Urban poverty:
Theory and evidence from gentrifying American cities

Francesco Andreoli* Mauro Mussini†

Preliminary version not be cited or quoted: February 2018

Abstract

The proportion of poor people living in neighborhoods where poverty is highly concentrated is widely accepted as a policy-relevant measure of urban poverty in large metro areas. We challenge this view by developing new measures of urban poverty that i) capture aspects of the incidence and distribution of poverty across neighborhoods and ii) are consistent with the idea that poor people living in high-poverty neighborhoods face a double welfare burden of poverty. We demonstrate that there is only a measure that is consistent with a parsimonious axiomatic model for urban poverty. Panel variations of this measure are additively decomposed into the contribution of demographic, spatial and neighborhood-level poverty convergence effects. We collect new evidence of heterogenous patterns and trends of urban poverty across American metro areas over the last 35 years. Reduced form models allow to recover the implications of (income) sorting, affordable housing and of rising gentrification on different components of urban poverty across American cities.

Keywords: Concentrated poverty, axiomatic, decomposition, American metro areas, gentrification.

JEL codes: C34, D31, H24, P25.

*Corresponding author. Luxembourg Institute of Socio-Economic Research, LISER. MSH, 11 Porte des Sciences, L-4366 Esch-sur-Alzette/Belval Campus, Luxembourg. E-mail: francesco.andreoli@liser.lu.

†DSE, University of Verona. Via Cantarane 24, 37129 Verona, Italy. E-mail: mauro.mussini@univr.it.

1 Introduction

Cities are the most unequal places in America (Moretti 2013, Baum-Snow and Pavan 2013, Andreoli and Peluso 2017) and, increasingly so, the places where opportunities are created and redistributed. Variability in earnings inequality across cities is associated with heterogeneity in the economic mobility prospects of the children exposed to these inequalities (Chetty, Hendren, Kline and Saez 2014, Chetty and Hendren 2016). Inequalities and opportunities, however, are not equally distributed within the cities: in large American metro areas one can observe both neighborhoods where the local income distribution represents roughly the income distribution observed in the city, and more segregated neighborhoods where residents are systematically exposed to the extremes of the income distribution.

People living in poor neighborhoods face and indirect burden of local poverty via neighborhood and peer effects on their labor market outcomes (Conley and Topa 2002) and economic mobility prospects (Chetty et al. 2014). Even more so, poor people living in these neighborhood face a *double* burden of poverty, being poor in places where poverty is substantial. Evidence from the Moving to Opportunities experiment, offering vouchers for housing to poor people living in high poverty neighborhoods, highlights that income externalities generated from poverty concentration in the neighborhood are rather intergenerational (Chetty, Hendren and Katz 2016), while most of the within-generation effects pass though mechanisms affecting individual health and other life dimensions (Ludwig, Duncan, Gennetian, Katz, Kessler, Kling and Sanbonmatsu 2013).

Many contributions have suggested using information on the presence of ghettos and enclaves of poor individuals in the city to identify the extent at which poverty in the neighborhood can affect residents' outcomes. These are well captured by measures of the degree of concentrated poverty in the neighborhood. Following the American Census definition, concentrated poverty can be measured as the fraction of poor people in the city that live in neighborhoods where more than 40% of residents is poor. Implementing this measure requires knowing an absolute poverty line (possibly household type- and city-specific, provided by the Census) and counting individuals living in families with income smaller than the reference poverty line. Information on poverty incidence are then aggregated at the census tract level, a good approximation of neighborhood size. Recent studies (Kneebone 2016) have highlighted persistence and trends of concentrated poverty across largest American cities.

We argue here that traditional measures of concentrated poverty are not consistent with very simple concepts illustrating what urban poverty is, implying that ranking cities on the based on concentrated poverty, or studying its trends across time, or assessing correlations with other dimensions, might be flawed. We introduce a simple axiomatic model that generalize concentrated poverty measures toward what we call a *urban poverty index*. We require this index to satisfy some relevant properties, adopted from the study of inequality and poverty, as well as a basic principle: there is a double welfare burden of poverty concentration, meaning that local residents exposed to more local poverty suffer, ceteris paribus, a welfare drop. The largest and generalized the welfare drop, produce by externalities associated with local concentration of poverty in the neighborhood, the larger is urban poverty.

