neighborhoods as given, and we study the distribution of poor and non-poor people therein.

Let $i \in \{1, \dots, n\}$, with n a positive natural number, denote a neighborhood of a city and $N_i \in \mathbb{R}_+$ be the individuals living in that neighborhood, with $N = \sum_{i=1}^n N_i$.³ An individual is poor when living in a household whose total disposable income is smaller than an exogenous poverty threshold (such as the federal income poverty line provided by the American Census Bureau), calculated in a given year for that specific type of household (for instance, depending on the size and the age structure). The analysis of urban poverty is hence conditional on the definition of poverty status, which we take as given. In our application, for instance, residents are poor if they live in households with an equivalent income smaller than 100% of the federal poverty line. Furthermore, let P_i denote the number of poor individuals living in neighborhood i , while $P = \sum_{i=1}^n P_i$ is the number of poor individuals in the city. A *urban poverty configuration* is a collection of counts of poor and non-poor individuals distributed across neighborhoods and it is denoted

³In order to simplify the exposition population aggregates are assumed to take real values. The assumption accommodates the case where representative population counts are estimated from underlying sample data (such as the ACS).

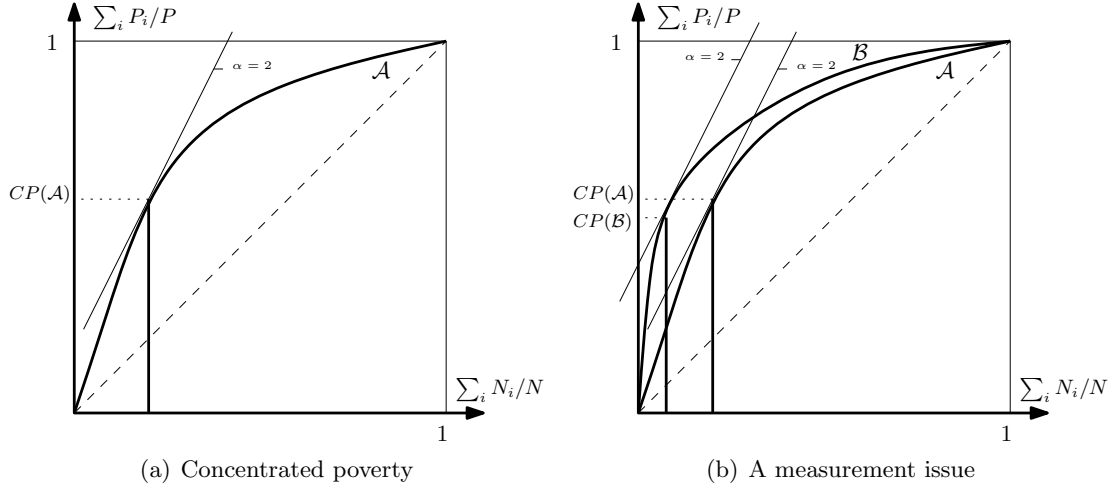
by $\mathcal{A} = \{P_i^{\mathcal{A}}, N_i^{\mathcal{A}}\}_{i=1}^n$, with average incidence $P^{\mathcal{A}}/N^{\mathcal{A}}$. In what follows, a configuration always represents a city in a given year, and we use superscripts to indicate a specific urban poverty configuration only when disambiguation is needed.

The ratio $\frac{P_i}{N_i}$ measures the incidence of poverty in neighborhood i . The ratio $\frac{P}{N}$ measures instead the incidence of poverty in the city, and is equivalent to the average of poverty incidences across neighborhoods, weighted by the respective population proportions, i.e. $\frac{P}{N} = \sum_{i=1}^n \frac{N_i}{N} \frac{P_i}{N_i}$. We use $\zeta \in [0, 1)$ to define a *urban poverty threshold*, which is a cutoff point that allows to identify the neighborhoods where poverty is over-concentrated. The urban poverty threshold incorporates an exogenous normative judgment about the level of poverty concentration that can be tolerated in a given neighborhood: when $\frac{P_i}{N_i} \geq \zeta$ then poverty in neighborhood i exceeds the tolerance level and contributes to urban poverty. In this case, neighborhood i is addressed to as a highly concentrated poverty neighborhood. When $\zeta = 0$, tolerance is set to a minimum, indicating that every neighborhood of the city contributes to generate urban poverty.

For a given urban poverty threshold ζ , neighborhoods can be ranked by poverty incidence in non-increasing order: $\frac{P_1}{N_1} \geq \frac{P_2}{N_2} \geq \dots \geq \frac{P_z}{N_z} \geq \zeta > \frac{P_{z+1}}{N_{z+1}} \geq \dots \frac{P_n}{N_n}$. For simplicity, labels $1, 2, \dots, n$ are assumed to coincide with the ranks of neighborhoods, ordered by non-increasing poverty incidence. Among all neighborhoods in the city, z identifies the neighborhood where poverty incidence is the closest to the urban poverty threshold. The neighborhood z serves as a benchmark. In fact, poverty is over-represented in neighborhoods $i \in \{1, \dots, z\}$.

In this paper, the urban poverty threshold is exogenously given and represents a normative stance about the maximum level of poverty which can be tolerated in a neighborhood without triggering poverty concentration. The Census Bureau, for instance, makes use of the 20% and 40% thresholds to identify places where poverty is highly concentrated and ghettos, respectively. As a result, if a city displays higher poverty incidence on average than another, then that city should also display larger deviations from the urban poverty threshold, hence larger urban poverty: even if poor people are evenly spread across neighborhoods, residents in the first city have larger chances to be exposed to poverty in their neighborhoods compared to residents in the second city (for a discussion, see Ravallion

Figure 1: Urban poverty curve and concentrated poverty



Note: CP index values (vertical black solid lines) for two configurations \mathcal{A} and \mathcal{B} .

and Chen 2011).

2.2 Concentrated poverty and its critical aspects

A convenient way to represent the distribution of the poor population in the city is to plot the cumulative proportion of the poor against the proportion of the overall population living in the neighborhoods displaying higher incidence of poverty, i.e. ranked by decreasing $\frac{P_i}{N_i}$. The cumulative proportion of poor people in neighborhood j is given by $\sum_{i=1}^j \frac{P_i}{P}$ and the cumulative proportion of residents therein is $\sum_{i=1}^j \frac{N_i}{N}$. Consider plotting points with coordinates $\left(\sum_{i=1}^j \frac{N_i}{N}, \sum_{i=1}^j \frac{P_i}{P}\right)$ with $j = 1, \dots, n$ on a graph. The curve starting from the origin and interpolating these points is a concentration curve denoted the *urban poverty curve*. The urban poverty curve of an hypothetical configuration \mathcal{A} is reported in panel (a) of Figure 1. Its graph is concave and always lies above the unit square diagonal, implying that in configuration \mathcal{A} there are neighborhoods with poverty incidence smaller than $\frac{P}{N}$ and other neighborhoods with poverty incidence greater than $\frac{P}{N}$.⁴

The lack of intersections of urban poverty curves is a natural criterion to rank distributions by the degree of urban poverty they display. If the urban poverty curve of

⁴This curve can be interpreted as the Lorenz curve of the distribution of poor population proportions $\frac{P_i}{N_i}$ across the city neighborhoods, each weighted by $\frac{N_i}{N}$. The curve of a configuration in which poor people are evenly spread across neighborhoods of the city, that is $\frac{P_i}{N_i} = \frac{P}{N}$ for every neighborhood i , coincides with the unit square diagonal. For simplicity, we assume that the city has many neighborhoods that differ in terms of poverty shares, so that the urban poverty curve appears smooth.

configuration \mathcal{B} lies nowhere below and somewhere above that of \mathcal{A} , then any proportion of the population living in high-poverty neighborhoods in \mathcal{B} is systematically exposed to a larger fraction of poverty than the corresponding population proportion in \mathcal{A} .

In our graphical analysis, we always assume that the distributions under comparison display the same poverty incidence, i.e. $\frac{P^{\mathcal{A}}}{N^{\mathcal{A}}} = \frac{P^{\mathcal{B}}}{N^{\mathcal{B}}}$. In this case, the urban poverty threshold could be expressed in relative terms as $\zeta = \alpha \frac{P}{N}$, where $\alpha \geq 0$ is a parameter expressing a normative view about sensitivity of urban poverty to the incidence of poverty in the city. Larger values of α imply that urban poverty evaluations should focus on neighborhoods where poverty is highly concentrated. The coefficient α straightforwardly relates to the urban poverty curve. For instance, in a city with $\frac{P^{\mathcal{A}}}{N^{\mathcal{A}}} = 0.2$, one can set $\alpha = 2$ to have $\zeta = 0.4$. On the graph, the coefficient α gives the slope of a line tangent to the urban poverty curve, as in Figure 1, panel a). The tangent point identifies the neighborhood z displaying poverty incidence of about $\zeta = 2 \frac{P^{\mathcal{A}}}{N^{\mathcal{A}}}$, the urban poverty threshold.⁵

Urban poverty curves are also related to the measurement of *concentrated poverty*, which is identified by the index $CP(\mathcal{A}) := \sum_{i=1}^z \frac{P_i}{P}$. The index coincides with the level of the curve at abscissa $\sum_{i=1}^z \frac{N_i}{N}$. Graphically, it is identified by the length of the vertical line segment on the same figure. The index CP measures the proportion of poor people who live in high-poverty neighborhoods, defined according to the threshold ζ . According to the American census, concentrated poverty corresponds to the proportion of poor residents that live in census tracts where at least 20% or 40% of inhabitants fall below the urban poverty threshold (i.e., $\zeta = 0.2$ or $\zeta = 0.4$ respectively).

The concentrated poverty index misses some important aspects of the distribution of poverty across the city neighborhoods and, as a consequence, it may rank cities inconsistently with non-intersecting urban poverty curves. Panel b) of Figure 1 draws an example. In the figure we consider two configurations \mathcal{A} and \mathcal{B} where $\frac{P^{\mathcal{B}}}{N^{\mathcal{B}}} = \frac{P^{\mathcal{A}}}{N^{\mathcal{A}}}$. The distribution of poverty across the neighborhoods of city \mathcal{B} is more uneven than that in city \mathcal{A} , in the sense that in \mathcal{B} a larger fraction of the poor population is concentrated in high poverty neighborhoods, compared to \mathcal{A} . As a consequence, the urban poverty curve of the former

⁵To see this, denote with δx and δy variations in the coordinates of the urban poverty curves on the horizontal and vertical axis. Moving along the curve from the tangency point implies $\delta x = N_z/N$ and $\delta y = P_z/P$ which gives the slope of the curve $\delta y/\delta x = \frac{P_z}{N_z} \frac{N}{P} \approx \alpha$.

lies always above that of the latter. Nonetheless, $CP(\mathcal{B}) < CP(\mathcal{A})$ for $\alpha = 2$.

In this paper, we introduce a new urban poverty index which is inspired by social welfare and inequality analysis and is consistent with the ranking of configurations predicted by non-intersecting urban poverty curves. The index we study compounds, with proper normative weights, two aspects of the spatial distribution of poverty: on the one hand, the extent of poverty incidence in places where poverty is highly concentrated; on the other hand, aspects of the distribution of poverty among high poverty neighborhoods as well as across high and low poverty neighborhoods. Both components positively contribute to urban poverty. While the former component has to do with the incidence of poverty among neighborhoods $1, \dots, z$, the latter component captures inequality in the distribution of poverty proportions $\frac{P_1}{N_1}, \dots, \frac{P_z}{N_z}$ and is related to the *Gini coefficient* $G(.; \zeta)$ defined here:⁶

$$G(.; \zeta) := \frac{1}{2 \sum_{i=1}^z P_i / \sum_{i=1}^z N_i} \sum_{i=1}^z \sum_{j=1}^z \frac{N_i N_j}{(\sum_{i=1}^z N_i)^2} \left| \frac{P_i}{N_i} - \frac{P_j}{N_j} \right|. \quad (1)$$

Our contribution relates to other contributions which examine weaknesses of concentrated poverty, already identified by Massey and Eggers (1990). They also suggest valuing the intensity *and* the distribution of poverty in the city as relevant aspects generating the double burden of poverty. The approach they propose, based on mixtures of dissimilarity and interaction indices, is interesting and related to the urban poverty curve ordering, but it is not based on normative grounds. In the next section, we provide a parsimonious axiomatic approach incorporating the idea that poverty concentration gives rise to a double burden of poverty.

3 A characterization result

3.1 Axioms

Denote the set of admissible urban poverty configurations $\Omega = \bigcup_{n \in \mathbb{N}} \Omega(n)$ with $\Omega(n) := \{P_i, N_i\}_{i=1}^n$. A urban poverty index is a function $UP(\mathcal{A}; \zeta) : \Omega \times [0, 1) \rightarrow \mathbb{R}_+$ assigning a non-negative real number to a configuration $\mathcal{A} \in \Omega(n)$, interpreted as the level of urban

⁶The index $G(.; \zeta)$ is related to the area comprised between the urban poverty curve and the unit square diagonal, up to a proportion $\sum_{i=1}^z \frac{N_i}{N}$ of the overall population. A related index adopted in income inequality analysis is discussed in Zoli (1999) and Andreoli (2018).

poverty in that configuration. We write $UP(., \zeta)$ to explicitly recall that the measurement of urban poverty is conditional on the exogenous urban poverty threshold ζ . We develop an axiomatic approach for the measurement of urban poverty.

A convenient way to incorporate concerns for the effects of a transfer of poverty operation on the measured level of urban poverty is to focus on urban poverty indices that explicitly depend on the *urban poverty shortfall* $\left(\frac{P_i}{N_i} - \zeta\right)$. The shortfall is non-negative in every neighborhood i where poverty is highly concentrated (that is, $i \leq z$) and increases as the proportion of poor residents $\frac{P_i}{N_i}$ grows.⁷

The first axiom introduces structure. It assumes that for any configuration $\mathcal{A} \in \Omega$ and urban poverty threshold $\zeta \in [0, 1)$, $UP(\mathcal{A}; \zeta)$ is a normalized (weighted) average of urban poverty shortfalls of each highly concentrated poverty neighborhood with $\frac{P_i}{N_i} \geq \zeta$, where each of these neighborhoods is weighted according to its position in the ordered distribution of poor neighborhoods and on population shares $\frac{N_1}{N}, \dots, \frac{N_n}{N}$. The whole measure is scaled according to a normalization factor that depends on the aggregate statistics. The aggregate normalization and the neighborhoods weighting function are continuous in their arguments. Let Δ_n denote the unit simplex in the n -dimensional space whose elements are all positive, that is $\Delta_n := \{d_1, d_2, \dots, d_n : d_i > 0, \sum_{i=1}^n d_i = 1 \text{ for } i = 1, 2, \dots, n\}$. To ease notation, we also denote $\bar{N}_z = \sum_{i=1}^z N_i$ and $\bar{P}_z = \sum_{i=1}^z P_i$, while \mathbb{N}_+ is the set of positive natural numbers.

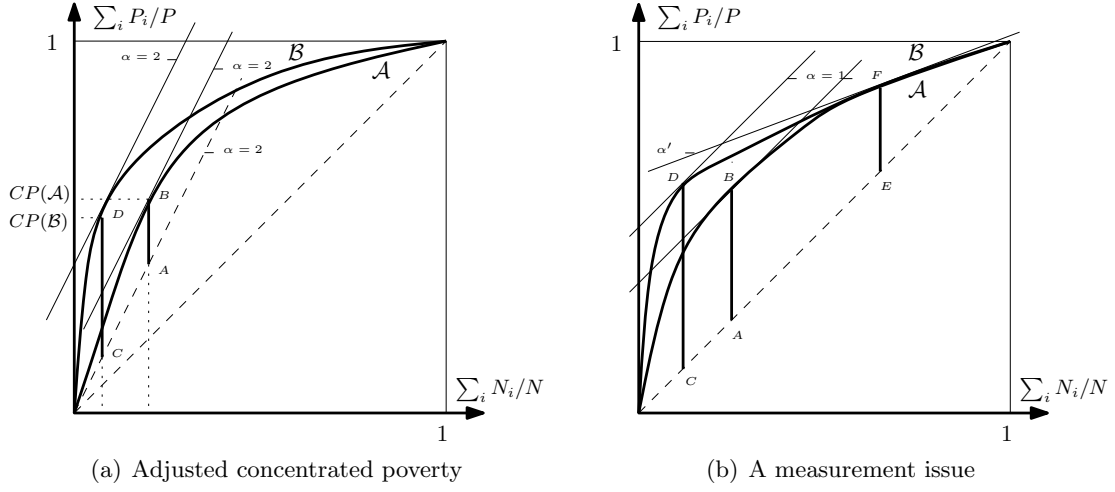
Axiom (AGG)regation. *$UP(., \zeta)$ satisfies AGG if for any $\mathcal{A} \in \Omega$ and $\zeta \in [0, 1)$ where $z \geq 1$, there exist a continuous function $A : [0, 1]^2 \rightarrow \mathbb{R}_+$ and a sequence of continuous functions $w_i : \Delta_n \rightarrow \mathbb{R}$ for each $i = 1, 2, \dots, n$ and each $n \in \mathbb{N}_+$ such that*

$$UP(\mathcal{A}; \zeta) := A\left(\frac{P}{N}, \frac{\bar{P}_z}{\bar{N}_z}\right) \cdot \sum_{i=1}^z \frac{N_i}{N} \cdot \left(\frac{P_i}{N_i} - \zeta\right) \cdot w_i\left(\frac{N_1}{N}, \dots, \frac{N_i}{N}, \dots, \frac{N_n}{N}\right).$$

Note that the AGG property holds if there exists at least a neighborhood with $\frac{P_i}{N_i} \geq \zeta$. The case where $\zeta > \frac{P_1}{N_1}$ will be considered when normalizing the index. According

⁷Notice that the urban poverty shortfall could never exceed zero if $\zeta = 1$. To avoid such situation, we maintain that $\zeta < 1$.

Figure 2: Urban poverty curve and adjusted concentrated poverty



Note: The adjusted concentrated poverty index CP^* corresponds to the vertical black solid line segments marked in the figure. In panel (a), the index is computed for both configurations \mathcal{A} (line segment AB) and \mathcal{B} (line segment CD) also reported in Figure 1 for $\alpha = 2$. In panel (b), the urban poverty curve of the hypothetical configuration \mathcal{B} lies nowhere below and somewhere above the curve of the hypothetical configuration \mathcal{A} .

to AGG, the function $A\left(\frac{P}{N}, \frac{\bar{P}_z}{\bar{N}_z}\right)$ is the aggregate normalization factor and the functions $w_i\left(\frac{N_1}{N}, \dots, \frac{N_i}{N}, \dots, \frac{N_n}{N}\right)$ denote the normative weights attached to the neighborhoods.

The AGG axiom imposes considerable structure, albeit it represents an encompassing model for a variety of indicators consistent with the ranking of urban poverty curves, stemming from choices of normalization and weighting parameters. Given the linearity of the components considered in AGG, the concerns about the poverty distribution across neighborhoods are formalized by the choice of the weighting functions $w_i(\cdot)$.

Let consider evaluations that are normalized by the incidence of poverty in the city, that is $A(P/N, \bar{P}_z/\bar{N}_z) = \frac{1}{P/N}$ and there are no concerns about the unequal distribution of poverty across neighborhoods, that is $w_i(N_1/N, \dots, N_n/N) = 1$ for every neighborhood i . This parametric choice retains exclusively concerns for the incidence of concentrated poverty and can be related to the concentrated poverty index as follows:

$$\begin{aligned} UP(\mathcal{A}; \zeta) &= \frac{1}{P/N} \sum_{i=1}^z \frac{N_i}{N} \left(\frac{P_i}{N_i} - \zeta \right) = \frac{\bar{P}_z - \zeta \bar{N}_z}{P} \\ &= CP(\mathcal{A}) - \zeta \frac{N}{P} \sum_{i=1}^z \frac{N_i}{N} =: CP^*(\mathcal{A}; \zeta) \end{aligned} \quad (2)$$

The result (2) shows that the index CP is consistent with AGG only up to an additive

correction factor $\zeta \frac{N}{P} \sum_{i=1}^z \frac{N_i}{N}$, which gives the *adjusted concentrated poverty* index CP^* . Similarly to the concentrated poverty measure, the index CP^* is related to the urban poverty curves. Differently from CP , the index CP^* always ranks configurations consistently with the ordering produced by non-intersecting urban poverty curves. This is illustrated in Figure 2, where we consider the special case in which $\zeta = \alpha \frac{P}{N}$, which gives $CP^*(\mathcal{A}; \zeta) = CP(\mathcal{A}) - \alpha \bar{N}_z$.

In panel (a) of Figure 2 we show the same urban poverty curves as in Figure 1, and we denote with bold solid lines the adjusted concentrated poverty indices $CP^*(\mathcal{A}; \zeta)$ (segment AB) and $CP^*(\mathcal{B}; \zeta)$ (segment CD).⁸ The adjusted concentrated poverty index ranks $CP^*(\mathcal{B}; \zeta) > CP^*(\mathcal{A}; \zeta)$, coherently with the ordering of configurations induced by the urban poverty curves. Since every urban poverty curve is concave and lies above the diagonal, the index CP^* is always positive and bounded above by CP .