This simple setting allows us to characterize, along with technical properties a *unique* measure of urban poverty. This takes the form of the Gini index of the distribution of poverty shares across the city neighborhoods. The more unequally distributed are poverty proportions across the city neighborhoods with respect to the citywide distribution, the larger is urban poverty. For two cities displaying same poverty incidence (share of poor people over the total), the city with dispersion of poverty shares across neighborhood display larger urban poverty. This inequality captures a form of segregation of poor people across the city neighborhoods: when urban poverty is high, then there are neighborhoods displaying very high shares of local residents that are poor, and neighborhoods that are virtually poverty-free.

Our index, which is always consistent with the underlying axiomatic model, can be additively decomposed along different dimensions, notably space and time. In this way, we can assert whether urban poverty is mostly generated by neighborhoods that are spatially related, unveiling local poverty traps and persistence, from the case where urban poverty is idiosyncratic to the neighborhoods, i.e. the are very specific local characteristics of the neighborhood that drive poverty concentration.

We use our measurement apparatus to assess the dynamic of poverty across all American metro areas over the last 35 years exploiting rich data from the Census and the American Community Survey (AS). We use Census data for 1980, 1990 and 2000 to measure poverty concentration across census tract of American Metropolitan Statistical Areas (MSA) at census year. We also exploit the 5-years estimates from the American Community survey waves 2006-2010, 2010-2014 and 2012-2016 to obtain estimates of concentrated poverty and census tract level for each American MSA for representative years 2008, 2012 and 2014, roughly corresponding to the onset, the striking and the early aftermath of the Great Recession. Census and ACS data come in the form of tables: for each census tract we gather information on proportion of households whose income falls above a certain poverty threshold (which depends on household composition, we consider different thresholds) as well as other characteristics of the residents and housing stock in the census tract (including income distribution, and information about rents, housing values and density). We use these data in three separate ways.

First, we elicit facts about urban poverty. We use the year-specific census tracts partition to capture trends in urban poverty in each MSA spanning 1980, 1990, 2000, 2008, 2012, 2014. We find that XXX

Second, we exploit cross-walks files to construct a partition of the urban space into census tracts that are stable in size across time (the US Census revises census tracts definition to account for population chnages). We can construct a panel of census tracts and assess urban poverty across MSA. Given the panel structure, we can extrapolate different components related to changes across time in the urban poverty index. We study separately the trends and patterns of these changes across MSA. Results show that XXX

Finally, we investigate the drivers of urban poverty and of its components. We use reduced form regressions to assess the contribution of socio-economic composition, housing market functioning, local public policies, supply of educational goods and aspects of the urban income distribution affect urban poverty in American MSA. We place particular emphasis on the role of *gentrification*, which is widely recognized as the major driver of sorting patterns of poor population across neighborhoods in large metro areas. We identify gentrifying neighborhoods in each city by looking at decennial changes in demographic structure and housing market conditions in those neighborhoods.

TO BE CONCLUDED.

2 Measuring urban poverty

2.1 Setting

We assume from the outset that a city can be partitioned into n neighborhoods. For instance, neighborhoods can coincide with an administrative division of the territory, such as the census tracts partition of American cities provided by the Census Bureau. We assume that the spatial organization of neighborhoods in a city is given, and we study how poor people are distributed therein.

Let $i \in \{1, \ldots, n\}$ indicate a neighborhood. There are N_i individuals residing in neighborhood i and $N = \sum_{i=1}^{n} N_i$ individuals in the city. An individual is poor when living in a household whose total disposable income is smaller than the federal monetary poverty line provided by the American Census Bureau, calculated in a given year for that specific type of family. The poverty status of each individual is hence defined by an exogenous poverty line, which adjusts across time but not across cities. The analysis of urban poverty is hence conditional on the definition of poverty status. Let use P_i to denote the number of individuals that are poor and live in neighborhood i, while $P = \sum_{i=1}^{n} P_i$ denotes the total number of poor in the city. For a given city, the *urban poverty configuration* (denoted, for instance, $\mathcal{A}$, $\mathcal{B}$,...) consists in a collection of counts of poor and non-poor residents across the city neighborhoods, i.e. $\mathcal{A} = \{P_i^{\mathcal{A}}, N_i^{\mathcal{A}}\}_{i=1}^{n}$. In what follows, the superscript always indicates a urban poverty configuration, and we explicitly use it only when disambiguation is needed.