While the index CP^* can be regarded to as a natural extension of the CP index, it is far from being an ideal measure of urban poverty for a generic configuration, for at least two reasons. First, the index measures the degree of concentration of poverty by focusing on a particular point of the urban poverty curve. Hence, there are cases in which the index may not be able to rank configurations even if they are unambiguously ordered by the urban poverty curves. Panel (b) in Figure 2 reports one of such cases.⁹

The second critical aspect of CP^* is that the index does not address heterogeneity in the concentration of poor individuals across the city's neighborhoods. There are two potential sources of heterogeneity. The first source is due to heterogeneity in $\frac{P_i}{N_i}$ ratios for neighborhoods $i \leq z$. When these ratios are homogenous across neighborhoods where poverty is concentrated, i.e., $\frac{P_1}{N_1} = \dots = \frac{P_z}{N_z} \geq \zeta$, the index CP^* is a sufficient statistic for urban poverty. If they are not, the index CP^* may rank as indifferent configurations that can be unambiguously ranked according to the urban poverty curve (see Figure 3).¹⁰

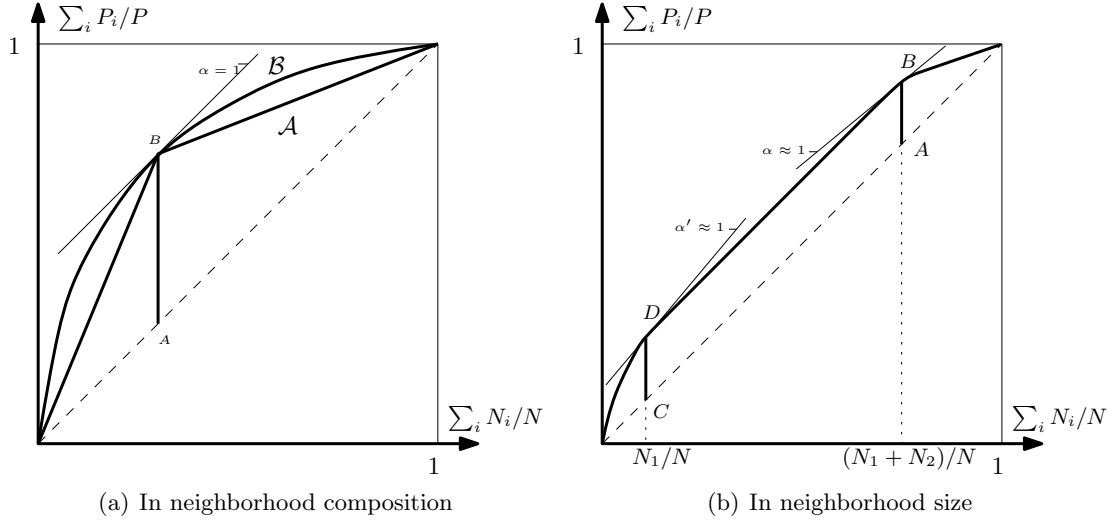
The second source of heterogeneity is due to the distribution of demographic sizes of the

⁸To see this, note that the length of the line segments starting from points A and C and intersecting the horizontal axis is $\alpha \bar{N}_z^A$ and $\alpha \bar{N}_z^B$, respectively.

⁹The curve of configuration \mathcal{B} lies above that of \mathcal{A} almost everywhere. For $\alpha = 1$, $CP^*(\mathcal{B}; P/N) > CP^*(\mathcal{A}; P/N)$. For $\zeta' = \alpha' \frac{P}{N}$ and α' small enough, however, $CP^*(\mathcal{B}; \zeta') = CP^*(\mathcal{A}; \zeta')$ and the two configurations become indistinguishable despite a larger fraction of the poor population of \mathcal{B} is concentrated in poor neighborhoods compared to \mathcal{A} .

¹⁰The graph in panel (a), Figure 3, provides an example where urban poverty is unambiguously larger in configuration \mathcal{B} than in configuration \mathcal{A} for $\alpha = 1$, but $CP^*(\mathcal{B}; P/N) = CP^*(\mathcal{A}; P/N)$.

Figure 3: Adjusted concentrated poverty and neighborhood structure heterogeneity



Note: Adjusted concentrated poverty measures for $\alpha = 1$ (i.e., at a poverty threshold $\zeta = \frac{P}{N}$) are given by solid line segments AB in both graphs.

neighborhoods, $\frac{N_i}{N}$. The index CP^* is insensitive to marginal changes in the poverty threshold that are due to changes in the demographic size of the neighborhoods. Panel (b) of Figure 3 reports an example of a city with many small neighborhoods, with an aggregate population share of N_1/N , and one large neighborhood of size N_2/N with a proportion of poor people equal to that in the population as a whole (i.e., P/N). The adjusted concentrated poverty measure is unaffected by small changes in the poverty threshold from ζ to ζ' . While this property of CP^* is appealing in some cases, it also implies that concentrated poverty evaluations neglect the size effects of the population that is actually exposed to poverty in the neighborhood of residence. In the figure, a large proportion of the population, $(N_1 + N_2)/N$, is concerned with concentrated poverty when the poverty threshold is $\zeta = \alpha \frac{P}{N}$, whereas only a minor share of the population seems to be exposed to high poverty when the poverty threshold marginally reduces to $\zeta = \alpha' \frac{P}{N}$.¹¹

More structure is needed in order to address distributional concerns. We consider additional axioms, characterizing the behaviour of any urban poverty measure vis-à-vis the effects of meaningful transformations of the data that affect heterogeneity. When paired with AGG, these axioms characterize the weighting scheme.

¹¹For a poverty threshold with α marginally larger than 1, the adjusted concentrated poverty index is the segment AB . For a poverty threshold with α' marginally smaller than 1, the adjusted concentrated poverty index, now identified by the segment CD , does not change.

The next axiom introduces a form of invariance of urban poverty measures with respect to the demographic structure of neighborhoods. To do so, we introduce a new operation, denoted by the *neighborhood splitting*, which reshapes the demographic size and geographic boundaries of any neighborhood i by splitting i into two new neighborhoods i' and i'' of smaller geographic and demographic size. We postulate invariance of the urban poverty index to any split operation or sequence thereof. This postulate owes its normative appeal to replication invariance properties formulated in inequality (Atkinson 1970, Cowell 2000) and segregation analysis (Andreoli and Zoli 2014).

Axiom *INV-S: INVariance to neighborhood Splitting.* $UP(., \zeta)$ satisfies *INV-S* if for any $\mathcal{A} \in \Omega(n)$ and $\mathcal{A}' \in \Omega(n+1)$ such that \mathcal{A}' is obtained from \mathcal{A} by splitting neighborhood i into two neighborhoods i' and i'' with $N_i = N_{i'} + N_{i''}$, $P_i = P_{i'} + P_{i''}$ and $\frac{P_i}{N_i} = \frac{P_{i'}}{N_{i'}} = \frac{P_{i''}}{N_{i''}}$, then $UP(\mathcal{A}; \zeta) = UP(\mathcal{A}'; \zeta)$.

Every urban poverty index should obey a simple *principle of transfers*, stating that urban poverty in configuration \mathcal{A} should be smaller than in configuration \mathcal{A}' whenever \mathcal{A}' is obtained from \mathcal{A} by a (regressive) transfer of poor people from a neighborhood with a lower poverty incidence to a neighborhood with a higher poverty incidence which is paired by a transfer in the opposite direction of non-poor people. The population size and the ranking of each neighborhood are unaffected by the transfer.

Inequality aversion with respect to poverty incidence of the neighborhoods is formalized by imposing next axiom. The axiom invokes a *principle of transfers*, stating that urban poverty in configuration \mathcal{A} should be smaller than in configuration \mathcal{A}' whenever \mathcal{A}' is obtained from \mathcal{A} by a (regressive) transfer of poor people from a neighborhood with a lower poverty incidence to a neighborhood with a higher poverty incidence which is paired by a transfer in the opposite direction of the same number of non-poor people. The population size and the ranking of each neighborhood are unaffected by the transfer.

Axiom *Principle of (TRAN)sfers.* $UP(., \zeta)$ satisfies *TRAN* if for $\mathcal{A}, \mathcal{A}' \in \Omega(n)$ there exist $i, j \in \{1, 2, \dots, n\}$ where $i < j \leq z$, such that $N_k^{\mathcal{A}'} = N_k^{\mathcal{A}}$, $P_k^{\mathcal{A}'} = P_k^{\mathcal{A}}$ for $k \in \{1, 2, \dots, n\}$, $k \neq i, j$, and $N_i^{\mathcal{A}'} = N_j^{\mathcal{A}'} = N_i^{\mathcal{A}} = N_j^{\mathcal{A}}$, $P_i^{\mathcal{A}'} = P_i^{\mathcal{A}} + \varepsilon$, $P_j^{\mathcal{A}'} = P_j^{\mathcal{A}} - \varepsilon$, for $\varepsilon > 0$ s.t. $P_h^{\mathcal{A}'} / N_h^{\mathcal{A}'} \in [0, 1]$, and $P_h^{\mathcal{A}'} / N_h^{\mathcal{A}'} \geq P_{h+1}^{\mathcal{A}'} / N_{h+1}^{\mathcal{A}'}$ for all $h \in \{1, 2, \dots, n-1\}$, then

$$UP(\mathcal{A}; \zeta) \leq UP(\mathcal{A}'; \zeta).$$

Note that in the definition of TRAN it is assumed that both i and j are highly concentrated poverty neighborhoods. However, it could be the case that because of the transfer the poverty incidence in neighborhood j falls below ζ (which is set exogenous), implying that $z' = z - 1$ (since z is endogenous). According to AGG, a similar transfer if implemented among neighborhoods that are not with highly concentrated poverty and that remain as such, does not modify $UP(\cdot; \zeta)$.

An urban poverty index satisfies TRAN when it ranks distributions consistently with non-intersecting urban poverty curves, since an operation underlying TRAN always implies an upward shift of the curve. The concentrated poverty index may violate TRAN, insofar a movement of poor people from a lower poverty neighborhood towards a higher poverty neighborhood can reduce concentrated poverty. A numeric example clarifies this point. Consider a city with $n = 3$ neighborhoods, where the urban poverty configuration \mathcal{A} is $(N_1, N_2, N_3) = (10, 10, 10)$ and $(P_1, P_2, P_3) = (7, 5, 3)$, implying $P/N = 15/30 = 0.5$. For a urban poverty threshold equal to $\zeta = 0.4$, we have that $z = 2$ (neighborhoods 1 and 2 have poverty incidences equal to $7/10 = 0.7$ and $5/10 = 0.5$, respectively) and $CP = 12/15$. Suppose now that two poor residents move from neighborhood 2 to neighborhood 1 and one poor resident switches from neighborhood 3 to neighborhood 1, and opposite transfers of non-poor people take place such that the population size of each neighborhood is not affected. The new urban poverty configuration \mathcal{A}' is such that $(P'_1, P'_2, P'_3) = (10, 3, 2)$ and $(N'_1, N'_2, N'_3) = (10, 10, 10)$. The overall poverty incidence for \mathcal{A}' is still $P'/N = 15/30 = 0.5$ but $z' = 1$, given that only neighborhood 1 has a poverty incidence greater than 0.4. We have that $CP' = 10/15 < CP$, i.e., concentrated poverty has decreased, despite the urban poverty curve of \mathcal{A}' lies above that of \mathcal{A} .

Next, we analyze the consequences of a *combined transfer*, generated by combining a regressive and a progressive transfer (i.e. a transfer of opposite sign obtained by setting $\varepsilon < 0$) of similar proportions of poor and non-poor individuals occurring in high poverty neighborhoods (that is, only across neighborhoods $1, \dots, z$). In our setting, any combined transfer does not affect poverty incidence in high poverty neighborhoods, but only its distribution. For ease of exposition, we assume that combined transfers always occur on

high poverty neighborhoods that occupy adjacent positions.

Different views may prevail when analyzing the effects of combined transfers on urban poverty. When a regressive transfer (involving neighborhoods i and $i + 1$) takes place *earlier* than a progressive transfer (involving neighborhoods j and $j + 1$) in the ranking of neighborhoods (that is, $i < j$), then the burden of concentrated poverty shifts more heavily on extreme poverty neighborhoods, while the population in some high poverty neighborhoods is relieved from it. Since poverty gets even more concentrated in extreme poverty neighborhoods, urban poverty is not bound to decrease. Conversely, a similar argument leads to conclude that when a progressive transfer is followed by a regressive one, urban poverty cannot increase after the transfer. The following axiom takes a neutral stance with respect to the effects of a combined transfer on the distribution of poverty across high poverty neighborhoods, and hence on urban poverty.

Axiom INV-T: (INV)ariance w.r.t. combined (T)ransfers. Let $z \geq 2$, $UP(\cdot; \zeta)$ satisfies INV-T if for $\mathcal{A}, \mathcal{A}' \in \Omega(n)$ there exist $i, j \in \{1, 2, \dots, z - 1\}$ such that $N_k^{\mathcal{A}'} = N_k^{\mathcal{A}}$, $P_k^{\mathcal{A}'} = P_k^{\mathcal{A}}$ for $k \in \{1, 2, \dots, n\}$, $k \neq i, j, i + 1, j + 1$ and $N_i^{\mathcal{A}'} = N_j^{\mathcal{A}'} = N_{i+1}^{\mathcal{A}'} = N_{j+1}^{\mathcal{A}'} = N_i^{\mathcal{A}} = N_j^{\mathcal{A}} = N_{i+1}^{\mathcal{A}} = N_{j+1}^{\mathcal{A}}$, $P_i^{\mathcal{A}'} = P_i^{\mathcal{A}} + \varepsilon$, $P_{i+1}^{\mathcal{A}'} = P_{i+1}^{\mathcal{A}} - \varepsilon$, $P_j^{\mathcal{A}'} = P_j^{\mathcal{A}} - \varepsilon$, $P_{j+1}^{\mathcal{A}'} = P_{j+1}^{\mathcal{A}} + \varepsilon$ for $\varepsilon > 0$ s.t. $P_h^{\mathcal{A}'} / N_h^{\mathcal{A}'} \in [0, 1]$, and $P_h^{\mathcal{A}'} / N_h^{\mathcal{A}'} \geq P_{h+1}^{\mathcal{A}'} / N_{h+1}^{\mathcal{A}'}$ for all $h \in \{1, 2, \dots, n - 1\}$, and $P_{i+1}^{\mathcal{A}'} / N_{i+1}^{\mathcal{A}'} \geq \zeta$, then $UP(\mathcal{A}; \zeta) = UP(\mathcal{A}'; \zeta)$.

Axiom INV-T postulates that the combined effect of transfers of population of poor people across adjacent highly concentrated neighborhoods with the same population size is not affecting urban poverty. This is the case irrespective of whether the progressive or regressive transfers of population of poor people take place between neighborhood with higher or lower poverty incidence. If $i = j$, INV-T is satisfied by definition because the two transfers cancel out, in all the other cases the impact of the transfer taking place between two adjacent neighborhoods in the ranking based on poverty incidence, is not affected by the position covered by the neighborhoods in that ranking.

Next, we introduce additional properties that define the cardinal features of the urban poverty indices. The first invariance axiom considers situations where the poverty threshold ζ is modified. It considers different effects of combined changes in P_i and in ζ . In order to

simplify the exposition, we assume that this invariance condition holds for $n = 2$. Axioms can be readily generalized.

Axiom INV-PL: (INV)ariance to (P)overty (L)ine modifications. $UP(., \zeta)$ satisfies INV-PL if for configurations $\mathcal{A}, \mathcal{A}' \in \Omega(2)$ such that $\frac{P_1^{\mathcal{A}}}{N_1^{\mathcal{A}}} \geq \zeta$ the following conditions hold:

(i) let $\zeta \in [0, 1)$, if there exists $\lambda > 0$ s.t. $P_i^{\mathcal{A}'} = \lambda P_i^{\mathcal{A}}$ and $N_i^{\mathcal{A}'} = N_i^{\mathcal{A}}$ for $i = 1, 2$ and $\lambda \zeta \in [0, 1)$, then $UP(\mathcal{A}; \zeta) = UP(\mathcal{A}'; \lambda \zeta)$.

(ii) let $\zeta \in (0, 1)$, if $\frac{P_2^{\mathcal{A}}}{N_2^{\mathcal{A}}} < \zeta$ and there exists $\theta \in \mathbb{R}$ s.t. $P_1^{\mathcal{A}'} / N_1^{\mathcal{A}'} = P_1^{\mathcal{A}} / N_1^{\mathcal{A}} + \theta$, $P_2^{\mathcal{A}'} / N_2^{\mathcal{A}'} = P_2^{\mathcal{A}} / N_2^{\mathcal{A}}$ and $\zeta + \theta \in [0, 1)$ then $UP(\mathcal{A}; \zeta) = UP(\mathcal{A}'; \zeta + \theta)$.

The two conditions in INV-PL require respectively that (i) if the number of poor individuals in each neighborhood is scaled by the same factor $\lambda > 0$ and similarly also the threshold ζ scaled by the same factor then UP is not affected, (ii) if one neighborhood exhibits a high concentration of poverty while the other does not, if both the poverty incidence in the first neighborhood and the poverty threshold change by the same amount and the overall poverty incidence in the population is not affected, then UP does not change.

The two invariance conditions imply respectively that what matters are the ratios between P_i / N_i and ζ and that the differences between P_i / N_i and ζ are informative only if non-negative for a given level P / N of average poverty in the population.

Next property requires that if poverty increases proportionally in each neighborhood then UP should not decrease. Along with TRAN, these two axioms incorporate the features of transformations of the data that rise poverty in high poverty neighborhoods while unambiguously rising the double burden due to poverty concentration. As a consequence, urban poverty is bound to rise.

Axiom (MON)otonicity: $UP(., \zeta)$ satisfies MON if for configurations $\mathcal{A}, \mathcal{A}' \in \Omega(2)$ where $z = 2$, and there exists $\nu > 1$ s.t. $P_i^{\mathcal{A}'} = \nu P_i^{\mathcal{A}}$ and $N_i^{\mathcal{A}'} = N_i^{\mathcal{A}}$ for $i = 1, 2$ then $UP(\mathcal{A}; \zeta) \leq UP(\mathcal{A}'; \zeta)$.

To conclude, we set an axiom that quantifies the lower bound of the UP index.

Axiom (NOR)malization: for any $\mathcal{A} \in \Omega$ and $\zeta \in (0, 1)$ where $\zeta > \frac{P_1}{N_1}$, $UP(\mathcal{A}; \zeta) := k \in \mathbb{R}$ where $k := \inf\{UP(\mathcal{A}'; \zeta) : \mathcal{A}' \in \Omega \text{ with } z \geq 1\}$.

The NOR condition specifies the value of the index for configurations where the poverty incidence in each neighborhood is below the threshold ζ . In this case the value of the index is constant for each configuration and coincides with the infimum of all values of $UP(\cdot; \zeta)$ that could be obtained in any other alternative configuration in Ω where at least for one neighborhood the poverty incidence is not below ζ .

When assuming AGG and considering configurations in which poverty incidence is evenly distributed across high-poverty neighborhoods (i.e. $P_i/N_i = \bar{P}_z/\bar{N}_z$ for all $i \leq z$), axioms NOR and TRAN jointly imply that urban poverty evaluations should be normalized by the average poverty incidence. This leads to the main result.

3.2 Main result and discussion

Theorem 1 Let $\mathcal{A} \in \Omega$, $\zeta \in [0, 1)$, $UP(\cdot; \zeta)$ satisfies AGG, INV-S, INV-T, INV-PL, MON, TRAN and NOR if and only if there exist $\beta, \gamma \geq 0$ such that:

$$UP(\mathcal{A}; \zeta) := \beta \cdot \frac{\bar{P}_z - \zeta \bar{N}_z}{P} + \gamma \cdot \left(\frac{\bar{N}_z}{N}\right) \cdot \frac{\bar{P}_z}{P} \cdot G(\mathcal{A}, \zeta) + \gamma \cdot \left(\frac{N - \bar{N}_z}{N}\right) \cdot \frac{\bar{P}_z - \zeta \bar{N}_z}{P} \quad (3)$$

if $z \geq 1$, otherwise $UP(\mathcal{A}; \zeta) = 0$.

The proof of the theorem is in Appendix A.1, where we also demonstrate the independence of the axioms. Theorem 1 shows that the urban poverty axioms characterize exactly one parametric urban poverty index. For any given poverty threshold ζ , this index depends on three components and two parameters (besides the urban poverty threshold), capturing respectively the incidence (β) and the aversion to dispersion (γ) of poverty. The three components of the index contribute positively to the measured level of urban poverty. Their relative importance depends on the weight they receive. The first component, weighted by β , captures the average excess of poverty (with respect to the tolerance level ζ) in high poverty neighborhoods and it is related to the adjusted concentrated poverty index. In fact, when evaluations do not express distributional concerns, the index CP^* becomes the relevant measure of urban poverty.

Corollary 2 *Let $\mathcal{A} \in \Omega$, $\zeta \in [0, 1)$ and $\gamma = 0$, then $UP(\mathcal{A}; \zeta) = \beta \cdot CP^*(\mathcal{A}; \zeta)$.*

The second component, weighted by γ , captures aspects of inequality in the distribution of poverty across high poverty neighborhoods through the Gini index $G(\mathcal{A}, \zeta)$. The third component captures inequality in the distribution of the poor population between high poverty (with weight $\frac{\bar{N}_z}{N}$) and low poverty (with weight $1 - \frac{\bar{N}_z}{N}$) neighborhoods.¹² The inequality component of urban poverty measures the relative extent of dissimilarity between the actual distribution of poor residents across neighborhoods of the city and the distribution of the population of poor and non-poor residents across the same neighborhoods (see Andreoli and Zoli 2014).