The ratio $\frac{P_i}{N_i}$ indicates the share of the population of a given city that is poor and that lives in neighborhood i. The proportion $\frac{P}{N}$ measures instead the incidence of poverty in the city and is an average of the share of poverty measured in each neighborhood of the city ($\frac{P}{N} = \sum_{i=1}^{n} \frac{N_i}{N} \frac{P_i}{N_i}$). The number $\frac{P}{N}$ defines an interesting cutoff point, discriminating between neighborhoods where the poor are over-represented, and neighborhoods where the poor are under-represented compared to the relative incidence of poverty in the city. In a slightly more general setting, we use $\zeta \in [0, 1]$ as a *urban poverty line*, i.e. a cutoff point identifying those neighborhoods where poverty is over-concentrated. If $\frac{P_i}{N_i} \geq \zeta$, then i is addressed to as a high concentrated poverty neighborhood, since the proportion of residents in that neighborhood that are also poor is larger than the threshold ζ. The latter defines a normative judgement about tolerance to poverty concentration.[1]

For given urban poverty line ζ, neighborhoods can be ranked according to the incidence of poverty therein:

$$\frac{P_1}{N_1} \geq \frac{P_2}{N_2} \geq \ldots \geq \frac{P_z}{N_z} \geq \zeta \geq \ldots \frac{P_n}{N_n}.$$

[1] The case $\zeta \approx 0$ reflects very low tolerance towards concentration of poverty, implying that even the existence of poverty is detrimental in the perspective of the evaluator, let aside the disproportional burden of poverty generated in those neighborhoods where poverty is slightly concentrated. Conversely, the case $\zeta \approx 1$ expresses hight tolerance to poverty, since a neighborhood is targeted as poor if and only if all his residents are poor.

For simplicity, the labels $1, 2, \ldots, n$ are assumed to coincide with the rank of the neighborhoods, ordered by decreasing magnitude of concentrated poverty. Among all neighborhoods in the metro area, we identify with z the neighborhood where poverty concentration coincides (or is approximately as large as) the urban poverty line. This neighborhood z will serve as a benchmark, so that poor are over-represented in neighborhood i if and only if $i \leq z$.

2.2 A relative urban poverty line

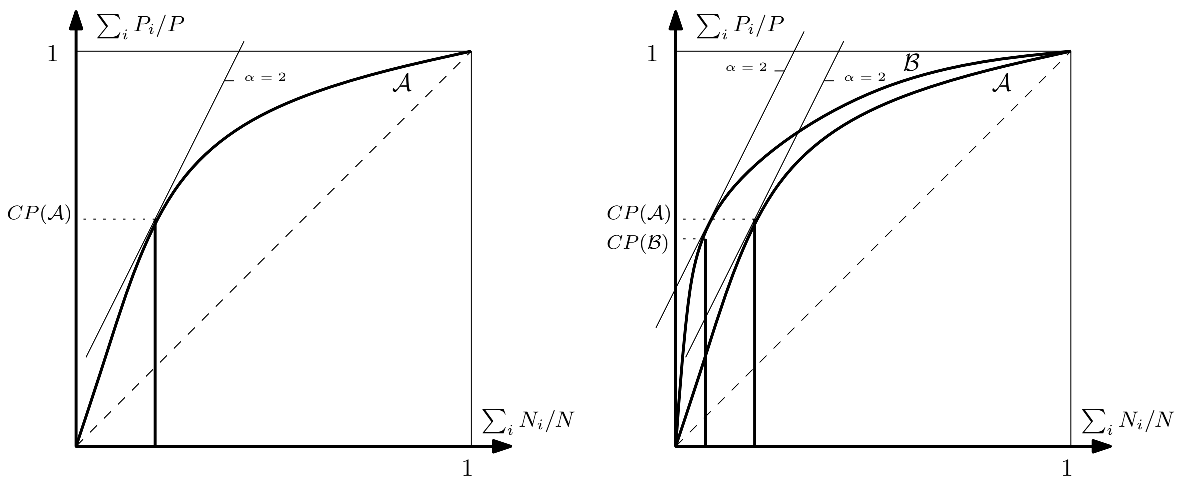
The cutoff poverty line ζ defines an absolute concept of poverty, i.e., neighborhoods are target as poor if the proportion of poor resident is larger than a pre-selected threshold ζ. Adopting an absolute urban poverty threshold implies that configurations displaying different poverty incidence should be evaluated with respect to the same threshold ζ.