Each component of urban poverty captures a specific feature of changes in urban poverty distribution. For instance, a regressive transfer of poor population occurring between neighborhoods where poverty is highly concentrated does only affect the distribution of poverty in those neighborhoods. This effect is reflected on the component $G(\cdot; \zeta)$. Consider instead a situation where poverty is evenly distributed in high poverty neighborhoods (i.e. $\frac{P_i}{N_i} = \frac{\bar{P}_z}{\bar{N}_z}$ for each $i \leq z$, implying $G(\cdot; \zeta) = 0$) and assume that some poor individuals in z are evenly distributed across neighborhoods $1, \dots, z - 1$, so that poverty rises uniformly across these neighborhoods. This change rises both the first component (poverty incidence rises) and the third component (poverty is more unevenly distributed across high and low poverty neighborhoods) of the index when such movement shifts z towards $z' = z - 1$. Extending the weighting scheme over $\{1, \dots, n\}$, as postulated in axiom AGG, guarantees that urban poverty evaluations are robust with respect to marginal changes in the distribution of poverty around the (exogenous) urban poverty threshold ζ which may nonetheless affect the (endogenous) threshold neighborhood z . Finally, consider again a situation in which poverty is evenly distributed and new non-poor individuals flew in low poverty neighborhoods. While this change does not affect \bar{P}_z , \bar{N}_z , P and z , it increases N and therefore it increases the segregation of poverty in high poverty neighborhoods by rising the number $N - \bar{N}_z$ of residents that are least exposed to high poverty concentration. Arguably, this change rises inequalities between neighborhoods and rises measured urban

¹²The third component is the value of the absolute Gini coefficient of a counterfactual distribution of poverty incidence shortfalls across two types of neighborhoods: high poverty neighborhoods receive the average shortfall, whereas shortfall realizations in low poverty neighborhoods are $(\zeta - \zeta) = 0$. As such, this component captures segregation of poverty across high and low poverty neighborhoods.

poverty through the third component of the index.

Evaluations of urban poverty are conditional on the poverty threshold. When the tolerance threshold approaches zero, concerns for even small levels of poverty concentration rise. This allows to take into account the fact that increments of poverty in those neighborhoods where poverty is more concentrated prevents other people living in neighborhoods where the poor are under-represented to be exposed to the double burden of poverty. By setting the urban poverty threshold at $\zeta = 0$, concerns about the distribution of poverty are extended to all neighborhoods of the city. Notice that if $\zeta = 0$, then $z = n$, $\bar{N}_z = N$ and $\bar{P}_z = P$, it follows that:

Corollary 3 *Let $\mathcal{A} \in \Omega$ and $\zeta = 0$, then $UP(\mathcal{A}; 0) = \beta + \gamma \cdot G(\mathcal{A}; 0)$ with $\beta, \gamma \geq 0$.*

When the urban poverty threshold is inclusive of all neighborhoods of the city, the first component of urban poverty, measuring incidence, reaches its maximum level, whereas the urban poverty index becomes ordinal equivalent to the Gini inequality index $G(\mathcal{A}; 0)$.

We explore the decomposition properties of the Gini index to address some key measurement issues in urban poverty. A first issue is that urban poverty may be insensitive to the depth of poverty, insofar P_i is identified on the basis of an exogenous poverty threshold. In the American case, for instance, urban poverty patterns of extremely poor families (with equivalent income below 75% of the federal income poverty line) may differ from urban poverty of the average family in need (comprising all families with income below 200% of the federal income poverty line). The factor income decomposition in Shorrocks (1982) can be used to linearly decompose the urban poverty index into the contribution of different subgroups identified by varying the poverty threshold, such as for families in extreme (below 75%), severe (between 75% and 100%) or mild (between 100% and 200%) poverty. Such a decomposition may be useful in drawing empirical robustness checks.

A second issue concerns the possibility of factorizing longitudinal variations in urban poverty within the same city between two periods t and t' into the contribution of demographic growth, poverty growth and poverty relocation across neighborhoods. Based on Corollary 3, the focus should be on the quantity $\Delta UP = UP(\mathcal{A}') - UP(\mathcal{A}) = G(\mathcal{A}'; 0) - G(\mathcal{A}; 0)$ for the arbitrary selection $\gamma = 1$ and $\beta = 0$. The following section provides two relevant decomposition results.

4 Decomposing changes in urban poverty

This section builds on Corollary 3 and provides two additional corollaries to the main result, showing that changes in urban poverty can be decomposed linearly into convergence, re-ranking and poverty growth components, as well as into the contribution of changes taking place within or across spatial clusters of neighborhoods. Both decompositions are relevant for describing the dynamics of urban poverty.

4.1 Convergence, re-ranking and growth components of urban poverty

Consider a city, where each neighborhood i is observed in both periods t and t' . The overall population and the number of poor residents in i are denoted $N_i^{\mathcal{A}}$ and $P_i^{\mathcal{A}}$ in period t and $N_i^{\mathcal{A}'}$ and $P_i^{\mathcal{A}'}$ in period t' , respectively. Let c stand for the change in poverty incidence in the whole city, so that:

$$c := \left(\frac{P^{\mathcal{A}'}}{N^{\mathcal{A}'}} - \frac{P^{\mathcal{A}}}{N^{\mathcal{A}}} \right) / \frac{P^{\mathcal{A}}}{N^{\mathcal{A}}}.$$

The next corollary shows that changes in urban poverty can be linearly decomposed into the contribution of demographic (W), re-ranking (R) and convergence ($C \cdot E$) components.

Corollary 4 *Let $\mathcal{A}, \mathcal{A}' \in \Omega(n)$ represent respectively configurations in time t and in time t' , then $\Delta UP = G(\mathcal{A}'; 0) - G(\mathcal{A}; 0) = W + R + C \cdot E$, where $C = 1/(1 + c)$.*

Proof. See appendix A.2. ■

The first component captures the effect of changes in the *demographic weights* of neighborhoods, and is denoted by W . In empirical applications, it is generally the case that $\frac{N_i^{\mathcal{A}}}{N^{\mathcal{A}}} \neq \frac{N_i^{\mathcal{A}'}}{N^{\mathcal{A}'}}$ for some neighborhoods. The changes in demographic weights may have non-trivial effects on the accountancy of urban poverty changes. The demographic component contributes positively to changes in urban poverty ($W > 0$) if the demographic weights of the neighborhoods that are more dissimilar in terms of poverty incidence increase, whereas the weights of neighborhoods that are less dissimilar decrease. The component W isolates the effect of population changes from changes in the distribution of poverty across neighborhoods.

Component C captures the effect of the *change in poverty incidence* in the city. It measures the implication of a citywide expansion (or reduction) of poverty incidence on urban poverty, thus allowing to separate the contribution of a change in poverty incidence in the same proportion c for all neighborhoods¹³ from the contribution of neighborhood-specific changes in poverty incidence (occurring when poverty incidence changes across neighborhoods in a disproportionate way).

Components R and E measure different distributional effects due to disproportionate changes in neighborhood poverty rates. The component R measures the effect of re-ranking of the neighborhoods, based on poverty incidence, from t to t' . The component E measures the effect of convergence (or divergence) in poverty incidence among neighborhoods. Neighborhoods diverge when the poverty rates of neighborhoods with high (low) poverty incidence in t increase (decrease) faster than the poverty rates in low (high) poverty neighborhoods. In this case, $E > 0$ and urban poverty level increases by $C \cdot E$. Otherwise, $E < 0$.

The implications of convergence in poverty incidence on changes in urban poverty can be more complex. For instance, a strong convergence of poverty rates across neighborhoods may induce a re-ranking of neighborhoods in terms of poverty incidence, implying that the reduction in urban poverty due to convergence (i.e., $C \cdot E$) can be partially offset by the re-ranking effect measured by R , which is always non-negative. Component R is null when there is no re-ranking, and is positive when at least two neighborhoods exchange their positions in the ranking of neighborhoods by poverty rate from t to t' .¹⁴

The result in Corollary 4 is useful for decomposing additively the contribution of poverty incidence and demographic changes at neighborhood and city level on the dynamics of urban poverty. The decomposition displays advantages over other methods. First, the decomposition allows to factor out the effect of demographic changes (W) on urban poverty,

¹³For each neighborhood i , $\frac{P_i^{A'}}{N_i^{A'}} = (1 + c) \frac{P_i^A}{N_i^A}$

¹⁴We borrow the terminology from the analysis of panel income growth (Jenkins and Van Kerm 2016). Component E is computed by comparing the relative disparities between the neighborhood poverty rates in t and those in t' , under the assumption that the ranking of neighborhoods remains constant over time to that observed in t . The effect of E can be either magnified or mitigated by C , since the latter component reflects the change in citywide poverty incidence. For instance, the potential effect of a convergence in poverty incidence among neighborhoods ($E < 0$) is reduced when changes in neighborhoods poverty rates lead to increasing the citywide poverty incidence ($C < 1$).

thus disentangling the effect of changes in poverty from the effect of demographic shifts and growth across neighborhoods. Second, components R and $C \cdot E$ pick up specific aspects of changes in poverty concentration that cannot be inferred just by looking at ΔUP . For instance, consider two cities \mathcal{A} and \mathcal{B} displaying no changes in urban poverty ($\Delta UP^{\mathcal{A}} = \Delta UP^{\mathcal{B}} = 0$), with $R^{\mathcal{A}} = C^{\mathcal{A}} \cdot E^{\mathcal{A}} = 0$ for the first city, while $R^{\mathcal{B}} = -C^{\mathcal{B}} \cdot E^{\mathcal{A}} > 0$ for the second city. While the poor population is immobile in the first city \mathcal{A} , poverty deconcentrates in some neighborhoods and reconcentrates in others in the second city \mathcal{B} , despite the change does not imply a neat form of convergence in the degree of poverty concentration, but rather a shift of poverty across the neighborhoods of the city (large $R^{\mathcal{B}}$).

4.2 Spatial components of urban poverty

The index UP , ΔUP and their components are, by construction, invariant to changes in the spatial configuration of poverty within the city, and hence unaffected by the implicit degree of spatial association of poverty incidence across neighborhoods. Building on the Rey and Smith (2013) spatial decomposition of the Gini index, we obtain a two-term additive decomposition of the urban poverty index, in which a “neighborhood component” measures distributional changes originating from neighborhoods that are spatially close and a “non-neighborhood component” measures changes concerning neighborhoods that are not in spatial proximity. The two components reveal the contribution of spatial association to measured urban poverty. The spatial decomposition is conditional on the knowledge of a proximity matrix \mathbf{N} , whose generic binary element $n_{ij} \in [0, 1]$ indicates whether neighborhoods i and j are neighbors according to a given criterion. The matrix \mathbf{N} can be constructed from the data and is assumed fixed throughout the comparisons (in our setting, the spatial structure of a city does not change across time), but is specific to the city.

Corollary 5 *Let $\mathcal{A}, \mathcal{A}' \in \Omega(n)$ represent respectively configurations in time t and in time t' , then $\Delta UP = G(\mathcal{A}'; 0) - G(\mathcal{A}; 0) = (G_N(\mathcal{A}'; 0) + G_{nN}(\mathcal{A}'; 0)) - (G_N(\mathcal{A}; 0) + G_{nN}(\mathcal{A}; 0)) = (W_N + W_{nN}) + (R_N + R_{nN}) + C \cdot (E_N + E_{nN})$.*

Proof. See appendix A.3. ■

Corollary 5 delivers two important results. First, it shows that UP can be exactly decomposed into neighborhood N and non-neighborhoods nN components. When G_N is large relative to G_{nN} , most of inequality in urban poverty occurs in neighborhoods that are located in spatial proximity. In this case, high and low poverty neighborhoods tend to belong to the same spatial cluster. The converse holds when G_N is small compared to G_{nN} , in which case there is a positive spatial autocorrelation in the distribution of poverty among neighborhoods.

The clustering dimension of urban poverty is relevant for policy analysis for at least two reasons. First, spatial clustering of high poverty neighborhoods may decrease the likelihood of access to transportation, to the job market, to high-quality supply of public goods and definitely to economic and social opportunities for the residents, thus amplifying the double burden from poverty already experienced by neighborhood residents. Second, when clusters of high poverty neighborhoods overlap with administrative divisions of the territory, such as counties or school districts, more economically vulnerable residents might face poverty traps that extend their effects both on long-term poverty status of the residents as well as on inter-generational mobility prospects of the children living therein.

The second important result of Corollary 5 is that changes over time in urban poverty can be also linearly decomposed along the spatial dimension. In this way, we can disentangle the contribution of changes in poverty within clusters from that of changes across clusters, which are more relevant for understanding spatial drivers of urban poverty.

5 Patterns, trends and components of urban poverty in American cities: 1980-2014

5.1 Data

We use data from the U.S. Census Bureau to study patterns and trends of urban poverty in American cities. Data for 1980, 1990 and 2000 are from the decennial census Summary Tape File 3A. The STF 3A data come in the form of poverty counts at the census tract level. After 2000, the STF 3A files have been replaced with survey-based estimates of the income tables from the American Community Survey (ACS), which runs annually since

2001 on representative samples of the U.S. resident population. We focus on three waves of the 5-years ACS module: 2006-2010, 2010-2014 and 2012-2016. We interpret data from each wave as representative for the mid-interval year, i.e. 2008, 2012 and 2014 respectively. These years roughly correspond to the onset, the striking and the early aftermath of the Great Recession period (Jenkins, Brandolini, Micklewright and Nolan 2013, Thompson and Smeeding 2013).

Poverty incidence at the census tract level is measured by the number of individuals in families with total income below the poverty threshold, which varies by family size, number of children, and age of the family members.¹⁵ Poverty status is determined for all families (and, by implication, all family members).¹⁶ The census reports poverty counts at census tract level for various poverty thresholds. In this paper, we consider as poor all individuals living in households with income below the 100% federal income poverty line. In the supplemental appendix, we provide robustness checks for poverty status determined by equivalent family income below 75%, 100% (baseline) and 200% of the federal poverty line.

Following Andreoli and Peluso (2018), we consider the 2016 Census Bureau definition of American Metropolitan Statistical Areas (MSA) to group census tracts into cities. The number and geographic size of the census tracts vary substantially across time within the same MSA. Some census tracts experience demographic growth and are split into smaller tracts. Some other census tracts are consolidated to account for demographic shifts. While raw data allow to estimate urban poverty at the city level, they cannot be used to perform the decomposition exercise, insofar the definition of neighborhood is not constant over time. We resort on the Longitudinal Tract Data Base (LTDB), which provides crosswalk files to estimate population counts statistics within 2010 tract boundaries for any tract-

¹⁵Both Census 1990 and 2000 and ACS determine a family poverty threshold by multiplying the base-year poverty thresholds (1982) by the average of the monthly inflation factors for the 12 months preceding the data collection. The poverty thresholds in 1982, by size of family and number of related children under 18 years can be found on the Census Bureau web-site: <https://www.census.gov/data/tables/time-series/demo/income-poverty/historical-poverty-thresholds.html>. For a four persons household with two under-age children, the 1982 threshold is \$9,783. Using the inflation factor of 2.35795 gives a poverty threshold for this family in 2013 of \$23,067. If the disposable household income is below this threshold, then all four members of the household are recorded as poor in the census tract of residence, and included in the 2014 wave of ACS.

¹⁶Poverty status is also determined for persons not in families, except for inmates of institutions, members of the Armed Forces living in barracks, college students living in dormitories, and unrelated individuals under 15 years old.

level data that are available for prior years as well as in ACS for the period after 2010 (Logan, Xu and Stults 2014).¹⁷ We calculate poverty incidence in each census tract/year and then construct measures of urban poverty and concentrated poverty in high (i.e. where poverty incidence is above 20% of the resident population) and extreme (i.e. where poverty incidence is above 40% of the resident population) poverty neighborhoods.

The balanced panel enables us to further decompose changes in urban poverty in its underlying components for 395 American MSAs.¹⁸ Census tracts are geo-localized, implying that measures of proximity of these tracts can be further produced and used to disentangle the neighborhood and non-neighborhood components of urban poverty across all years and all MSAs.

5.2 Results

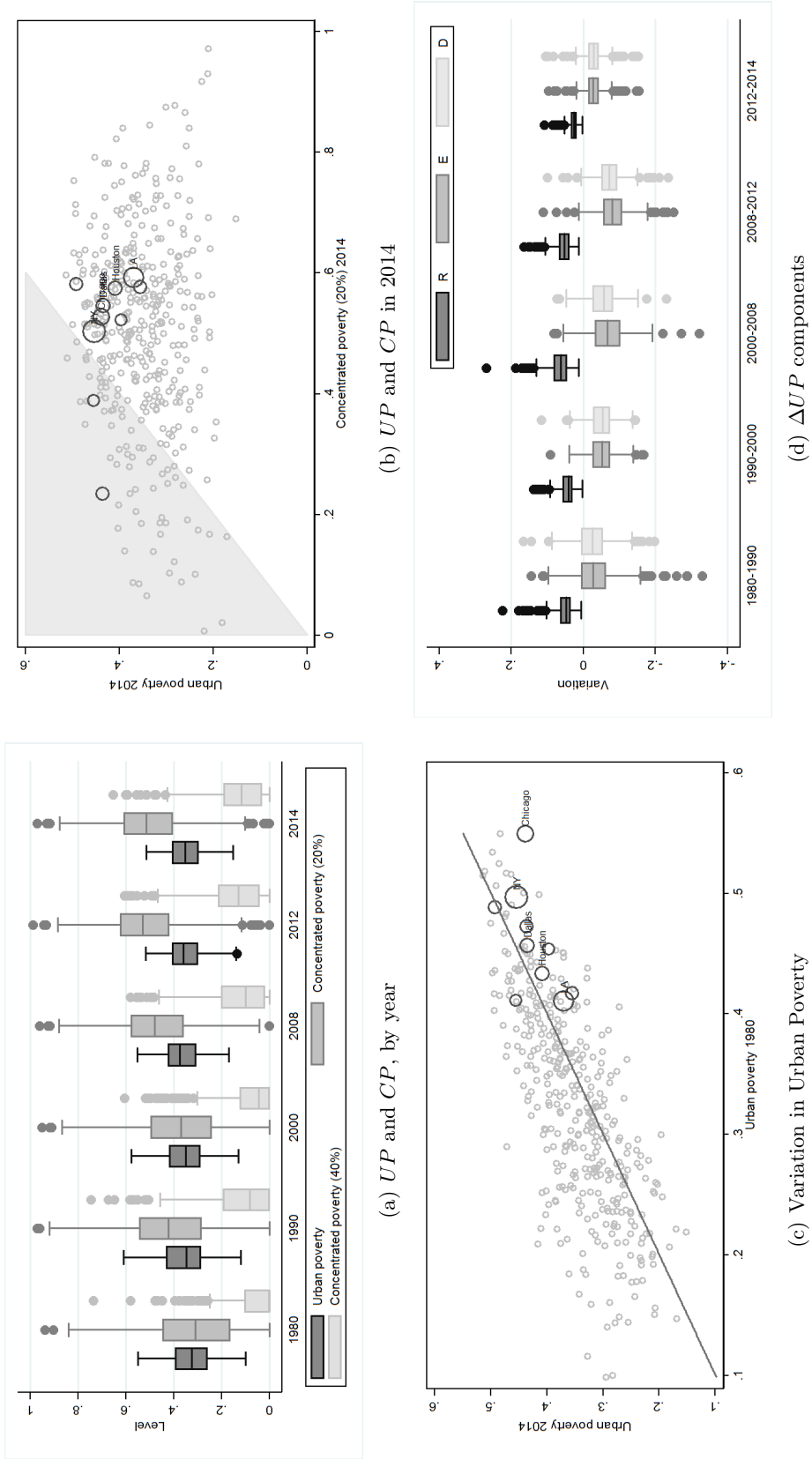
Panel a) of Figure 4 describes the levels and trends of urban poverty and concentrated poverty in American MSAs over 1980-2014. In line with the literature, we find that concentrated poverty is high in American cities, ranging from 26% to 51% on average over the period. Concentrated poverty has increased since the onset of the Great Recession, and it has remained stable in the aftermath. Conversely, the distribution of urban poverty among MSAs reveals a more stable pattern over the last 35 years we consider. Small changes in urban poverty may however be the outcome of the offsetting contributions of re-ranking and changes in disparities between census tract poverty rates. The decomposition in Corollary 4 is useful to separate these effects.

Panels b) and c) of Figure 4 display the extent of heterogeneity in the distribution of concentrated poverty and urban poverty over the whole period considered. Data suggest that concentrated poverty and urban poverty indices capture uncorrelated aspects of the urban distribution of poor. Larger metro areas, denoted by circles of larger size on the graph, display proportionally more concentrated poverty than urban poverty. Urban

¹⁷These files make use of re-weighting methods to assign each census and ACS year population to the exact census tract boundary defined in 2010 census. We obtain a balanced longitudinal dataset of census tracts for 395 American Metropolitan Areas (those with at least 10 census tracts according to 2010 census) for years 1980, 1990, 2000, 2008, 2012 and 2014.

¹⁸Figure 6 in the Appendix B displays urban poverty calculated on balanced longitudinal data against urban poverty calculated on raw data. Estimates of urban poverty based on the two methods largely coincide. Urban poverty estimates are also unrelated to the incidence of poverty in the city, as shown in Figure 7.

Figure 4: Urban poverty distribution among American MSAs, from 1980 to 2014



Note: Levels of urban poverty and concentrated poverty (concentration of poverty at neighborhood level at 20% and 40%), and of urban poverty components (R , E and $D = C \cdot E$), 1980-2014. Data for 395 selected MSAs. Solid line represents no changes in concentrated poverty or in urban poverty.