Poverty literature (Cowell 2000, Marx, Nolan and Olivera 2015) strongly advocates for relative concepts of poverty lines. We endorse this view in the analysis of urban poverty. The urban poverty cutoff ζ is assumed to be proportional to the citywide average poverty, $\frac{P}{N}$, by a positive real coefficient α, so that

$$\zeta = \alpha \frac{P}{N}. \tag{1}$$

Urban poverty hence depends on the citywide incidence of poverty, as measured by $\frac{P}{N}$, that defines a benchmark at the city level and that might differ across cities. The coefficient α expresses 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 that display extreme poverty, hence magnifying the roles of urban ghettos and enclaves. For instance, poverty incidence among the poorest American cities is approximatively 20%. By setting $\alpha = 2$, one would pick up urban poverty originating from those neighborhoods where more than 40% of the residents are poor. Conversely,

Figure 1: Urban poverty curve and concentrated poverty



(a) Concentrated poverty

(b) A measurement issue

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

small values of α put the emphasis also on the distribution of poverty across the neighborhoods. Comparisons of urban poverty across cities are conditional on the relative poverty threshold α, which is kept constant across cities.

2.3 Concentrated poverty and its critical aspects

Given the incidence of poverty in the city, $\frac{P}{N}$, there are neighborhoods displaying larger incidence of poverty ($\frac{P_i}{N_i} > \frac{P}{N}$) and neighborhoods displaying smaller incidence of poverty than the city as a whole. Urban poverty assessments focus on extent and distribution of poverty in those neighborhoods where poverty incidence is larger.[2]

A convenient way to represent the distribution of the poor population in the city is to plot the cumulative proportion of poor people against the proportion of the overall population living in the neighborhoods displaying higher incidence of poverty. These neighborhoods are ordered according to the ratio $\frac{P_i}{N_i}$ and cumulative proportions are calculated based on this order. The cumulative proportion of poor people in neighborhood j is

[2] Concentrated poverty evaluations are not directly concerned in comparisons of poverty (i.e., $P^{\mathcal{A}} > P^{\mathcal{B}}$) or poverty incidence (i.e., $\frac{P^{\mathcal{A}}}{N^{\mathcal{A}}} > \frac{P^{\mathcal{B}}}{N^{\mathcal{B}}}$) across cities.

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 the points with coordinates $\left(\sum_{i=1}^{j} \frac{N_i}{N}, \sum_{i=1}^{j} \frac{P_i}{P}\right)$ with $j = 1, \ldots, n$ on a graph. The curve starting from the origin and interpolating these points is the *urban poverty curve*. The urban poverty curve of an hypothetical configuration $\mathcal{A}$ is reported in panel (a) of Figure 1. 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. Its graph is concave and always lies above the unit square diagonal, implying that in configuration $\mathcal{A}$ there are neighborhoods with $\frac{P_i}{N_i} < \frac{P}{N}$ and other neighborhoods with $\frac{P_i}{N_i} > \frac{P}{N}$.[3] We will use this curve extensively throughout the paper.

We can relate this curve to the measurement of urban poverty in a city. Literature has focused on a particular aspect of urban poverty, denoted *concentrated poverty*, which is measured by the incidence of poverty in those neighborhoods of the city where poverty is more concentrated *vis-à-vis* a urban poverty line ζ. According to the American census, concentrated poverty corresponds to the proportion of poor residents that live in census tracts where at least 40% of inhabitants fall below the poverty line (i.e., $\zeta = 0.4$). A convenient measure of concentrated poverty is $CP := \sum_{i=1}^{z} \frac{P_i}{P}$, where $\frac{P_z}{N_z} \approx \zeta$. The concentrated poverty measure CP was first proposed by Wilson (1987) to pick up spatial trends in urban poverty and thus highlight distressed census tracts characterized by extreme poverty. **?**, **?** and **?** have documented the dynamics of concentrated poverty (using CP index) in larger American cities, warning about the re-concentration patters registered in the last decades.