	Spatial independence, 1980			Spatial independence, 2014			Population (mln)	CT
	Reject 1%	Reject 10%	Accept	Reject 1%	Reject 10%	Accept	Avg.	Avg.
Q1	0.08	0.17	0.83	0.05	0.18	0.82	0.08	19
Q2	0.11	0.23	0.77	0.16	0.39	0.61	0.15	36
Q3	0.24	0.43	0.57	0.37	0.56	0.44	0.33	74
Q4	0.69	0.80	0.20	0.79	0.82	0.18	2.02	455
All	0.28	0.41	0.59	0.34	0.48	0.52	0.64	145

Table 1: Proportion of acceptances (p-values > 0.1) and weak (p-values < 0.1) and strong (p-values < 0.01) rejections of spatial independence assumption, based on Moran-I tests with order-1 nearest neighborhood spatial weighting matrices.

poverty is persistent over the period, with most MSAs grouped along the figure bisector. Panel d) of Figure 4 breaks down heterogeneity of year-to-year variation in urban poverty into its components, computed separately for each MSA.¹⁹ The little variability in urban poverty can be explained by the trends in its components R and $D := C \cdot E$. The component D is negative for a majority of MSAs in each sub-period, indicating that relative disparities in poverty incidence across census tracts have decreased over time. Such a pattern suggests convergence in neighborhood poverty and decreasing urban poverty. This equalizing effect is partially offset by the re-ranking component R , which indicates that initially low poverty census tracts catch up high poverty census tracts, indicating substantial re-ranking which contributes to rise urban poverty. The contribution of W is negligible.

Overall, the analysis of urban poverty suggests a major trend of convergence in poverty across American MSA neighborhoods. Poverty has grown everywhere in American MSAs after the Great Recession, but less so in high poverty neighborhoods, while concentrating into historically middle-class, low poverty neighborhoods.

We examine the decomposition of urban poverty into neighborhood and non-neighborhood components. A proximity matrix describing the spatial relations between census tracts is obtained for each city resorting on the notion of *critical cut-off neighborhood*, according to which two census tracts are neighbors if their distance is equal or less than a given cut-off distance.²⁰

¹⁹Summary statistics for the distribution of year-by-year variations in concentrated poverty, in urban poverty and in its components across American MSA, are reported in Table 2 in Appendix B.

²⁰We use the minimum cut-off distance criterion (Espa, Filipponi, Giuliani and Piacentino 2014), that defines the cut-off distance for a city as the minimum distance for which every census tract in the city has at least one neighbor. A minimum cut-off distance is set for each MSA in the sample, with the threshold varying across MSAs.

We use the Moran-I index to test for spatial dependence in urban poverty rates (setting spatial independence at the null) and register for each MSA the p-value of the test, computed separately in 1980 and 2014. In Table 1 we report proportions of cases of weak (at 10% significance level) and strong (at 1% significance level) rejections of the null hypothesis, alongside the proportion of acceptances (with p-value larger than 0.1). The Moran-I statistics can be highly influenced by the population size of the city and the number of neighborhoods. We hence report rejections and acceptance rates by quartiles (Q1,...,Q4) of MSAs ranked by population size.

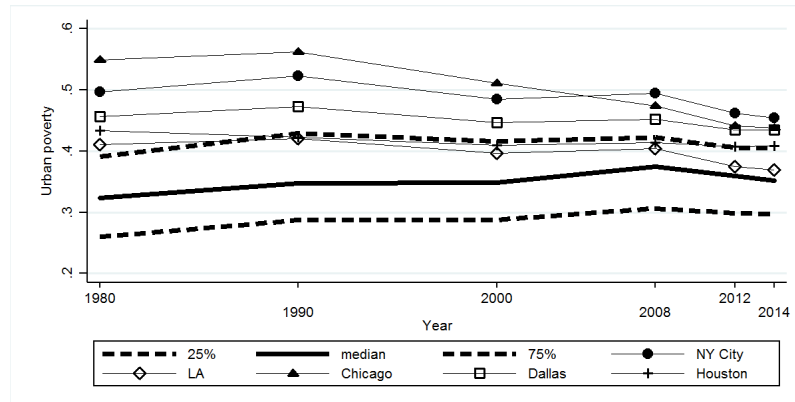
Results support the hypothesis of spatial independence for the Q1 and Q2 cities (with average population size smaller than 0.15mln). Patterns is less clear for the Q3 cities, where rates of rejections and acceptance of spatial independence in the occurrence of poverty are mixed, with weak rejection rates ranging from 43% in 1980 to 56% in 2014. For large MSAs (with about 2mln residents on average) included in Q4, data weakly reject the null hypothesis of spatial independence in about 80% of the cases (in both years alike) in favor of positive spatial autocorrelation. In these cities, neighboring census tracts tend to have similar poverty rates, thus rising the risk of presence of spatial poverty traps.

In Figure 5, we separately analyze the patterns of urban poverty in the largest five American MSAs and further decompose the changes of urban poverty into neighborhood and non-neighborhood components. Overall, we find that urban poverty has increased from 1980 to 1990, with largest MSAs in the top of urban poverty distribution. Urban poverty in largest cities has declined after 2000, slowly converging towards the rest of the MSAs we consider. This change is mostly driven by the non-neighborhood component of urban poverty (panel c) of Figure 5), which is generally high and explains most of urban poverty in these cities. The minor role of the neighborhood component in large MSAs confirms that census tract poverty rates are more similar among neighboring tracts than among non-neighboring tracts.

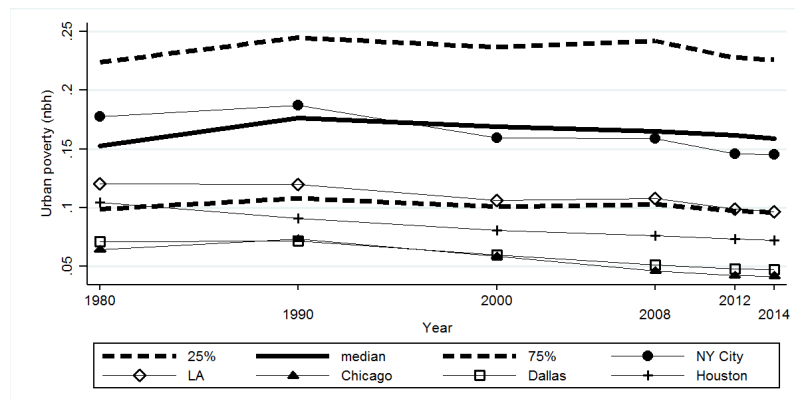
6 Concluding remarks and extensions

This paper introduces a parsimonious axiomatic approach to characterize a new parametric urban poverty measure. This measure weights the contribution of poverty incidence and

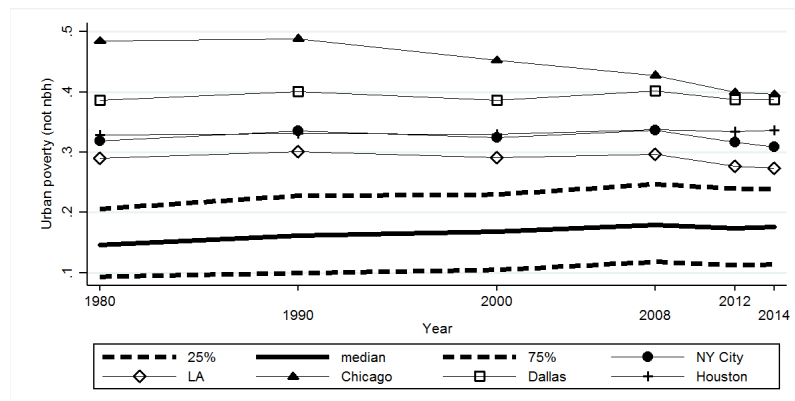
Figure 5: Urban poverty across American MSAs, 1980 to 2014



(a) *UP* by year



(b) Neighborhood component



(c) Non-neighborhood component

Note: Urban poverty levels and components of changes in urban poverty (R , E and $D = C \cdot E$) for median, top and bottom quartile cities in the sample, 1980-2014. Data for 395 selected MSA and five selected MSA (with largest population as of 2012-2016 ACS module): New York-Newark-Jersey City, NY-NJ-PA (NY City); Los Angeles-Long Beach-Anaheim, CA (LA); Chicago-Naperville-Elgin, IL-IN-WI (Chicago); Dallas-Fort Worth-Arlington, TX (Dallas); Houston-The Woodlands-Sugar Land, TX (Houston).

inequality in poverty distribution across neighborhoods of a city. The latter component depends on the way poverty is unequally distributed across high poverty neighborhoods and on the way the poor population is split across clusters of high and low poverty. The approach builds on the idea that concentration of poverty across neighborhoods can produce welfare losses for those exposed to it. The concentrated poverty index, the official measure of urban poverty adopted by the Census Bureau to assess urban poverty, may fail to satisfy this basic requirement.

We use our urban poverty measure to highlight patterns, trends and components of urban poverty using census and ACS data for the largest 395 American MSAs over the last 35 years. While there is evidence that concentrated poverty has increased after the onset of the Great Recession, we find no systematic trends in the evolution of urban poverty. This apparent steadiness masks the implications of ongoing changes in the geography of poverty within MSAs, with poverty rising and falling heterogeneously across census tracts. The data we use do not allow to distinguish whether trends in urban poverty are driven by relocation of chronically poor individuals across census tracts, or rather by the fact that the likelihood of occurrence of poverty spells is unevenly distributed across census tracts, possibly affected by unobservable factors that are also relevant for the way rich and poor households sort in space. Distinguishing the two effects would require knowledge of individual-level incidence of poverty spells alongside residential decisions.

The urban poverty index we characterize focuses on the incidence and distributional inequality of poverty among neighborhoods where poverty is highly concentrated, as identified by an exogenously given urban poverty threshold (for instance above 20% or 40% of resident population). When the threshold is set to zero, urban poverty evaluations are based on all neighborhoods of the city. In this case, urban poverty evaluation depend only on the distribution of poverty across all neighborhoods of the city and the index converges to the Gini coefficient. Building on this result, we investigate a spatial decomposition of the implied urban poverty index, which is additive in the contribution to urban poverty of high-poverty clusters and the contribution of distant neighborhoods. This decomposition is relevant for analyzing the spatial dynamics of urban poverty, which may vary across similar cities on the basis of the quality of housing stock, the distribution of public goods

and the extent of affordability of neighborhoods. We find that in the largest MSAs, the non-neighborhood component of urban poverty is dominating. Trends are less clearcut for the rest of the MSAs.

The results presented in this paper can be extended in a variety of directions which are useful to study the incidence of covariates on urban poverty.

First, notice that the urban poverty index takes the poverty status identification as given, while it evaluates the distribution of poverty across neighborhoods. Different criteria can be used to identify the poor, for instance using different poverty thresholds. As a robustness check, we assess urban poverty looking at different populations with income below 75%, 100% (baseline) and 200% of the poverty line. Results reported in the appendix (Figure 8) show that urban poverty grows as the severity of poverty rises (panel a)), albeit uniformly across MSAs (panels b), c), d)). The ranking of MSAs based on different measures of poverty depth is robust (the indices obtained with different poverty groups display a rank correlation larger than 89% in 2014) and uncorrelated with concentrated poverty. While heterogeneity in distribution of poverty across neighborhoods is substantial, it is unlikely driven by the severity of poverty status.

Second, the urban poverty index obtained when setting the urban poverty threshold to zero, i.e. G , can be decomposed along the lines of the factor income decomposition by Shorrocks (1982) to analyze the contribution of various groups identified by varying degrees of poverty depth, as well as to analyze the contribution to urban poverty of different social groups in which the poor and non-poor populations can be further partitioned into, for instance along the lines of race, human capital or income. Furthermore, the urban poverty index G can be weakly (additively) decomposed as in Ebert (2010) into two components. A component captures the incidence of urban poverty among neighborhoods identified by some common trait, such as quality of housing stock, affordability, supply of local public goods. Another component measures instead the average contribution to urban poverty of differences between poverty incidence in any given neighborhood and poverty incidence in neighborhoods displaying different characteristics.

Finally, we acknowledge that poverty may be multidimensional, insofar individuals can be deprived in dimensions other than income and these dimensions are relevant to

explain the urban distribution (Decancq, Fleurbaey and Maniquet 2019). Borrowing on the relation with the Gini index, the urban poverty index can be generalized to the multidimensional setting following the approach in Koshevoy and Mosler (1997), based on a multivariate extension of the Gini index. As motivated in Andreoli and Zoli (2020), such approach consists in measuring the extent of dissimilarity between the distribution of the population across the city neighborhoods and the distributions of people that are poor in any given dimension, such as income (as we do here), housing or education. These extensions are left for future investigations.

Appendix

A Proofs

A.1 Proof of Theorem 1

We will prove the theorem making use of a sequence of lemmas that will highlight the role of the different axioms in the derivation of the final result.

Lemma 1 *Let $\mathcal{A} \in \Omega$, $\zeta \in [0, 1)$, and $z \geq 1$, $UP(\cdot; \zeta)$ satisfies AGG and INV-S if and only if there exist a continuous function $A : [0, 1]^2 \rightarrow \mathbb{R}_+$ and a function $h : [0, 1] \rightarrow \mathbb{R}$ continuous in $(0, 1)$ with $h(0) = 0$ such that:*

$$UP(\mathcal{A}; \zeta) := A\left(\frac{P}{N}, \frac{\bar{P}_z}{\bar{N}_z}\right) \sum_{i=1}^z \left(\frac{P_i}{N_i} - \zeta\right) \cdot \left[h\left(\frac{\bar{N}_i}{N}\right) - h\left(\frac{\bar{N}_{i-1}}{N}\right)\right] \quad (4)$$

with $\bar{N}_0 := 0$.

Proof. The proof combines the effect of AGG with INV-S by deriving a functional restriction on the class of weighting functions $w_i\left(\frac{N_1}{N}, \dots, \frac{N_i}{N}, \dots, \frac{N_n}{N}\right)$ that appear in the definition of AGG. We leave to the reader to verify that the index in (4) satisfies AGG and INV-S, here we focus on the proof of the (only if) part of the statement in the lemma.

First recall that, given AGG, we can write

$$UP(\mathcal{A}; \zeta) := A\left(\frac{P}{N}, \frac{\bar{P}_z}{\bar{N}_z}\right) \sum_{i=1}^z \frac{N_i}{N} \cdot \left(\frac{P_i}{N_i} - \zeta\right) \cdot w_i\left(\frac{N_1}{N}, \dots, \frac{N_i}{N}, \dots, \frac{N_n}{N}\right) \quad (5)$$

where $A : [0, 1]^2 \rightarrow \mathbb{R}_+$ and $w_i : \Delta_n \rightarrow \mathbb{R}$ satisfy the conditions specified in AGG.

Let $z \geq 1$, we apply INV-S. Note that because of the definition of INV-S, the scaling component $A\left(\frac{P}{N}, \frac{\bar{P}_z}{\bar{N}_z}\right)$ of (5) is not affected by splitting operations. Thus INV-S only affects the component $\sum_{i=1}^z \frac{N_i}{N} \cdot \left(\frac{P_i}{N_i} - \zeta\right) \cdot w_i\left(\frac{N_1}{N}, \dots, \frac{N_i}{N}, \dots, \frac{N_n}{N}\right)$.

We construct the proof in two steps. We first derive the restrictions on the function $w_1(\cdot)$ and then in a recursive manner we derive also the restrictions on all the other functions $w_i(\cdot)$ for $i = 2, 3, \dots, n$.

We first note that the function $w_i(\cdot)$ does not depend on ζ , and then we set ζ such that for a given $\mathcal{A} \in \Omega$ we have that $n = z$. Note that for any $\mathcal{A} \in \Omega$ there exist values of ζ such that $n = z$, for instance this is the case if we let $\zeta = 0$.

Step 1. Suppose that $n = z = 2$, and assume that $\frac{P_1}{N_1} > \zeta$, while $\frac{P_2}{N_2} = \zeta$. Apply repeatedly the splitting operations over the neighborhood indexed by $i = 2$. Because of the invariance requirement in INV-S and the specification in (5), if we denote by $\hat{n}_i := \frac{N_i}{N}$ we obtain that $\hat{n}_1\left(\frac{P_1}{N_1} - \zeta\right) \cdot w_1(\hat{n}_1, 1 - \hat{n}_1) = \hat{n}_1\left(\frac{P_1}{N_1} - \zeta\right) \cdot w_1(\hat{n}_1, \hat{n}_2, \hat{n}_3, \dots, \hat{n}_z)$ with $\hat{n}_2 + \hat{n}_3 + \dots + \hat{n}_z = 1 - \hat{n}_1$, this result holds for all $z = n \geq 2$. Recalling that $\frac{P_1}{N_1} - \zeta > 0$, we then obtain

$$\hat{n}_1 w_1(\hat{n}_1, 1 - \hat{n}_1) = \hat{n}_1 w_1(\hat{n}_1, \hat{n}_2, \hat{n}_3, \dots, \hat{n}_z)$$

with $\hat{n}_2 + \hat{n}_3 + \dots + \hat{n}_z = 1 - \hat{n}_1$, for all $z = n \geq 2$ and $\hat{n}_1 \in (0, 1)$. We can thus define the

function $h : [0, 1] \rightarrow \mathbb{R}$ such that $h(\hat{n}) := \hat{n}w_1(\hat{n}, 1 - \hat{n})$. It then follows that by definition

$$\hat{n}_1 w_1(\hat{n}_1, \hat{n}_2, \hat{n}_3, \dots, \hat{n}_z) = h(\hat{n}_1)$$

for all $z = n \geq 2$ and $\hat{n}_1 \in (0, 1)$. Given that by AGG $\hat{n}_1 w_1(\hat{n}_1, 1 - \hat{n}_1)$ is continuous for $\hat{n}_1 \in (0, 1)$ then $h(\cdot)$ is continuous on $(0, 1)$.

Step 2. Let $z = n = 1$, and assume to split into two neighborhoods the neighborhood 1 where $\frac{P_1}{N_1} - \zeta > 0$, then one obtains two neighborhoods of relative sizes \hat{n}_1 and $1 - \hat{n}_1$. INV-S then implies that $w_1(1) = \hat{n}_1 w_1(\hat{n}_1, 1 - \hat{n}_1) + (1 - \hat{n}_1) w_2(\hat{n}_1, 1 - \hat{n}_1)$. Let $h(1) := w_1(1)$, then one obtains for $z = n = 2$, $(1 - \hat{n}_1) w_2(\hat{n}_1, 1 - \hat{n}_1) = w_1(1) - \hat{n}_1 w_1(\hat{n}_1, 1 - \hat{n}_1)$, that is $(1 - \hat{n}_1) w_2(\hat{n}_1, 1 - \hat{n}_1) = h(1) - h(\hat{n}_1)$, in other words $(1 - \hat{n}_1) w_2(\hat{n}_1, 1 - \hat{n}_1) = h(\hat{n}_1 + (1 - \hat{n}_1)) - h(\hat{n}_1)$. This gives the definition of $w_2(\cdot)$ for $z = n = 2$.

The argument could be further generalized. Let $z = n = 2$, assume that $\frac{P_1}{N_1} > \zeta$, while $\frac{P_2}{N_2} = \zeta$. Then split neighborhood 1 of relative size \hat{n} into two neighborhoods of relative sizes respectively \hat{n}_1 and \hat{n}_2 such that $\hat{n}_1 + \hat{n}_2 = \hat{n}$, and, either leave neighborhood 2 unaffected, or split it into many others. According to INV-S it follows that $\hat{n} \left(\frac{P_1}{N_1} - \zeta \right) \cdot w_1(\hat{n}, 1 - \hat{n}) = \hat{n}_1 \left(\frac{P_1}{N_1} - \zeta \right) w_1(\hat{n}_1, \hat{n}_2, \hat{n}_3, \dots, \hat{n}_{z'}) + \hat{n}_2 \left(\frac{P_1}{N_1} - \zeta \right) w_2(\hat{n}_1, \hat{n}_2, \hat{n}_3, \dots, \hat{n}_{z'})$ where $z' = n \geq 3$ and $\hat{n}_3 + \dots + \hat{n}_{z'} = 1 - \hat{n} = 1 - \hat{n}_1 - \hat{n}_2$.

That is, $(\hat{n}_1 + \hat{n}_2) \cdot w_1(\hat{n}_1 + \hat{n}_2, 1 - \hat{n}_1 - \hat{n}_2) = \hat{n}_1 w_1(\hat{n}_1, \hat{n}_2, \hat{n}_3, \dots, \hat{n}_{z'}) + \hat{n}_2 w_2(\hat{n}_1, \hat{n}_2, \hat{n}_3, \dots, \hat{n}_{z'})$. Recalling that $\hat{n}_1 w_1(\hat{n}_1, \hat{n}_2, \hat{n}_3, \dots, \hat{n}_z) = h(\hat{n}_1)$ for all $z = n \geq 2$ and $\hat{n}_1 \in (0, 1)$, one obtains

$$\hat{n}_2 w_2(\hat{n}_1, \hat{n}_2, \hat{n}_3, \dots, \hat{n}_{z'}) = h(\hat{n}_1 + \hat{n}_2) - h(\hat{n}_1)$$

for all $z' = n \geq 2$ and $\hat{n}_1 + \hat{n}_2 \in (0, 1]$, $\hat{n}_1, \hat{n}_2 \in (0, 1)$.

By replicating the same logic and splitting into three neighborhoods the first one, then one can derive the definition of $w_3(\cdot)$ from

$$\hat{n}_3 w_3(\hat{n}_1, \hat{n}_2, \hat{n}_3, \dots, \hat{n}_{z'}) = h(\hat{n}_1 + \hat{n}_2 + \hat{n}_3) - h(\hat{n}_1 + \hat{n}_2)$$

for all $z' = n \geq 3$ and $\hat{n}_1 + \hat{n}_2 + \hat{n}_3 \in (0, 1]$, $\hat{n}_1, \hat{n}_2, \hat{n}_3 \in (0, 1)$.

We can then obtain in general that $\frac{N_i}{N_z} \cdot w_i \left(\frac{N_1}{N_z}, \dots, \frac{N_i}{N_z}, \dots, \frac{N_z}{N_z} \right) = h \left(\frac{\bar{N}_i}{N_z} \right) - h \left(\frac{\bar{N}_{i-1}}{N_z} \right)$ for $i = 1, 2, \dots, z$ and $z = n$ where $\frac{\bar{N}_0}{N_z} := 0$ and $h(0) := 0$. If $z = n$ then we have that $\bar{N}_z = N$, leading to

$$\frac{N_i}{N} \cdot w_i \left(\frac{N_1}{N}, \dots, \frac{N_i}{N}, \dots, \frac{N_n}{N} \right) = h \left(\frac{\bar{N}_i}{N} \right) - h \left(\frac{\bar{N}_{i-1}}{N} \right)$$

for $i = 1, 2, \dots, n$ where $\frac{\bar{N}_0}{N} := 0$ and $h(0) := 0$.

As pointed out the function $w_i(\cdot)$ does not depend on ζ , therefore even if it is derived under the assumption that ζ is such that $z = n$, the specification also hold for any $\zeta \in [0, 1)$, and therefore for any $z \leq n$, provided that $z \geq 1$ as required in the definition of AGG. ■

Lemma 2 *Let $\mathcal{A} \in \Omega$, $\zeta \in [0, 1)$, and $z \geq 1$, $UP(\cdot; \zeta)$ satisfies AGG, INV-S, INV-T if and only if there exist a continuous functions $A : [0, 1]^2 \rightarrow \mathbb{R}_+$ and $\beta_0, \gamma_0 \in \mathbb{R}$ such that:*

$$UP(\mathcal{A}; \zeta) := A \left(\frac{P}{N}, \frac{\bar{P}_z}{N_z} \right) \sum_{i=1}^z \frac{N_i}{N} \left(\frac{P_i}{N_i} - \zeta \right) \cdot \left[\beta_0 + \frac{\gamma_0}{2} \left(\frac{(N - \bar{N}_i) - \bar{N}_{i-1}}{N} \right) \right] \quad (6)$$

with $\bar{N}_0 := 0$.

Proof. We take the result from Lemma 1 and investigate the implications on the specification of $UP(\cdot; \zeta)$ generated by further imposing INV-T. We leave to the reader to check that the obtained specification of $UP(\cdot; \zeta)$ satisfies all axioms, here we focus on the "only if" part of the lemma.

For $z = 1$, INV-T does not hold, note that when $z = 1$ the specification of $UP(\cdot; \zeta)$ in the lemma is consistent with the one derived in Lemma 1 where $h(1) = \beta_0$ if $z = n = 1$. While the specification in the lemma for $h(\cdot)$ that is valid also when $z = 1 < n$, will be obtained in the next general part of the proof.

We set $z \geq 2$ and consider the transfers involved in the definition of INV-T. Note that with $z = 2$, the axiom is satisfied by construction given that it involves two transfers of population taking place in opposite directions and therefore their effects cancel out leading to the initial configuration \mathcal{A} .

Without loss of generality we assume that there are $z \geq 2$ neighborhoods with highly concentrated poverty with $\frac{P_i}{N_i} \geq \zeta$ and such that their population size is equal, that is $N_i = N_0$ for $i = 1, 2, \dots, z$. It follows that their relative population size within this set of neighborhoods is $\frac{N_i}{N_z} = \frac{1}{z}$, with $\frac{\bar{N}_i}{\bar{N}_z} = \frac{i}{z}$.

Moreover, we consider first the case where $\zeta \in [0, 1)$ is such that for a given $\mathcal{A} \in \Omega$ we have $z = n \geq 2$.

Consider the effect of the combined transfers of population in INV-T, and apply them to the specification derived in Lemma 1. Note that these transfers take place among neighborhoods in $\{1, 2, \dots, z\}$ and do not affect the components $A\left(\frac{P}{N}, \frac{\bar{P}_z}{\bar{N}_z}\right)$ and $\left[h\left(\frac{\bar{N}_i}{N}\right) - h\left(\frac{\bar{N}_{i-1}}{N}\right)\right]$ but only the distributions of $\left(\frac{P_i}{N_i} - \zeta\right)$. The application of the transfers in INV-T leads to the following condition

$$\begin{aligned} & \left[h\left(\frac{i}{z}\right) - h\left(\frac{i-1}{z}\right) \right] \varepsilon - \left[h\left(\frac{i+1}{z}\right) - h\left(\frac{i}{z}\right) \right] \varepsilon \\ & - \left[h\left(\frac{j}{z}\right) - h\left(\frac{j-1}{z}\right) \right] \varepsilon + \left[h\left(\frac{j+1}{z}\right) - h\left(\frac{j}{z}\right) \right] \varepsilon = 0 \end{aligned} \quad (7)$$

for $\varepsilon > 0$, satisfying the conditions specified in INV-T for all $i, j \in \{1, 2, \dots, z-1\}$ with $z = n \geq 2$.

It then follows that $\left[h\left(\frac{j+1}{z}\right) - h\left(\frac{j}{z}\right) \right] - \left[h\left(\frac{j}{z}\right) - h\left(\frac{j-1}{z}\right) \right] = \left[h\left(\frac{i+1}{z}\right) - h\left(\frac{i}{z}\right) \right] - \left[h\left(\frac{i}{z}\right) - h\left(\frac{i-1}{z}\right) \right]$ for all $i, j \in \{1, 2, \dots, z-1\}$, and for all $z = n \geq 2$. Thus, we have that $\left[h\left(\frac{i+1}{z}\right) - h\left(\frac{i}{z}\right) \right] - \left[h\left(\frac{i}{z}\right) - h\left(\frac{i-1}{z}\right) \right]$ does not depend on i but eventually only on $1/z$. In general, there exists a function $g(1/z)$ such that

$$\left[h\left(\frac{i+1}{z}\right) - h\left(\frac{i}{z}\right) \right] - \left[h\left(\frac{i}{z}\right) - h\left(\frac{i-1}{z}\right) \right] = g\left(\frac{1}{z}\right)$$

for all $i \in \{1, 2, \dots, z-1\}$, all $z = n \geq 2$.

To simplify the exposition, denote $\frac{1}{z} = \sigma$ and let $f(i) := h(i\sigma)$ for a fixed σ . We then have

$$[f(i+1) - f(i)] - [f(i) - f(i-1)] = g(\sigma)$$

for all $i \in \{1, 2, \dots, z-1\}$, all $z = n \geq 2$.

Let $d(i) := f(i) - f(i-1)$ for $i \in \{1, 2, \dots, z-1\}$ with by construction $f(0) = h(0) = 0$. Thus, $d(1) + d(2) + \dots + d(i) = f(i) - f(0) = f(i)$. We then have

$$d(i+1) - d(i) = g(\sigma)$$

for $i \in \{1, 2, \dots, z-1\}$, that leads to $d(j) = d(1) + (j-1)g(\sigma)$ for $i \in \{1, 2, \dots, z\}$ and all $z = n \geq 2$.

Thus, $f(i) = \sum_{j=1}^i d(j) = \sum_{j=1}^i d(1) + (j-1)g(\sigma)$, with $f(1) = d(1)$. It follows that

$$f(i) = if(1) + \frac{i(i-1)}{2}g(\sigma)$$

for $i \in \{1, 2, \dots, z\}$, $z = n \geq 2$. Recalling that $f(i) := h(i/z)$ we have

$$h\left(\frac{i}{z}\right) = ih\left(\frac{1}{z}\right) + \frac{i(i-1)}{2}g(1/z)$$

for $i \in \{1, 2, \dots, z\}$, $z = n \geq 2$. Given that i and z can take any pair of natural number values such that $i \leq z = n$, and that $h(0) = 0$ by construction, then the above formula is a functional equation that allow to specify the value of the function $h(\cdot)$ for all rational numbers in $[0, 1]$.

Let $i = z$, we then obtain

$$h(1) = zh\left(\frac{1}{z}\right) + \frac{z(z-1)}{2}g\left(\frac{1}{z}\right),$$

note that the value of $h(1)$ is constant and therefore independent from z , rearranging it then follows that for any $z = n \geq 2$ it holds that

$$h\left(\frac{1}{z}\right) = \frac{h(1)}{z} - \frac{(z-1)}{2}g\left(\frac{1}{z}\right).$$

Recalling the derivation of $h\left(\frac{i}{z}\right)$ and inserting $h\left(\frac{1}{z}\right)$ one obtains

$$h\left(\frac{i}{z}\right) = h(1)\frac{i}{z} - \frac{i(z-i)}{2}g\left(\frac{1}{z}\right)$$

for $i \in \{1, 2, \dots, z\}$, $z = n \geq 2$.

Note that $\frac{i}{z}$ are unaffected if both i and $z = n$ are replicated r times for $r \in \{1, 2, 3, 4, \dots\}$, we then obtain $h\left(\frac{i}{z}\right) = h\left(\frac{ri}{rz}\right) = h(1)\frac{ri}{rz} - \frac{ri(rz-ri)}{2}g\left(\frac{1}{rz}\right) = h(1)\frac{i}{z} - r^2\frac{i(z-i)}{2}g\left(\frac{1}{rz}\right)$ for all r . As a result it should hold that

$$r^2g\left(\frac{1}{rz}\right) = g\left(\frac{1}{z}\right)$$

for all $r, z \in \{2, 3, 4, \dots\}$ with $z = n$, thus, $g\left(\frac{1}{rz}\right) = \frac{1}{r^2}g\left(\frac{1}{z}\right)$. By switching z with r one obtains $g\left(\frac{1}{rz}\right) = \frac{1}{z^2}g\left(\frac{1}{r}\right)$. As a result $\frac{1}{r^2}g\left(\frac{1}{z}\right) = \frac{1}{z^2}g\left(\frac{1}{r}\right)$ for all $r, z \in \{2, 3, 4, \dots\}$, that is, $z^2g\left(\frac{1}{z}\right) = r^2g\left(\frac{1}{r}\right)$ for all r, z . Thus, there exists a constant $\gamma_0 \in \mathbb{R}$ such that $z^2g\left(\frac{1}{z}\right) = -\gamma_0$, leading to $g\left(\frac{1}{z}\right) = -\frac{\gamma_0}{z^2}$ for all $z \in \{2, 3, 4, \dots\}$. By substituting into the definition of $h\left(\frac{i}{z}\right)$ and letting $h(1) = \beta_0 \in \mathbb{R}$ it follows that $h\left(\frac{i}{z}\right) = \beta_0\frac{i}{z} + \frac{\gamma_0}{2}\frac{i(z-i)}{z}$. Note that we have derived the specification of $h(\cdot)$ under the assumption that $z = n \geq 2$, then replacing z with n we obtain

$$h\left(\frac{i}{n}\right) = \beta\frac{i}{n} + \frac{\gamma_0}{2}\frac{i(n-i)}{n}.$$

Recall that $\frac{i}{n}$ by construction could be any rational number in $(0, 1]$, with $h(0) = 0$ already

set in Lemma 1. Given that the set of rational numbers is dense in $(0, 1]$ and that $h(\cdot)$ is continuous in that interval the result could be extended to all real numbers in $[0, 1]$, with $h(0) = 0$. Recalling that $\frac{\bar{N}_i}{N} = \frac{i}{n}$ we can then write more generally

$$h\left(\frac{\bar{N}_i}{N}\right) = \beta_0 \frac{\bar{N}_i}{N} + \frac{\gamma_0}{2} \frac{\bar{N}_i}{N} \frac{(N - \bar{N}_i)}{N}.$$

Consider the weighting function $\left[h\left(\frac{\bar{N}_i}{N}\right) - h\left(\frac{\bar{N}_{i-1}}{N}\right)\right]$ from Lemma 1, it can then be specified as:

$$\begin{aligned} h\left(\frac{\bar{N}_i}{N}\right) - h\left(\frac{\bar{N}_{i-1}}{N}\right) &= \left(\beta_0 + \frac{\gamma_0}{2}\right) \frac{\bar{N}_i}{N} - \frac{\gamma_0}{2} \left(\frac{\bar{N}_i}{N}\right)^2 - \left(\beta_0 + \frac{\gamma_0}{2}\right) \frac{\bar{N}_{i-1}}{N} + \frac{\gamma_0}{2} \left(\frac{\bar{N}_{i-1}}{N}\right)^2 \\ &= \left(\beta_0 + \frac{\gamma_0}{2}\right) \frac{N_i}{N} - \frac{\gamma_0}{2} \left(\frac{\bar{N}_i - \bar{N}_{i-1}}{N}\right) \left(\frac{\bar{N}_i + \bar{N}_{i-1}}{N}\right) \\ &= \left(\beta_0 + \frac{\gamma_0}{2}\right) \frac{N_i}{N} - \frac{\gamma_0}{2} \frac{N_i}{N} \left(\frac{\bar{N}_i + \bar{N}_{i-1}}{N}\right) \\ &= \beta_0 \frac{N_i}{N} - \frac{\gamma_0}{2} \frac{N_i}{N} \left(\frac{\bar{N}_i + \bar{N}_{i-1}}{N} - 1\right) \\ &= \beta_0 \frac{N_i}{N} + \frac{\gamma_0}{2} \frac{N_i}{N} \left(\frac{(N - \bar{N}_i) - \bar{N}_{i-1}}{N}\right). \end{aligned}$$

By substituting into the specification of $UP(\mathcal{A}; \zeta)$ in Lemma 1, one obtains the results presented in this lemma.

Recall that we have derived the result under the assumption that for $\zeta \in [0, 1]$ we have $z = n \geq 2$ and note that the function $h(\cdot)$ does not depend on ζ . In order to extend the result to all cases where $n \geq z \geq 2$ it needs to be checked that the obtained functional form for $h(\cdot)$ allows to satisfy INV-T also when $z < n$. Note that, as in (7), the application of the transfers in INV-T, when $n > z \geq 2$ requires that the following condition

$$\begin{aligned} &\left[h\left(\frac{i}{n}\right) - h\left(\frac{i-1}{n}\right)\right] \varepsilon - \left[h\left(\frac{i+1}{n}\right) - h\left(\frac{i}{n}\right)\right] \varepsilon \\ &- \left[h\left(\frac{j}{n}\right) - h\left(\frac{j-1}{n}\right)\right] \varepsilon + \left[h\left(\frac{j+1}{n}\right) - h\left(\frac{j}{n}\right)\right] \varepsilon = 0 \end{aligned}$$

has to be satisfied for all $i, j \in \{1, 2, \dots, z-1\}$, for $n > z \geq 2$, for $\varepsilon > 0$. That is, after substituting for the derived specification of $h\left(\frac{\bar{N}_i}{N}\right) - h\left(\frac{\bar{N}_{i-1}}{N}\right)$, it should be verified that

$$\begin{aligned} &\left[h\left(\frac{\bar{N}_i}{N}\right) - h\left(\frac{\bar{N}_{i-1}}{N}\right)\right] - \left[h\left(\frac{\bar{N}_{i+1}}{N}\right) - h\left(\frac{\bar{N}_i}{N}\right)\right] = \\ &\left[h\left(\frac{\bar{N}_j}{N}\right) - h\left(\frac{\bar{N}_{j-1}}{N}\right)\right] - \left[h\left(\frac{\bar{N}_{j+1}}{N}\right) - h\left(\frac{\bar{N}_j}{N}\right)\right]. \end{aligned}$$

Note that $\left[h\left(\frac{\bar{N}_i}{N}\right) - h\left(\frac{\bar{N}_{i-1}}{N}\right)\right] - \left[h\left(\frac{\bar{N}_{i+1}}{N}\right) - h\left(\frac{\bar{N}_i}{N}\right)\right]$ equals $\beta_0 \frac{N_i}{N} + \frac{\gamma_0}{2} \frac{N_i}{N} \left(\frac{(N - \bar{N}_i) - \bar{N}_{i-1}}{N}\right) - \beta_0 \frac{N_{i+1}}{N} - \frac{\gamma_0}{2} \frac{N_{i+1}}{N} \left(\frac{(N - \bar{N}_{i+1}) - \bar{N}_i}{N}\right)$, and recall that according to INV-T $N_i = N_j = N_{i+1} =$

N_{j+1} it then follows that

$$\left[h\left(\frac{\bar{N}_i}{N}\right) - h\left(\frac{\bar{N}_{i-1}}{N}\right) \right] - \left[h\left(\frac{\bar{N}_{i+1}}{N}\right) - h\left(\frac{\bar{N}_i}{N}\right) \right] = 0.$$

This is similarly the case if we consider the neighborhood with index j . As a result INV-T holds also for $n > z \geq 2$.

To complete the exposition we consider the case where $n > z = 1$. In this case INV-T cannot be applied, however we have already derived the required specifications for function $h(\cdot)$ from the previous steps of the proof. ■

Lemma 3 *Let $\mathcal{A} \in \Omega$, $\zeta \in [0, 1)$, $UP(\cdot; \zeta)$ satisfies AGG, INV-S, INV-T, INV-PL, MON, TRAN and NOR if and only if there exist $\beta, \gamma \geq 0$ such that:*

$$UP(\mathcal{A}; \zeta) := \beta \cdot \frac{\bar{P}_z - \zeta \bar{N}_z}{P} + \gamma \cdot \frac{\bar{N}_z}{N} \cdot \sum_{i=1}^z \frac{P_i}{P} \cdot \left(\frac{(\bar{N}_z - \bar{N}_i) - \bar{N}_{i-1}}{\bar{N}_z} \right) + \gamma \cdot \left(\frac{N - \bar{N}_z}{N} \right) \cdot \frac{\bar{P}_z - \zeta \bar{N}_z}{P} \quad (8)$$

with $\bar{N}_0 := 0$, if $z \geq 1$, otherwise $UP(\mathcal{A}; \zeta) = 0$.

Proof. We consider the result from Lemma 2 and investigate the implications on the specification of $UP(\cdot; \zeta)$ generated by further imposing INV-PL, MON, TRAN and NOR. We leave to the reader to check that the obtained specification of $UP(\cdot; \zeta)$ satisfies all axioms, here we focus on the "only if" part of the lemma. Recall that, if $z \geq 1$, then according to Lemma 2 it is possible to write

$$UP(\mathcal{A}; \zeta) := A\left(\frac{P}{N}, \frac{\bar{P}_z}{\bar{N}_z}\right) \sum_{i=1}^z \frac{N_i}{N} \left(\frac{P_i}{N_i} - \zeta\right) \cdot \left[\beta_0 + \frac{\gamma_0}{2} \left(\frac{(N - \bar{N}_i) - \bar{N}_{i-1}}{N}\right) \right].$$

We first consider INV-PL(i). By applying the scale component $\lambda > 0$ one obtains that from \mathcal{A} to \mathcal{A}' the values $\left(\frac{P}{N}, \frac{\bar{P}_z}{\bar{N}_z}, \frac{P_i}{N_i}, \zeta\right)$ are scaled to $\left(\lambda \frac{P}{N}, \lambda \frac{\bar{P}_z}{\bar{N}_z}, \lambda \frac{P_i}{N_i}, \lambda \zeta\right)$, it then follows that $UP(\mathcal{A}; \zeta) = UP(\mathcal{A}'; \lambda \zeta)$ if and only if $A\left(\frac{P}{N}, \frac{\bar{P}_z}{\bar{N}_z}\right) = \lambda A\left(\lambda \frac{P}{N}, \lambda \frac{\bar{P}_z}{\bar{N}_z}\right) \geq 0$.

We take into account two cases, first when $\zeta = 0$ and then when $\zeta \in (0, 1)$.

Case 1: $\zeta = 0$. In this case holds only INV-PL(i). If $\zeta = 0$ then $\frac{P}{N} = \frac{\bar{P}_z}{\bar{N}_z}$. By applying INV-PL(i) it then follows that $A\left(\frac{P}{N}, \frac{P}{N}\right) = \lambda A\left(\lambda \frac{P}{N}, \lambda \frac{P}{N}\right)$ for $\lambda > 0$. Let $\lambda \frac{P}{N} = 1$, we obtain $A\left(\frac{P}{N}, \frac{P}{N}\right) = \frac{A(1,1)}{\frac{P}{N}}$, letting $A(1,1) := K \geq 0$, then $A\left(\frac{P}{N}, \frac{P}{N}\right) = \frac{K}{\frac{P}{N}}$.

Case 2: $\zeta \in (0, 1]$. In this case $\frac{P}{N} \leq \zeta \leq \frac{\bar{P}_z}{\bar{N}_z} \leq 1$. INV-PL(i) holds if and only if $A\left(\frac{P}{N}, \frac{\bar{P}_z}{\bar{N}_z}\right) = \lambda A\left(\lambda \frac{P}{N}, \lambda \frac{\bar{P}_z}{\bar{N}_z}\right) \geq 0$ for $\lambda > 0$. Moreover, according to INV-PL(ii) one obtains $A\left(\frac{P}{N}, \frac{\bar{P}_z}{\bar{N}_z}\right) = A\left(\frac{P}{N}, \frac{\bar{P}_z}{\bar{N}_z} + \theta\right)$ where $\frac{P}{N} < \zeta \leq \frac{\bar{P}_z}{\bar{N}_z}$. By INV-PL(ii) it follows that there exists a continuous function $H(\cdot)$ such that $A\left(\frac{P}{N}, \frac{\bar{P}_z}{\bar{N}_z}\right) := H\left(\frac{P}{N}\right)$ whenever $\frac{P}{N} < \zeta \leq \frac{\bar{P}_z}{\bar{N}_z}$. Note that $\frac{P}{N} = \frac{\bar{P}_z}{\bar{N}_z} \frac{\bar{N}_z}{N} + \left(1 - \frac{\bar{N}_z}{N}\right) \zeta'$ for $\zeta' < \zeta$, where ζ' denotes the average poverty incidence of the neighborhood with incidence below ζ . By letting $\bar{N}_z \rightarrow 0$ and $\zeta' \rightarrow \zeta$ one obtains the case where $\frac{P}{N} \rightarrow \zeta$. We then have that $A\left(\frac{P}{N}, \frac{\bar{P}_z}{\bar{N}_z}\right) = H\left(\frac{P}{N}\right)$ for $\frac{P}{N} < \frac{\bar{P}_z}{\bar{N}_z}$. Thus, by INV-PL(i) we have $\lambda A\left(\lambda \frac{P}{N}, \lambda \frac{\bar{P}_z}{\bar{N}_z}\right) = \lambda H\left(\lambda \frac{P}{N}\right) = H\left(\frac{P}{N}\right) = A\left(\frac{P}{N}, \frac{\bar{P}_z}{\bar{N}_z}\right)$ for $\frac{P}{N} < \frac{\bar{P}_z}{\bar{N}_z}$ and for $\lambda > 0$ such that $\lambda \frac{\bar{P}_z}{\bar{N}_z} \leq 1$. Suppose that there exists \bar{c} such that $\lambda \frac{P}{N} = \bar{c}$, then

$\lambda H\left(\lambda \frac{P}{N}\right) = H\left(\frac{P}{N}\right)$ implies that $\frac{\bar{c}H(\bar{c})}{\frac{P}{N}} = H\left(\frac{P}{N}\right)$. By letting $\bar{c}H(\bar{c}) := K \geq 0$ one obtains that $A\left(\frac{P}{N}, \frac{\bar{P}_z}{N_z}\right) = H\left(\frac{P}{N}\right) = \frac{K}{\frac{P}{N}}$ for $\frac{P}{N} < \frac{\bar{P}_z}{N_z}$.

Thus, we have that both

$$A\left(\frac{P}{N}, \frac{\bar{P}_z}{N_z}\right) = \frac{K}{\frac{P}{N}}$$

for $\frac{P}{N} < \frac{\bar{P}_z}{N_z} \leq 1$ and for $\frac{P}{N} = \frac{\bar{P}_z}{N_z} \leq 1$ that identify all possible ranges of values of the arguments of $A(\cdot)$.

By applying the result to the specification in Lemma 2, and letting $\beta := \beta_0 K$ and $\gamma := \gamma_0 K/2$ one obtains

$$\begin{aligned} UP(\mathcal{A}; \zeta) &:= \frac{1}{\frac{P}{N}} \cdot \sum_{i=1}^z \frac{N_i}{N} \left(\frac{P_i}{N_i} - \zeta \right) \cdot \left[\beta + \gamma \left(\frac{(N - \bar{N}_i) - \bar{N}_{i-1}}{N} \right) \right] \\ &= \sum_{i=1}^z \left(\frac{P_i}{P} - \zeta \frac{N_i}{P} \right) \cdot \left[\beta + \gamma \left(\frac{(N - \bar{N}_i) - \bar{N}_{i-1}}{N} \right) \right]. \end{aligned} \quad (9)$$

We now investigate the effects of MON and TRAN. According to MON considering that $UP(\mathcal{A}'; \zeta) = \sum_{i=1}^z \left(\frac{\nu P_i}{\nu P} - \zeta \frac{N_i}{\nu P} \right) \cdot \left[\beta + \gamma \left(\frac{(N - \bar{N}_i) - \bar{N}_{i-1}}{N} \right) \right]$ for $\nu > 1$, it should hold $UP(\mathcal{A}'; \zeta) - UP(\mathcal{A}; \zeta) \geq 0$, this requires that

$$\sum_{i=1}^z \left[\left(-\zeta \frac{N_i}{\nu P} \right) - \left(-\zeta \frac{N_i}{P} \right) \right] \cdot \left[\beta + \gamma \left(\frac{(N - \bar{N}_i) - \bar{N}_{i-1}}{N} \right) \right] \geq 0$$

for all $\nu > 1$, all $\zeta \in [0, 1)$, all $\mathcal{A} \in \Omega$. Rearranging the condition, it implies that

$$\frac{\zeta}{P/N} \left(-\frac{1}{\nu} + 1 \right) \sum_{i=1}^z \frac{N_i}{N} \cdot \left[\beta + \gamma \left(\frac{(N - \bar{N}_i) - \bar{N}_{i-1}}{N} \right) \right] \geq 0$$

for all $\nu > 1$, all $\zeta \in [0, 1)$, all $\mathcal{A} \in \Omega$.

This is the case if and only if $\sum_{i=1}^z \frac{N_i}{N} \cdot \left[\beta + \gamma \left(\frac{(N - \bar{N}_i) - \bar{N}_{i-1}}{N} \right) \right] \geq 0$ for all $\mathcal{A} \in \Omega$. This condition depends on the value of $z \geq 1$, and in particular, because of the construction of the weighting function $w_i(\cdot)$ that satisfies INV-S, the condition depends only on $\frac{\bar{N}_z}{N}$. In fact, in this case, because of INV-S, without loss of generality, one can consider distributions with two neighborhoods and $z = 1$. In this case $N_1 = \bar{N}_1 = \bar{N}_z$ and recall that $\bar{N}_0 = 0$. After substituting, one obtains the condition

$$\frac{\bar{N}_z}{N} \cdot \left[\beta + \gamma \left(\frac{N - \bar{N}_z}{N} \right) \right] \geq 0 \quad (10)$$

for all $\frac{\bar{N}_z}{N} \in (0, 1]$. Letting $\frac{\bar{N}_z}{N} = 1$, that is if $z = n$, it follows that a necessary condition for MON to hold is $\beta \geq 0$. Moreover, letting $\frac{\bar{N}_z}{N} \rightarrow 0$, the additional derived necessary condition is $\beta + \gamma \geq 0$, because otherwise, if $\gamma < -\beta$ for sufficiently small values of $\frac{\bar{N}_z}{N}$ is possible to violate the condition in (10). Both necessary conditions $\beta \geq 0$ and $\beta + \gamma \geq 0$ turn out to be sufficient for (10) to hold for all $\frac{\bar{N}_z}{N} \in (0, 1]$.

We consider now the restrictions required by axiom TRAN. First we consider the case where because of the transfer the poverty incidence in neighborhood j does not fall below ζ , that is $P_j^{\mathcal{A}'} / N_j^{\mathcal{A}'} \geq \zeta$.

Recall moreover, that according to TRAN the considered transfer does not affect the ranking of the neighborhoods. Consider (9) and note that according to TRAN, only P_i and P_j are modified by the transfer, it should then be verified that

$$\begin{aligned} & UP(\mathcal{A}'; \zeta) - UP(\mathcal{A}; \zeta) = \\ &= \frac{\varepsilon}{P} \left[\beta + \gamma \left(\frac{(N - \bar{N}_i) - \bar{N}_{i-1}}{N} \right) \right] - \frac{\varepsilon}{P} \left[\beta + \gamma \left(\frac{(N - \bar{N}_j) - \bar{N}_{j-1}}{N} \right) \right] \geq 0 \end{aligned}$$

for $j > i$, with $j \leq z$, and $\varepsilon > 0$. Thus

$$\gamma \frac{\varepsilon}{P} \left[\frac{\bar{N}_j - \bar{N}_i + \bar{N}_{j-1} - \bar{N}_{i-1}}{N} \right] \geq 0$$

with $\bar{N}_j - \bar{N}_i > 0$ and $\bar{N}_{j-1} - \bar{N}_{i-1} > 0$, which implies

$$\left(\frac{(N - \bar{N}_i) - \bar{N}_{i-1}}{N} \right) - \left(\frac{(N - \bar{N}_j) - \bar{N}_{j-1}}{N} \right) = \frac{\bar{N}_j - \bar{N}_i + \bar{N}_{j-1} - \bar{N}_{i-1}}{N} > 0,$$

as a result should hold $\gamma \geq 0$.

Note that the condition should hold for all $i < j \leq z$ with $z \leq 2$ and therefore letting $z = n$, should hold for all $i < j \leq n$.

We consider now the case where $P_j^{\mathcal{A}'} / N_j^{\mathcal{A}'} < \zeta$, with $j = z$, by applying TRAN, follows that

$$\begin{aligned} & UP(\mathcal{A}'; \zeta) - UP(\mathcal{A}; \zeta) = \\ &= \frac{\varepsilon}{P} \left[\beta + \gamma \left(\frac{(N - \bar{N}_i) - \bar{N}_{i-1}}{N} \right) \right] - \frac{\varepsilon'}{P} \left[\beta + \gamma \left(\frac{(N - \bar{N}_j) - \bar{N}_{j-1}}{N} \right) \right] \geq 0 \end{aligned}$$

where $0 \leq \varepsilon' \leq \varepsilon$ with $\varepsilon' = \min\{\varepsilon, P_j - \zeta N_j\}$. The condition can then be simplified as

$$(\varepsilon - \varepsilon') \beta + \gamma \left[\varepsilon \left(\frac{(N - \bar{N}_i) - \bar{N}_{i-1}}{N} \right) - \varepsilon' \left(\frac{(N - \bar{N}_j) - \bar{N}_{j-1}}{N} \right) \right] \geq 0.$$

Recalling that $\frac{(N - \bar{N}_i) - \bar{N}_{i-1}}{N} > \frac{(N - \bar{N}_j) - \bar{N}_{j-1}}{N}$ for $j > i$, that $\varepsilon' \leq \varepsilon$, and that $\beta \geq 0$, then $\gamma \geq 0$ is sufficient to verify that $UP(\mathcal{A}'; \zeta) \geq UP(\mathcal{A}; \zeta)$.

Thus, $\gamma \geq 0$ is necessary and sufficient for TRAN to hold. By combining with the parametric restrictions derived by applying MON one obtains $\beta \geq 0$, and $\gamma \geq 0$.

All derivations illustrated so far consider the case where $z \geq 1$. Note that for $\mathcal{A} \in \Omega$ and given $\zeta \in (0, 1)$ it is possible to take into account also configurations where $\frac{P_i}{N_i} < \zeta$ for all neighborhoods i . In this case the value of the index is derived by considering axiom NOR. For all these configurations the value of the index coincides with the infimum of the index taken over all the other possible configurations in \mathcal{A} where $z \geq 1$. Consider the obtained derivation of UP for $z \geq 1$, where

$$UP(\mathcal{A}; \zeta) = \sum_{i=1}^z \left(\frac{P_i - \zeta N_i}{P} \right) \cdot \left[\beta + \gamma \left(\frac{(N - \bar{N}_i) - \bar{N}_{i-1}}{N} \right) \right] \quad (11)$$

with $\beta, \gamma \geq 0$, note that the first term in the summation $\frac{P_i - \zeta N_i}{P} \geq 0$ is non-increasing in i , and that the term $\frac{(N - \bar{N}_i) - \bar{N}_{i-1}}{N}$ is also non-increasing in i . It follows that, given that

$\beta, \gamma \geq 0$ also $\beta + \gamma \left(\frac{(N - \bar{N}_i) - \bar{N}_{i-1}}{N} \right)$ is non-increasing in i . The summation in (11) is then minimized, for each $z \geq 1$ if the terms $\frac{P_i - \zeta N_i}{P}$ are equalized. Given that $\frac{P_i - \zeta N_i}{P} \geq 0$ then the minimum for each $z \geq 1$ is obtained for $P_i - \zeta N_i = 0$ for all $i \leq z$. It follows that in this case $UP = 0$.

Thus, by NOR the value of the index is 0 when $\frac{P_i}{N_i} < \zeta$ for all i .

To complete the proof we rearrange the specification of $UP(\cdot; \zeta)$ in (11). We can rewrite:

$$\begin{aligned}
UP(\mathcal{A}; \zeta) & : = \sum_{i=1}^z \left(\frac{P_i - \zeta N_i}{P} \right) \cdot \left[\beta + \gamma \left(\frac{(N - \bar{N}_i) - \bar{N}_{i-1}}{N} \right) \right] \\
& = \beta \sum_{i=1}^z \left(\frac{P_i - \zeta N_i}{P} \right) + \gamma \sum_{i=1}^z \left(\frac{P_i - \zeta N_i}{P} \right) \cdot \left(\frac{(N - \bar{N}_i) - \bar{N}_{i-1} + \bar{N}_z - \bar{N}_z}{N} \right) \\
& = \beta \cdot \frac{\bar{P}_z - \zeta \bar{N}_z}{P} + \gamma \frac{\bar{N}_z}{N} \sum_{i=1}^z \left(\frac{P_i - \zeta N_i}{P} \right) \cdot \left(\frac{(\bar{N}_z - \bar{N}_i)}{\bar{N}_z} - \frac{\bar{N}_{i-1}}{\bar{N}_z} + \frac{N - \bar{N}_z}{\bar{N}_z} \right) \\
& = \beta \cdot \frac{\bar{P}_z - \zeta \bar{N}_z}{P} + \gamma \frac{\bar{N}_z}{N} \sum_{i=1}^z \left(\frac{P_i - \zeta N_i}{P} \right) \cdot \left(\frac{(\bar{N}_z - \bar{N}_i) - \bar{N}_{i-1}}{\bar{N}_z} \right) + \\
& \quad + \gamma \left(\frac{N - \bar{N}_z}{N} \right) \frac{\bar{P}_z - \zeta \bar{N}_z}{P}.
\end{aligned}$$

Note that

$$\sum_{i=1}^z N_i (\bar{N}_z - \bar{N}_i - \bar{N}_{i-1}) = 0,$$

in fact $\sum_{i=1}^z N_i (\bar{N}_z - \bar{N}_i - \bar{N}_{i-1}) = \sum_{i=1}^z N_i \bar{N}_z - \sum_{i=1}^z N_i (\bar{N}_i + \bar{N}_{i-1}) = (\bar{N}_z)^2 - \sum_{i=1}^z (\bar{N}_i - \bar{N}_{i-1})(\bar{N}_i + \bar{N}_{i-1})$ where $\sum_{i=1}^z (\bar{N}_i - \bar{N}_{i-1})(\bar{N}_i + \bar{N}_{i-1}) = \sum_{i=1}^z (\bar{N}_i)^2 - (\bar{N}_{i-1})^2 = (\bar{N}_z)^2$. It then follows that the term $\sum_{i=1}^z \left(\frac{P_i - \zeta N_i}{P} \right) \cdot \left(\frac{(\bar{N}_z - \bar{N}_i) - \bar{N}_{i-1}}{\bar{N}_z} \right)$ simplifies to $\sum_{i=1}^z \frac{P_i}{P} \cdot \left(\frac{(\bar{N}_z - \bar{N}_i) - \bar{N}_{i-1}}{\bar{N}_z} \right)$. Thus, we obtain:

$$UP(\mathcal{A}; \zeta) = \beta \cdot \frac{\bar{P}_z - \zeta \bar{N}_z}{P} + \gamma \cdot \frac{\bar{N}_z}{N} \cdot \sum_{i=1}^z \left(\frac{P_i}{P} \right) \cdot \left(\frac{(\bar{N}_z - \bar{N}_i) - \bar{N}_{i-1}}{\bar{N}_z} \right) + \gamma \cdot \left(\frac{N - \bar{N}_z}{N} \right) \cdot \frac{\bar{P}_z - \zeta \bar{N}_z}{P}.$$

■

In order to complete the proof of the Theorem one has to link the result in Lemma 3 with the Gini index formula $G(\mathcal{A}; \zeta)$. Next lemma provides this link.

Lemma 4 *Let $\mathcal{A} \in \Omega$, $\zeta \in [0, 1)$, and $z \geq 1$, then*

$$\frac{\bar{P}_z}{P} G(\mathcal{A}; \zeta) = \sum_{i=1}^z \frac{P_i}{P} \left(\frac{\bar{N}_z - \bar{N}_i}{\bar{N}_z} - \frac{\bar{N}_{i-1}}{\bar{N}_z} \right).$$

Proof. The Gini index $G(\cdot; \zeta)$ can be written as follows:

$$\begin{aligned}
G(\mathcal{A}; \zeta) &= \frac{1}{2\frac{\bar{P}_z}{\bar{N}_z}} \sum_{i=1}^z \sum_{j=1}^z \frac{N_i N_j}{\bar{N}_z \bar{N}_z} \left| \frac{P_i}{N_i} - \frac{P_j}{N_j} \right| \\
&= \frac{\bar{N}_z}{2\bar{P}_z} \sum_{i=1}^z \sum_{j=1}^z \frac{N_i N_j}{\bar{N}_z \bar{N}_z} \left[2 \max \left\{ \frac{P_i}{N_i}, \frac{P_j}{N_j} \right\} - \frac{P_i}{N_i} - \frac{P_j}{N_j} \right] \\
&= \frac{\bar{N}_z}{2\bar{P}_z} \left[\sum_{i=1}^z \sum_{j=1}^z \frac{N_i N_j}{\bar{N}_z \bar{N}_z} 2 \max \left\{ \frac{P_i}{N_i}, \frac{P_j}{N_j} \right\} - 2 \sum_{i=1}^z \frac{N_i}{\bar{N}_z} \sum_{i=1}^z \frac{N_i P_i}{\bar{N}_z N_i} \right]. \quad (12)
\end{aligned}$$

Thus

$$\frac{\bar{P}_z}{\bar{N}_z} G(\mathcal{A}; \zeta) = \sum_{i=1}^z \sum_{j=1}^z \frac{N_i N_j}{\bar{N}_z \bar{N}_z} \max \left\{ \frac{P_i}{N_i}, \frac{P_j}{N_j} \right\} - \sum_{i=1}^z \frac{N_i P_i}{\bar{N}_z N_i}. \quad (13)$$

We now develop the first term appearing in squared brackets in (12), denoted *max* in short-hand notation, to show that it can be written as a function of the rank weights. First, let develop the double summations term as follows:

$$\begin{aligned}
max &= \sum_{i=1}^z \sum_{j=1}^z \frac{N_i N_j}{\bar{N}_z \bar{N}_z} \max \left\{ \frac{P_i}{N_i}, \frac{P_j}{N_j} \right\} \\
&= \frac{N_1 P_1}{\bar{N}_z N_1} \left(\sum_{j=1}^z \frac{N_j}{\bar{N}_z} + \sum_{j=2}^z \frac{N_j}{\bar{N}_z} \right) + \frac{N_2 P_2}{\bar{N}_z N_2} \left(\sum_{j=2}^z \frac{N_j}{\bar{N}_z} + \sum_{j=3}^z \frac{N_j}{\bar{N}_z} \right) + \\
&\quad \dots + \frac{N_{z-1} P_{z-1}}{\bar{N}_z N_{z-1}} \left(\sum_{j=z-1}^z \frac{N_j}{\bar{N}_z} + \frac{N_z}{\bar{N}_z} \right) + \frac{N_z P_z N_z}{\bar{N}_z N_z \bar{N}_z} \\
&= \sum_{i=1}^z \frac{N_i P_i}{\bar{N}_z N_i} \left(\frac{N_i}{\bar{N}_z} + 2 \sum_{j=i+1}^z \frac{N_j}{\bar{N}_z} \right). \quad (14)
\end{aligned}$$

After subtracting $\sum_{i=1}^z \frac{N_i P_i}{\bar{N}_z N_i}$ we obtain

$$\frac{\bar{P}_z}{\bar{N}_z} G(\mathcal{A}; \zeta) = \sum_{i=1}^z \frac{N_i P_i}{\bar{N}_z N_i} \left(\frac{N_i}{\bar{N}_z} + 2 \sum_{j=i+1}^z \frac{N_j}{\bar{N}_z} \right) - \sum_{i=1}^z \frac{N_i P_i}{\bar{N}_z N_i} \quad (15)$$

$$= \sum_{i=1}^z \frac{N_i P_i}{\bar{N}_z N_i} \left(\frac{N_i}{\bar{N}_z} + 2 \sum_{j=i+1}^z \frac{N_j}{\bar{N}_z} - 1 \right) \quad (16)$$

$$= \sum_{i=1}^z \frac{N_i P_i}{\bar{N}_z N_i} \left(\frac{N_i}{\bar{N}_z} + \sum_{j=i+1}^z \frac{N_j}{\bar{N}_z} - 1 + \sum_{j=i+1}^z \frac{N_j}{\bar{N}_z} \right) \quad (17)$$

$$= \sum_{i=1}^z \frac{N_i P_i}{\bar{N}_z N_i} \left(\frac{\bar{N}_z - \bar{N}_i}{\bar{N}_z} - \frac{\bar{N}_{i-1}}{\bar{N}_z} \right). \quad (18)$$

As a result $\bar{P}_z G(\mathcal{A}; \zeta) = \sum_{i=1}^z P_i \left(\frac{\bar{N}_z - \bar{N}_i}{\bar{N}_z} - \frac{\bar{N}_{i-1}}{\bar{N}_z} \right)$, after dividing both sides by P we obtain the result in the lemma. ■

By substituting from Lemma 4 into the specification of Lemma 3 in (8) for $z \geq 1$, we obtain the specification of $UP(\cdot; \zeta)$ in the Theorem for $z \geq 1$:

$$UP(\mathcal{A}; \zeta) = \beta \cdot \frac{\bar{P}_z - \zeta \bar{N}_z}{P} + \gamma \cdot \frac{\bar{P}_z}{P} \cdot \frac{\bar{N}_z}{N} \cdot G(\mathcal{A}; \zeta) + \gamma \left(\frac{N - \bar{N}_z}{N} \right) \frac{\bar{P}_z - \zeta \bar{N}_z}{P}.$$

To complete the proof we show that *all axioms are independent*, meaning that it is possible to derive alternative functional forms for $UP(\cdot; \zeta)$ by dropping one of the axioms and considering all the others.

Drop NOR: consider (3) for $z \geq 1$ and set $UP(\cdot; \zeta) = k < 0$ in all other cases.

Drop TRAN: consider (3) with $\gamma = -1$ and $\beta = 0$ for $z \geq 1$, and set $UP(\cdot; \zeta) = -\sup\{\frac{\bar{N}_z}{N} \frac{\bar{P}_z}{P} \cdot G(\mathcal{A}; \zeta) + \frac{N - \bar{N}_z}{N} \frac{\bar{P}_z - \zeta \bar{N}_z}{P} : \mathcal{A} \in \Omega \text{ with } z \geq 1\}$ in all other cases.

Drop MON: consider (3) with $\gamma = 0$ and $\beta = -1$ for $z \geq 1$, and set $UP(\cdot; \zeta) = -1$ in all other cases.

Drop INV-PL: consider (3) multiplied by P/N for $z \geq 1$, and set $UP(\cdot; \zeta) = 0$ in all other cases.

Drop INV-T: consider

$$UP(\mathcal{A}; \zeta) := \frac{1}{\frac{P}{N}} \sum_{i=1}^z \left(\frac{P_i}{N_i} - \zeta \right) \cdot \left[\left(\frac{\bar{N}_i}{N} \frac{(N - \bar{N}_i)}{N} \right)^2 - \left(\frac{\bar{N}_{i-1}}{N} \frac{(N - \bar{N}_{i-1})}{N} \right)^2 \right]$$

with $\bar{N}_0 := 0$ for $z \geq 1$, and set $UP(\cdot; \zeta) = 0$ in all other cases.

Drop INV-S: consider

$$UP(\mathcal{A}; \zeta) := \frac{1}{\frac{P}{N}} \sum_{i=1}^z \left(\frac{N_i}{N} \right)^2 \cdot \left(\frac{P_i}{N_i} - \zeta \right) = \sum_{i=1}^z \frac{N_i}{N} \cdot \left(\frac{P_i - \zeta N_i}{P} \right)$$

for $z \geq 1$, and set $UP(\cdot; \zeta) = 0$ in all other cases.

Drop AGG: consider

$$\begin{aligned} UP(\mathcal{A}; \zeta) & : = \frac{1}{\frac{P}{N}} \sum_{i=1}^z \frac{N_i}{\bar{N}_z} \cdot \left(\frac{P_i}{N_i} - \zeta \right) \\ & = \frac{1}{\frac{P}{N}} \sum_{i=1}^z \frac{P_i - \zeta N_i}{\bar{N}_z} = \frac{\bar{P}_z - \zeta}{\frac{P}{N}} \end{aligned}$$

for $z \geq 1$, and set $UP(\cdot; \zeta) = 0$ in all other cases. QED.

A.2 Proof of Corollary 4

Proof. Let $p_i = \frac{P_i}{N_i}$ and $s_i = \frac{N_i}{N}$ denote the poverty incidence and population share of neighborhood i , respectively.

Let $\mathbf{p} = (p_1, \dots, p_n)^T$ be the $n \times 1$ vector of neighborhood poverty rates sorted in decreasing order and $\mathbf{s} = (s_1, \dots, s_n)^T$ be the $n \times 1$ vector of the corresponding population shares. A urban poverty configuration is fully identified by the pair (\mathbf{s}, \mathbf{p}) , and is used interchangeably. Let $\mathbf{1}_n$ being the $n \times 1$ vector with each element equal to 1, \mathbf{P} is the $n \times n$

skew-symmetric matrix:

$$\mathbf{P} = \frac{1}{\bar{p}} (\mathbf{1}_n \mathbf{p}^T - \mathbf{p} \mathbf{1}_n^T) = \begin{bmatrix} \frac{p_1 - p_1}{\bar{p}} & \dots & \frac{p_n - p_1}{\bar{p}} \\ \vdots & \ddots & \vdots \\ \frac{p_1 - p_n}{\bar{p}} & \dots & \frac{p_n - p_n}{\bar{p}} \end{bmatrix}, \quad (19)$$

where \bar{p} is the overall poverty rate in the city. The elements of \mathbf{P} are the n^2 relative pairwise differences between the neighborhood poverty incidences as ordered in \mathbf{p} . Let $\mathbf{S} = \text{diag}\{\mathbf{s}\}$ be the $n \times n$ diagonal matrix with diagonal elements equal to the population shares in \mathbf{s} , and \mathbf{G} be a $n \times n$ G -matrix (a skew-symmetric matrix whose diagonal elements are equal to 0, with upper diagonal elements equal to -1 and lower diagonal elements equal to 1) (Silber 1989). The Gini index of urban poverty is expressed in matrix form:

$$G(\mathbf{s}, \mathbf{p}) = \frac{1}{2} \text{tr}(\tilde{\mathbf{G}} \mathbf{P}^T), \quad (20)$$

where the matrix $\tilde{\mathbf{G}} = \mathbf{S} \mathbf{G} \mathbf{S}$ is the weighting G -matrix, a generalization of the G -matrix introduced by Mussini and Grossi (2015) to add weights in the calculation of the Gini index. The change in urban poverty from t to t' is measured by the difference between the Gini index in t' and the Gini index in t :

$$\Delta UP = G(\mathbf{s}_{t'}, \mathbf{p}_{t'}) - G(\mathbf{s}_t, \mathbf{p}_t) = \frac{1}{2} \text{tr}(\tilde{\mathbf{G}}_{t'} \mathbf{P}_{t'}^T) - \frac{1}{2} \text{tr}(\tilde{\mathbf{G}}_t \mathbf{P}_t^T). \quad (21)$$

Equation (21) can be broken down into three components by applying the matrix approach used in Mussini and Grossi (2015) and in Mussini (2017). The three components separate the contributions of changes in neighborhood population shares, ranking of neighborhoods and disparities between neighborhood poverty rates. Let $\mathbf{s}_{t|t'}$ stand for the $n \times 1$ vector of the t population shares arranged by the decreasing order of the corresponding t' poverty rates. Let $\lambda = \bar{p}_{t'}/\bar{p}_{t|t}$ be the ratio of the actual t' overall poverty rate to the fictitious t' overall poverty rate which is the weighted average of t' poverty rates where the weights are the corresponding population shares in t . After defining $\mathbf{S}_{t|t'} = \text{diag}\{\mathbf{s}_{t|t'}\}$, the Gini index of t' neighborhood poverty rates calculated by using the t neighborhood population shares is

$$\begin{aligned} G(\mathbf{s}_{t|t'}, \mathbf{p}_{t'}) &= \frac{1}{2} \text{tr}(\mathbf{S}_{t|t'} \mathbf{G} \mathbf{S}_{t|t'} \lambda \mathbf{P}_{t'}^T) \\ &= \frac{1}{2} \text{tr}(\tilde{\mathbf{G}}_{t|t'} \lambda \mathbf{P}_{t'}^T) \end{aligned} \quad (22)$$

where $\tilde{\mathbf{G}}_{t|t'} = \mathbf{S}_{t|t'} \mathbf{G} \mathbf{S}_{t|t'}$ is the weighting G -matrix obtained by using the neighborhood population shares in t instead of those in t' . In equation (22), the multiplication of $\mathbf{P}_{t'}^T$ by λ ensures that the pairwise differences between the t' neighborhood poverty incidences are divided by $\bar{p}_{t|t}$ instead of $\bar{p}_{t'}$. By adding and subtracting $G(\mathbf{s}_{t|t'}, \mathbf{p}_{t'})$ in equation (21), the contribution to ΔUP due to changes in neighborhood population shares can be separated

from that attributable to changes in disparities between neighborhood poverty rates:

$$\begin{aligned}
\Delta UP &= \left[\frac{1}{2} \text{tr} \left(\tilde{\mathbf{G}}_{t'} \mathbf{P}_{t'}^T \right) - \frac{1}{2} \text{tr} \left(\tilde{\mathbf{G}}_{t|t'} \lambda \mathbf{P}_{t'}^T \right) \right] + \left[\frac{1}{2} \text{tr} \left(\tilde{\mathbf{G}}_{t|t'} \lambda \mathbf{P}_{t'}^T \right) - \frac{1}{2} \text{tr} \left(\tilde{\mathbf{G}}_t \mathbf{P}_t^T \right) \right] \\
&= \frac{1}{2} \text{tr} \left(\mathbf{W} \mathbf{P}_{t'}^T \right) + \left[\frac{1}{2} \text{tr} \left(\tilde{\mathbf{G}}_{t|t'} \lambda \mathbf{P}_{t'}^T \right) - \frac{1}{2} \text{tr} \left(\tilde{\mathbf{G}}_t \mathbf{P}_t^T \right) \right] \\
&= W + \left[\frac{1}{2} \text{tr} \left(\tilde{\mathbf{G}}_{t|t'} \lambda \mathbf{P}_{t'}^T \right) - \frac{1}{2} \text{tr} \left(\tilde{\mathbf{G}}_t \mathbf{P}_t^T \right) \right],
\end{aligned} \tag{23}$$

where $\mathbf{W} = \tilde{\mathbf{G}}_{t'} - \lambda \tilde{\mathbf{G}}_{t|t'}$. Component W measures the effect of changes in neighborhood population shares. A positive value of W indicates that the weights assigned to more unequal pairs of neighborhoods are larger in t' than in t , increasing urban poverty from t to t' . A negative value of W indicates that the weights assigned to more unequal pairs of neighborhoods are smaller in t' than in t , reducing urban poverty.

The difference enclosed within square brackets on the right-hand side of equation (23) can be additively split into two components: one component measuring the re-ranking of neighborhoods, a second component measuring the change in disparities between neighborhood poverty rates. Let $\mathbf{p}_{t'|t}$ be the $n \times 1$ vector of t' neighborhood poverty rates sorted in decreasing order of the respective t neighborhood poverty rates, and \mathbf{B} be the $n \times n$ permutation matrix re-arranging the elements of $\mathbf{p}_{t'}$ to obtain $\mathbf{p}_{t'|t}$, that is $\mathbf{p}_{t'|t} = \mathbf{B} \mathbf{p}_{t'}$. Matrix $\mathbf{P}_{t'|t} = (1/\bar{p}_{t'|t}) \left(\mathbf{1}_n \mathbf{p}_{t'|t}^T - \mathbf{p}_{t'|t} \mathbf{1}_n^T \right)$ contains the n^2 relative pairwise differences between the neighborhood poverty rates as arranged in $\mathbf{p}_{t'|t}$. The concentration index of the t' neighborhood poverty rates sorted by the t neighborhood poverty rates, calculated by using the t population shares, is defined as follows:

$$C \left(\mathbf{s}_t, \mathbf{p}_{t'|t} \right) = \frac{1}{2} \text{tr} \left(\tilde{\mathbf{G}}_t \mathbf{P}_{t'|t}^T \right). \tag{24}$$

By using permutation matrix \mathbf{B} , the concentration index $C \left(\mathbf{s}_t, \mathbf{p}_{t'|t} \right)$ can be re-written as a function of $\mathbf{P}_{t'}$ instead of $\mathbf{P}_{t'|t}$. Since $\mathbf{P}_{t'|t} = \mathbf{B} \lambda \mathbf{P}_{t'}^T \mathbf{B}^T$, the concentration index $C \left(\mathbf{s}_t, \mathbf{p}_{t'|t} \right)$ expressed as a function of $\mathbf{P}_{t'}$ becomes

$$\begin{aligned}
C \left(\mathbf{s}_t, \mathbf{p}_{t'|t} \right) &= \frac{1}{2} \text{tr} \left(\tilde{\mathbf{G}}_t \mathbf{B} \lambda \mathbf{P}_{t'}^T \mathbf{B}^T \right) \\
&= \frac{1}{2} \text{tr} \left(\mathbf{B}^T \tilde{\mathbf{G}}_t \mathbf{B} \lambda \mathbf{P}_{t'}^T \right).
\end{aligned} \tag{25}$$

By adding $C \left(\mathbf{s}_t, \mathbf{p}_{t'|t} \right)$ as expressed in (24) and subtracting it as expressed in (25) to the difference enclosed within square brackets on the right-hand side of equation (23), we

obtain

$$\begin{aligned}
\frac{1}{2}tr\left(\tilde{\mathbf{G}}_{t|t'}\lambda\mathbf{P}_{t'}^T\right) - \frac{1}{2}tr\left(\tilde{\mathbf{G}}_t\mathbf{P}_t^T\right) &= \left[\frac{1}{2}tr\left(\tilde{\mathbf{G}}_{t|t'}\lambda\mathbf{P}_{t'}^T\right) - \frac{1}{2}tr\left(\mathbf{B}^T\tilde{\mathbf{G}}_t\mathbf{B}\lambda\mathbf{P}_{t'}^T\right)\right] \\
&+ \left[\frac{1}{2}tr\left(\tilde{\mathbf{G}}_t\mathbf{P}_{t'|t}^T\right) - \frac{1}{2}tr\left(\tilde{\mathbf{G}}_t\mathbf{P}_t^T\right)\right] \\
&= \frac{1}{2}tr\left[\left(\tilde{\mathbf{G}}_{t|t'} - \mathbf{B}^T\tilde{\mathbf{G}}_t\mathbf{B}\right)\lambda\mathbf{P}_{t'}^T\right] \\
&+ \frac{1}{2}tr\left[\tilde{\mathbf{G}}_t\left(\mathbf{P}_{t'|t}^T - \mathbf{P}_t^T\right)\right] \\
&= \frac{1}{2}tr\left(\mathbf{R}\lambda\mathbf{P}_{t'}^T\right) + \frac{1}{2}tr\left(\tilde{\mathbf{G}}_t\mathbf{D}^T\right) \\
&= R + D,
\end{aligned} \tag{26}$$

where $\mathbf{R} = \tilde{\mathbf{G}}_{t|t'} - \mathbf{B}^T\tilde{\mathbf{G}}_t\mathbf{B}$ and $\mathbf{D} = \mathbf{P}_{t'|t} - \mathbf{P}_t$. Component R measures the effect of re-ranking of neighborhoods from t to t' and its contribution to the change in urban poverty is always non-negative. The nonzero elements of \mathbf{R} indicate the pairs of neighborhoods which have re-ranked from t to t' .

Component D measures the effect of disproportionate changes in neighborhood poverty rates. The generic (i, j) -th element of \mathbf{D} compares the relative difference between the t poverty rates of the neighborhoods in positions j and i in \mathbf{p}_t with the relative difference between the t' poverty rates of the same two neighborhoods in $\mathbf{p}_{t'|t}$. A positive (negative) value of D indicates that relative disparities in neighborhood poverty rates have increased (decreased) from t to t' , increasing (reducing) urban poverty. If all neighborhood poverty rates have changed in the same proportion from t to t' , then $D = 0$.

Given equations (23) and (26), a three-term decomposition of ΔUP is obtained:

$$\Delta UP = \frac{1}{2}tr\left(\mathbf{W}\mathbf{P}_{t'}^T\right) + \frac{1}{2}tr\left(\mathbf{R}\lambda\mathbf{P}_{t'}^T\right) + \frac{1}{2}tr\left(\tilde{\mathbf{G}}_t\mathbf{D}^T\right) = W + R + D. \tag{27}$$

Since component D would not reveal changes in neighborhood poverty rates if all neighborhood poverty rates changed in the same proportion, this component is split into two further terms: one measuring the change in the city poverty rate, the second measuring the changes in disparities between neighborhood poverty rates by assuming that the city poverty rate remains the same from t to t' . Let c stand for the change in the city poverty rate by assuming that neighborhood population shares are unchanged from t to t' :

$$c = \frac{\bar{p}_{t'|t} - \bar{p}_t}{\bar{p}_t}. \tag{28}$$

Let $\mathbf{p}_{t'|t}^c = \mathbf{p}_t + c\mathbf{p}_t$ be the vector of neighborhood poverty rates we would observe in t' if every neighborhood poverty rate changed by proportion c . This implies that $\bar{p}_{t'|t}^c = \bar{p}_{t'|t}$. Vector $\mathbf{p}_{t'|t}$ can be expressed as

$$\mathbf{p}_{t'|t} = \mathbf{p}_{t'|t}^c + \mathbf{p}_{t'|t}^\delta,$$

where the elements of vector $\mathbf{p}_{t'|t}^\delta$ are the element-by-element differences between vectors $\mathbf{p}_{t'|t}$ and $\mathbf{p}_{t'|t}^c$. Since $\mathbf{p}_{t'|t}^c = \mathbf{p}_t + c\mathbf{P}_t$, $\mathbf{p}_{t'|t}$ can be re-written as

$$\begin{aligned}\mathbf{p}_{t'|t} &= \underbrace{\mathbf{p}_t + \mathbf{p}_{t'|t}^\delta}_{\mathbf{p}_{t'|t}^e} + c\mathbf{P}_t \\ &= \mathbf{p}_{t'|t}^e + c\mathbf{P}_t,\end{aligned}\tag{29}$$

where the elements of $\mathbf{p}_{t'|t}^e$ account for disproportionate changes in neighborhood poverty rates from t to t' , as $\mathbf{p}_{t'|t}^e$ would equal \mathbf{p}_t if there were no disproportionate changes in neighborhood poverty rates. Given equations above, matrix $\mathbf{P}_{t'|t}$ can be written as

$$\begin{aligned}\mathbf{P}_{t'|t} &= (1/\bar{p}_{t'|t}) \left(\mathbf{1}_n \mathbf{P}_{t'|t}^T - \mathbf{p}_{t'|t} \mathbf{1}_n^T \right) \\ &= \frac{1}{1+c} \begin{bmatrix} \frac{p_{1,t'|t}^e - p_{1,t'|t}^c}{\bar{p}_t} & \dots & \frac{p_{n,t'|t}^e - p_{n,t'|t}^c}{\bar{p}_t} \\ \vdots & \ddots & \vdots \\ \frac{p_{1,t'|t}^e - p_{n,t'|t}^c}{\bar{p}_t} & \dots & \frac{p_{n,t'|t}^e - p_{n,t'|t}^c}{\bar{p}_t} \end{bmatrix} + \frac{c}{1+c} \begin{bmatrix} \frac{p_{1,t} - p_{1,t}}{\bar{p}_t} & \dots & \frac{p_{n,t} - p_{1,t}}{\bar{p}_t} \\ \vdots & \ddots & \vdots \\ \frac{p_{1,t} - p_{n,t}}{\bar{p}_t} & \dots & \frac{p_{n,t} - p_{n,t}}{\bar{p}_t} \end{bmatrix} \\ &= \frac{1}{1+c} \mathbf{P}_{t'|t}^e + \frac{c}{1+c} \mathbf{P}_t.\end{aligned}\tag{30}$$

Since matrix \mathbf{D} in equation (27) is obtained by subtracting \mathbf{P}_t from $\mathbf{P}_{t'|t}$, \mathbf{D} can be re-written as

$$\begin{aligned}\mathbf{D} &= \mathbf{P}_{t'|t} - \mathbf{P}_t \\ &= \frac{1}{1+c} \mathbf{P}_{t'|t}^e + \frac{c}{1+c} \mathbf{P}_t - \mathbf{P}_t \\ &= \underbrace{\left(\frac{1}{1+c} \right)}_C \underbrace{\left(\mathbf{P}_{t'|t}^e - \mathbf{P}_t \right)}_E \\ &= C \cdot \mathbf{E}\end{aligned}\tag{31}$$

By replacing \mathbf{D} in equation (27) with its expression in equation (31), the decomposition of the change in urban poverty becomes

$$\Delta UP = \frac{1}{2} \text{tr}(\mathbf{W} \mathbf{P}_{t'}^T) + \frac{1}{2} \text{tr}(\mathbf{R} \lambda \mathbf{P}_{t'}^T) + C \cdot \frac{1}{2} \text{tr}(\tilde{\mathbf{G}}_t \mathbf{E}^T) = W + R + C \cdot E.\tag{32}$$

■

A.3 Proof of Corollary 5

Proof. Building on the Rey and Smith (2013) spatial decomposition of the Gini index and the spatial decomposition of the change in inequality in Mussini (2020), ΔUP , W , R and E can be broken down into spatial components. Let \mathbf{N}_t be the $n \times n$ spatial weights matrix having its (i, j) -th entry equal to 1 if and only if the (i, j) -th element of \mathbf{P}_t is the relative difference between the poverty rates of two neighborhoods that are spatially close, otherwise the (i, j) -th element of \mathbf{N}_t is 0. Using the Hadamard product,²¹ the relative

²¹Let \mathbf{X} and \mathbf{Y} be $k \times k$ matrices. The Hadamard product $\mathbf{X} \odot \mathbf{Y}$ is defined as the $k \times k$ matrix with the (i, j) -th element equal to $x_{ij}y_{ij}$.

pairwise differences between the poverty rates of neighborhoods that are spatially close can be selected from \mathbf{P}_t :

$$\mathbf{P}_{N,t} = \mathbf{N}_t \odot \mathbf{P}_t. \quad (33)$$

For each pair of neighborhoods, the relative difference between their t' poverty rates in $\mathbf{P}_{t'|t}^e$ has the same position as the relative difference between their t poverty rates in \mathbf{P}_t . Thus, \mathbf{N}_t also selects the relative pairwise differences between neighbors from $\mathbf{P}_{t'|t}^e$:

$$\mathbf{P}_{N,t'|t}^e = \mathbf{N}_t \odot \mathbf{P}_{t'|t}^e. \quad (34)$$

Since $\mathbf{E} = \mathbf{P}_{t'|t}^e - \mathbf{P}_t$, the Hadamard product between \mathbf{N}_t and \mathbf{E} is a matrix with nonzero elements equal to the elements of \mathbf{E} pertaining to neighborhoods that are spatially close:

$$\mathbf{E}_N = \mathbf{P}_{N,t'|t}^e - \mathbf{P}_{N,t} = \mathbf{N}_t \odot (\mathbf{P}_{t'|t}^e - \mathbf{P}_t) = \mathbf{N}_t \odot \mathbf{E}. \quad (35)$$

Let $\mathbf{N}_{t'}$ be the $n \times n$ spatial weights matrix having its (i, j) -th entry equal to 1 if and only if the (i, j) -th element of $\mathbf{P}_{t'}$ is the relative difference between the poverty rates of two neighborhoods that are spatially close, otherwise the (i, j) -th element of $\mathbf{N}_{t'}$ is 0. The Hadamard product of $\mathbf{N}_{t'}$ and $\mathbf{P}_{t'}$ is the matrix

$$\mathbf{P}_{N,t'} = \mathbf{N}_{t'} \odot \mathbf{P}_{t'}. \quad (36)$$

The nonzero elements of $\mathbf{P}_{N,t'}$ are the relative pairwise differences between the t' poverty rates of neighborhoods that are in spatial proximity.

The decomposition of the change in the neighborhood component of urban poverty is obtained by replacing $\mathbf{P}_{t'}$ and \mathbf{E} in equation (32) with $\mathbf{P}_{N,t'}$ and \mathbf{E}_N , respectively:

$$\Delta UP_N = \frac{1}{2}tr(\mathbf{W}\mathbf{P}_{N,t'}^T) + \frac{1}{2}tr(\mathbf{R}\lambda\mathbf{P}_{N,t'}^T) + C \cdot \frac{1}{2}tr(\tilde{\mathbf{G}}_t\mathbf{E}_N^T) = W_N + R_N + C \cdot E_N. \quad (37)$$

Let \mathbf{J}_n be the matrix with diagonal elements equal to 0 and extra-diagonal elements equal to 1, the matrix with nonzero elements equal to the relative pairwise differences between the t' poverty rates of neighborhoods that are not in spatial proximity is

$$\mathbf{P}_{nN,t'} = (\mathbf{J}_n - \mathbf{N}_{t'}) \odot \mathbf{P}_{t'}. \quad (38)$$

The matrix selecting the elements of \mathbf{E} pertaining to the pairs of neighborhoods that are not spatially close is

$$\mathbf{E}_{nN} = (\mathbf{J}_n - \mathbf{N}_t) \odot \mathbf{E}. \quad (39)$$

The decomposition of the change in the non-neighborhood component of urban poverty is obtained by replacing $\mathbf{P}_{t'}$ and \mathbf{E} in equation (32) with $\mathbf{P}_{nN,t'}$ and \mathbf{E}_{nN} , respectively:

$$\Delta UP_{nN} = \frac{1}{2}tr(\mathbf{W}\mathbf{P}_{nN,t'}^T) + \frac{1}{2}tr(\mathbf{R}\lambda\mathbf{P}_{nN,t'}^T) + C \cdot \frac{1}{2}tr(\tilde{\mathbf{G}}_t\mathbf{E}_{nN}^T) = W_{nN} + R_{nN} + C \cdot E_{nN}. \quad (40)$$

Given equations (40) and (37), the spatial decomposition of the change in urban poverty is

$$\Delta UP = W_N + W_{nN} + R_N + R_{nN} + C \cdot (E_N + E_{nN}). \quad (41)$$

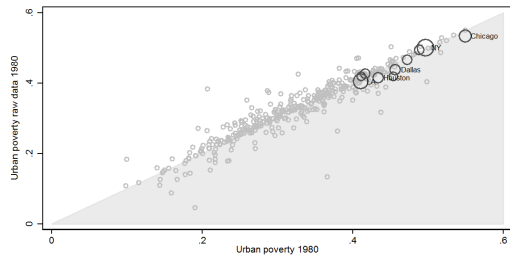
■

B Additional results

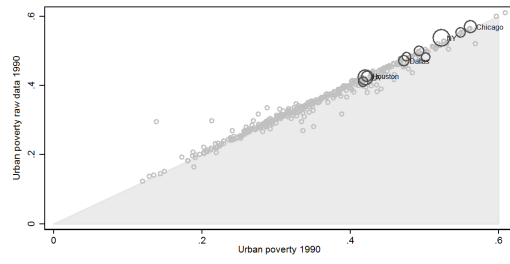
variable	statistics	1980-1990	1990-2000	2000-2008	2008-2012	2012-2014
ΔUP	mean	.0314	-.0038	.0153	-.016	-.0016
	pc(25)	.0026	-.0258	-.0109	-.0337	-.0135
	pc(75)	.0592	.015	.0401	.0009	.0097
$\Delta CP(20)$	mean	.0958	-.0474	.104	.0491	-.0153
	pc(25)	.0089	-.1	.0305	-.0028	-.0487
	pc(75)	.1722	.0105	.1697	.0997	.0123
$\Delta CP(40)$	mean	.0578	-.0393	.0461	.021	-.0118
	pc(25)	0	-.0736	0	-.0013	-.0325
	pc(75)	.0954	0	.0841	.0535	.0055
W	mean	.0041	.0008	.003	.0006	.0004
	pc(25)	-.0043	-.005	-.0018	-.0015	-.0007
	pc(75)	.0138	.0062	.0069	.0028	.0016
R	mean	.0525	.0462	.0664	.0556	.0277
	pc(25)	.0357	.0318	.0463	.0391	.0192
	pc(75)	.0628	.0567	.0808	.0663	.0325
E	mean	-.0328	-.0507	-.0682	-.082	-.0291
	pc(25)	-.0612	-.0733	-.1017	-.107	-.0411
	pc(75)	.0046	-.0275	-.034	-.0562	-.0157
C	mean	.8869	1.044	.8257	.8923	1.0289
	pc(25)	.7709	.9462	.7387	.8437	.9955
	pc(75)	1.0091	1.1325	.9054	.9308	1.0556
$D = C \cdot E$	mean	-.0252	-.0508	-.054	-.0722	-.0297
	pc(25)	-.0527	-.0727	-.0808	-.0925	-.0419
	pc(75)	.0042	-.028	-.0276	-.0515	-.016
β (log-log)	mean	-.1306	-.1935	-.1694	-.2666	-.1228
	pc(25)	-.2477	-.2679	-.2492	-.3386	-.1712
	pc(75)	-.0247	-.1094	-.0802	-.1892	-.067

Table 2: Summary statistics of changes in urban poverty concentration, all 395 American MSA.

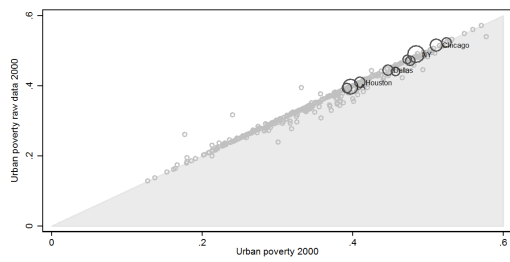
Figure 6: Urban poverty calculated by using balanced longitudinal data and raw data on census tracts in American MSAs, from 1980 to 2014.



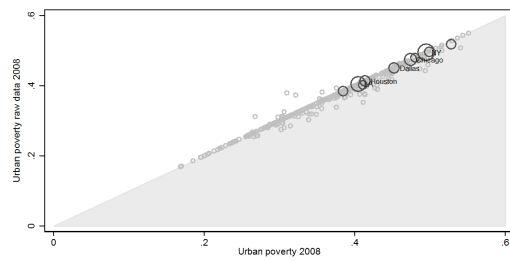
(a) 1980



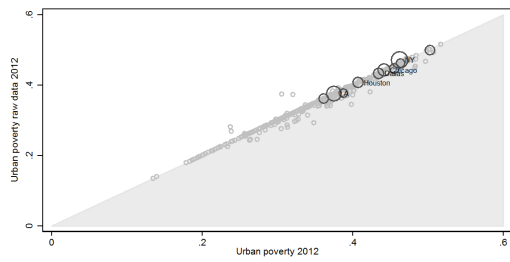
(b) 1990



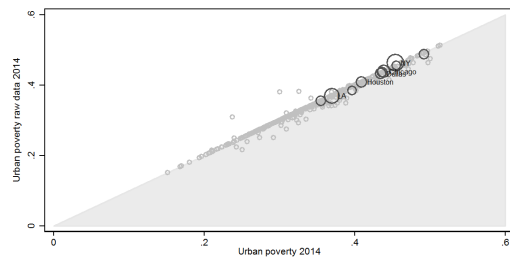
(c) 2000



(d) 2008



(e) 2012

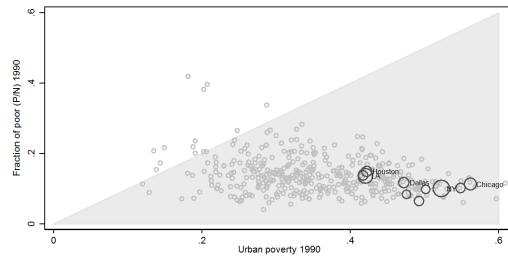


(f) 2014

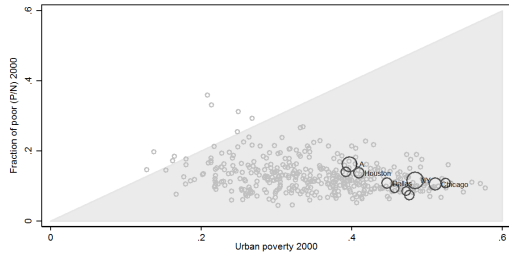
Figure 7: Urban poverty and poverty incidence in the city, from 1980 to 2014.



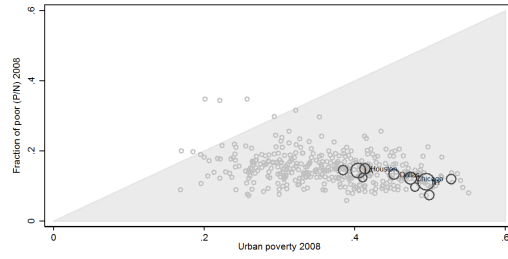
(a) 1980



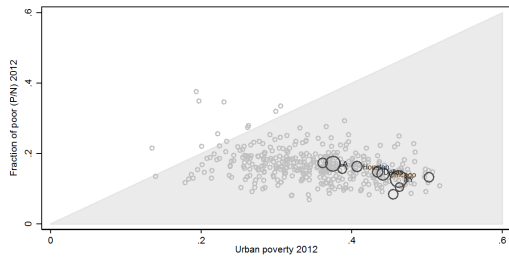
(b) 1990



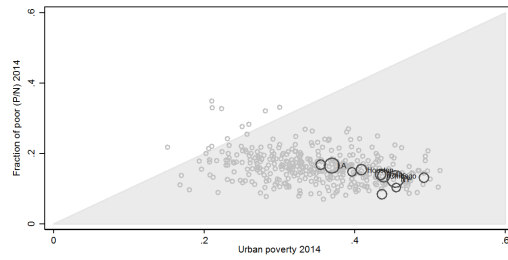
(c) 2000



(d) 2008

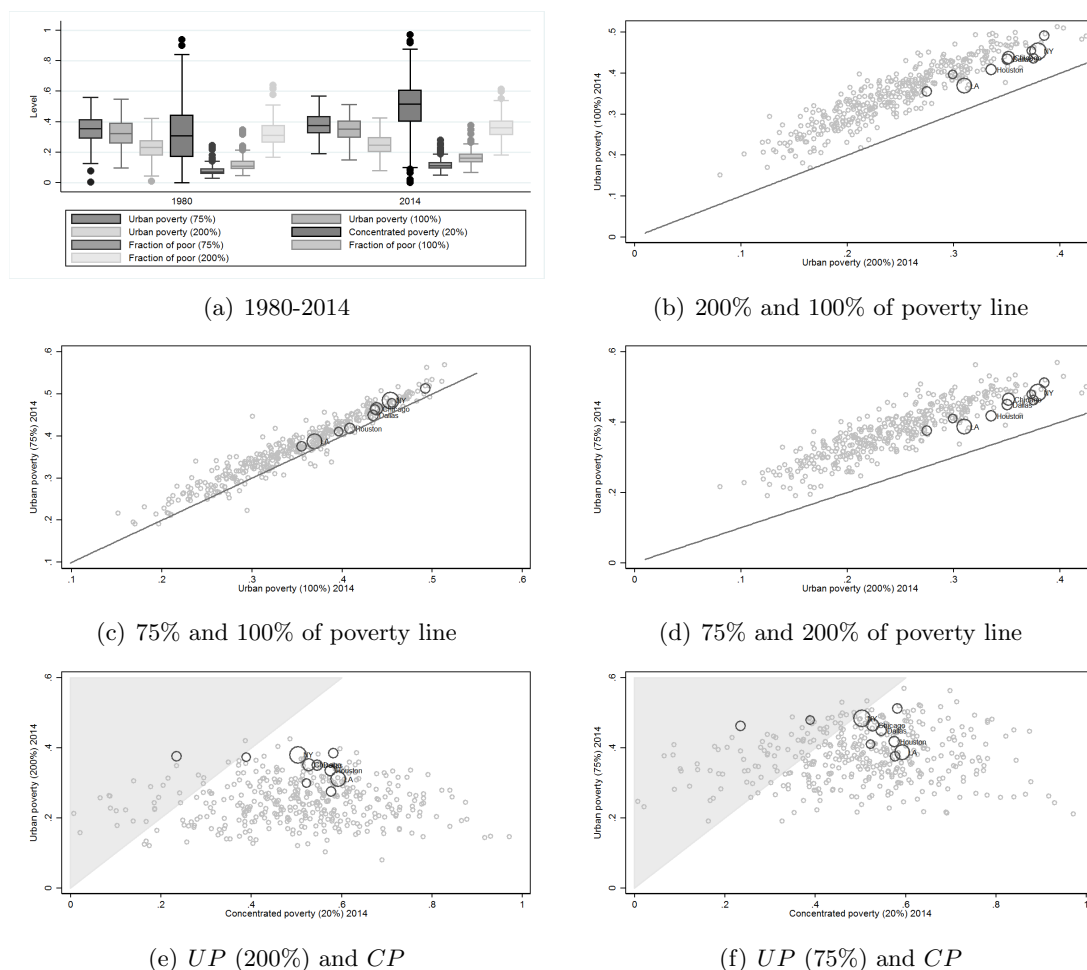


(e) 2012



(f) 2014

Figure 8: Urban poverty estimates based on poverty thresholds at 75%, 100% and 200% of the official poverty line, years 1980 and 2014.



References

- Andreoli, F. (2018). Robust inference for inverse stochastic dominance, *Journal of Business & Economic Statistics* **36**(1): 146–159.
- Andreoli, F. and Peluso, E. (2018). So close yet so unequal: Neighborhood inequality in American cities, *ECINEQ Working paper 2018-477*.
- Andreoli, F. and Zoli, C. (2014). Measuring dissimilarity, *Working Papers Series, Department of Economics, University of Verona, WP23*.
- Andreoli, F. and Zoli, C. (2020). From unidimensional to multidimensional inequality: a review., *METRON* **78**: 5–42.
- Atkinson, A. B. (1970). On the measurement of inequality, *Journal of Economic Theory* **2**: 244–263.
- Baum-Snow, N. and Pavan, R. (2013). Inequality and city size, *The Review of Economics and Statistics* **95**(5): 1535–1548.

- Bayer, P. and Timmins, C. (2005). On the equilibrium properties of locational sorting models, *Journal of Urban Economics* **57**(3): 462 – 477.
- Bosmans, K. (2014). Distribution-sensitivity of rank-dependent poverty measures, *Journal of Mathematical Economics* **51**(C): 69–76.
- Chetty, R. and Hendren, N. (2018). The impacts of neighborhoods on intergenerational mobility ii: County-level estimates*, *The Quarterly Journal of Economics* **133**(3): 1163–1228.
- Chetty, R., Hendren, N. and Katz, L. F. (2016). The effects of exposure to better neighborhoods on children: New evidence from the moving to opportunity experiment, *American Economic Review* **106**(4): 855–902.
- Conley, T. G. and Topa, G. (2002). Socio-economic distance and spatial patterns in unemployment, *Journal of Applied Econometrics* **17**(4): 303–327.
- Cowell, F. (2000). Measurement of inequality, Vol. 1 of *Handbook of Income Distribution*, Elsevier, chapter 2, pp. 87 – 166.
- Decancq, K., Fleurbaey, M. and Maniquet, F. (2019). Multidimensional poverty measurement with individual preferences, *The Journal of Economic Inequality* **17**(1): 29–49.
- Ebert, U. (2010). The decomposition of inequality reconsidered: Weakly decomposable measures, *Mathematical Social Sciences* **60**: 94–103.
- Espa, G., Filippini, D., Giuliani, D. and Piacentino, D. (2014). Decomposing regional business change at plant level in italy: A novel spatial shift-share approach, *Papers in Regional Science* **93**(S1): S113–S135.
- Glaeser, E., Resseger, M. and Tobio, K. (2009). Inequality in cities, *Journal of Regional Science* **49**(4): 617–646.
- Iceland, J. and Hernandez, E. (2017). Understanding trends in concentrated poverty: 1980-2014, *Social Science Research* **62**: 75 – 95.
- Jargowsky, P. A. (1996). Take the money and run: Economic segregation in U.S. metropolitan areas, *American Sociological Review* **61**(6): 984–998.
- Jargowsky, P. A. (1997). *Poverty and Place: Ghettos, Barrios, and the American City*, New York: Russell Sage Foundation.
- Jargowsky, P. A. (2015). Architecture of segregation. Civil unrest, the concentration of poverty, and public policy, *Technical report*, The Century Foundation.
- Jargowsky, P. A. and Bane, M. J. (1991). *Ghetto Poverty in the United States, 1970-1980*, Vol. The Urban Underclass, Washington, D.C.: The Brookings Institution., pp. 235–273.
- Jenkins, S. P., Brandolini, A., Micklewright, J. and Nolan, B. (2013). *The Great Recession and the distribution of household income*, Oxford University Press, Oxford, UK.
- Jenkins, S. P. and Van Kerm, P. (2016). Assessing individual income growth, *Economica* **83**(332): 679–703.

- Kneebone, E. (2016). *The changing geography of disadvantage*, Shared Prosperity in America's Communities, University of Pennsylvania Press, Philadelphia, chapter 3, pp. 41–56.
- Koshevoy, G. A. and Mosler, K. (1997). Multivariate Gini indices, *Journal of Multivariate Analysis* **60**(2): 252–276.
- Logan, J. R., Xu, Z. and Stults, B. J. (2014). Interpolating U.S. decennial census tract data from as early as 1970 to 2010: A longitudinal tract database, *The Professional Geographer* **66**(3): 412–420. PMID: 25140068.
- Ludwig, J., Duncan, G. J., Gennetian, L. A., Katz, L. F., Kessler, R. C., Kling, J. R. and Sanbonmatsu, L. (2012). Neighborhood effects on the long-term well-being of low-income adults, *Science* **337**(6101): 1505–1510.
- Ludwig, J., Duncan, G. J., Gennetian, L. A., Katz, L. F., Kessler, R. C., Kling, J. R. and Sanbonmatsu, L. (2013). Long-term neighborhood effects on low-income families: Evidence from Moving to Opportunity, *American Economic Review* **103**(3): 226–31.
- Ludwig, J., Sanbonmatsu, L., Gennetian, L., Adam, E., Duncan, G. J., Katz, L. F., Kessler, R. C., Kling, J. R., Lindau, S. T., Whitaker, R. C. and McDade, T. W. (2011). Neighborhoods, obesity, and diabetes - A randomized social experiment, *New England Journal of Medicine* **365**(16): 1509–1519. PMID: 22010917.
- Massey, D. S. and Eggers, M. L. (1990). The ecology of inequality: Minorities and the concentration of poverty, 1970-1980, *American Journal of Sociology* **95**(5): 1153–1188.
- Massey, D. S., Gross, A. B. and Eggers, M. L. (1991). Segregation, the concentration of poverty, and the life chances of individuals, *Social Science Research* **20**(4): 397 – 420.
- Moretti, E. (2013). Real wage inequality, *American Economic Journal: Applied Economics* **5**(1): 65–103.
- Mussini, M. (2017). Decomposing Changes in Inequality and Welfare Between EU Regions: The Roles of Population Change, Re-Ranking and Income Growth, *Social Indicators Research* **130**: 455–478.
- Mussini, M. (2020). Inequality and convergence in energy intensity in the European Union, *Applied Energy* **261**: 114371.
- Mussini, M. and Grossi, L. (2015). Decomposing changes in CO₂ emission inequality over time: the roles of re-ranking and changes in per capita CO₂ emission disparities, *Energy Economics* **49**: 274–281.
- Oreopoulos, P. (2003). The long-run consequences of living in a poor neighborhood, *The Quarterly Journal of Economics* **118**(4): 1533–1575.
- Ravallion, M. and Chen, S. (2011). Weakly relative poverty, *The Review of Economics and Statistics* **93**(4): 1251–1261.
- Rey, S. J. and Smith, R. J. (2013). A spatial decomposition of the Gini coefficient, *Letters in Spatial and Resource Sciences* **6**: 55–70.
- Sen, A. (1976). Poverty: An ordinal approach to measurement, *Econometrica* **44**(2): 219–231.

- Shorrocks, A. F. (1982). Inequality decomposition by factor components, *Econometrica* **50**(1): 193–211.
- Shorrocks, A. F. (1995). Revisiting the Sen poverty index, *Econometrica* **63**(5): 1225–1230.
- Silber, J. (1989). Factor components, population subgroups and the computation of the Gini index of inequality, *The Review of Economics and Statistics* **71**: 107–115.
- Thompson, J. P. and Smeeding, T. M. (2013). Inequality and poverty in the United States: The aftermath of the Great Recession, *FEDS Working Paper No. 2013-51* .
- Thon, D. (1979). On measuring poverty, *Review of Income and Wealth* **25**(4): 429–439.
- Watson, T. (2009). Inequality and the measurement of residential segregation by income in American neighborhoods, *Review of Income and Wealth* **55**(3): 820–844.
- Wheeler, C. H. and La Jeunesse, E. A. (2008). Trends in neighborhood income inequality in the U.S.: 1980–2000, *Journal of Regional Science* **48**(5): 879–891.
- Wilson, W. (1987). *The truly disadvantaged: The inner city, the underclasses and public policy*, University of Chicago Press, Chicago.
- Zoli, C. (1999). Intersecting generalized Lorenz curves and the Gini index, *Social Choice and Welfare* **16**: 183–196.