The index CP is related to the urban poverty curve of a given city: it is, in fact, the level of the curve corresponding to the proportion of the city population living in neighborhoods with at least 40% of inhabitants falling below the federal poverty line, that

[3]This 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 $\frac{N_i}{N}$. 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, would coincide with the unit square diagonal.

is $\sum_{i=1}^{z} \frac{N_i}{N}$. The index CP is calculated on the basis of an absolute urban poverty line. A relative version of the index, denoted $CP(\mathcal{A}; \alpha)$, can be also constructed. In this case, the urban poverty line is $\alpha \frac{P^{\mathcal{A}}}{N^{\mathcal{A}}}$ for a city with a configuration of urban poverty $\mathcal{A}$, and it changes across configurations depending on poverty incidence in the city. Consider now a city where $\frac{P^{\mathcal{A}}}{N^{\mathcal{A}}} = 0.2$ and $\alpha = 2$, which gives $\zeta = 0.4$ from (1) (implying that concentrated poverty calculations based on an absolute or a relative urban poverty thresholds coincide). The coefficient α identifies the slope of a line tangent to the urban poverty curve, as in Figure 1. The tangency point identifies the neighborhood z displaying a proportion of inhabitants falling below the federal poverty line that is approximatively $\alpha \frac{P^{\mathcal{A}}}{N^{\mathcal{A}}}$, the relative urban poverty threshold.

The concentrated poverty index might miss important aspects of the distribution of poverty across the city neighborhoods, implying that the ranking of the cities it produces might not be consistent with the changes in the geography of poverty registered by non-intersecting urban poverty curves. The example illustrated in panel b) of Figure 1 makes a case. We consider two cities, denoted by configurations $\mathcal{A}$ and $\mathcal{B}$ where, for simplicity, $\frac{P^{\mathcal{B}}}{N^{\mathcal{B}}} = \frac{P^{\mathcal{A}}}{N^{\mathcal{A}}}$. The distribution of poverty across the neighborhoods of city $\mathcal{B}$ is more uneven than in city $\mathcal{A}$, implying that the urban poverty curve of the former lies always above that of the latter. It seems uncontroversial to conclude that urban poverty is larger in $\mathcal{B}$ than it is in $\mathcal{A}$: for any share of the city population living in neighborhoods where poverty is more concentrated, the proportion of poor is always larger in $\mathcal{B}$ than in $\mathcal{A}$. Nonetheless, $CP(\mathcal{B}, \alpha) < CP(\mathcal{A}, \alpha)$ as shown in the figure for $\alpha = 2$.

The example above formalizes intuitions in **?**, who suggest valuing the intensity *and* the distribution of poverty in the city through the dissimilarity and the interaction indices, eliciting the degree of segregation of poor people across the neighborhoods. These indices are interesting, because they explicitly highlight that urban poverty is connected to the uneven distribution of poverty in the urban space, but have no clear connections with the welfare burden induced by concentrated poverty on the people exposed to it.

An interesting alternative approach consists in valuing the unequal distribution of poverty across the neighborhoods where poverty is more concentrated is the *Gini coefficient* $G(.;\alpha)$ of the vector of poverty proportions $\frac{P_1}{N_1}, \dots, \frac{P_z}{N_z}$, where neighborhood z is such that $\frac{P_z}{N_z} \approx \alpha \frac{P}{N}$. For a given configuration, the index is defined as follows:

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

The index $G(.;\alpha)$ 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.

The index is yet based on limited information about poverty distribution in the city and might rank configurations inconsistently with non-intersecting urban poverty curves. In what follows, we provide an axiomatic model for urban poverty that explicitly incorporates normative judgements about the welfare implications of concentrated poverty. We show that the *unique* index of concentrated poverty consistent with the setting is the index $G(\mathcal{A}) := G(\mathcal{A}; 0)$, which is a measure of the area comprised between the urban poverty curve and its reference diagonal.

2.4 Characterization of a family of urban poverty measures

A urban poverty index is a function $UP : \mathcal{P} \to \mathbb{R}_+$ (with $\mathcal{P}$ the set of urban configurations) assigning to each configuration a number, interpreted as the level of urban poverty in that configuration. We write $UP(\mathcal{A};\alpha)$ to explicitly mention that evaluations of urban poverty are conditional on a relative urban poverty line. Every urban poverty measure should obey a simple monotonicity principle: if the proportion of poor people living in neighborhoods where poverty concentration is high increases, then the urban poverty measure UP should not decrease.

Axiom *A1* (Monotonicity) *Other things being equal, an increase of the proportion of poor people in a neighborhood where poverty is concentrated ($i \leq z$) cannot reduce urban poverty.*

The $CP(.;\alpha)$ index does not satisfy this basic axiom, as motivated in previous examples. A convenient way to incorporate the implications of this axiom on urban poverty measurement is to focus on urban poverty indices that explicitly depend on the *urban poverty shortfall* $\frac{P_i/N_i}{P_z/N_z} - 1$, with z being the neighborhood identified by the urban poverty threshold α. The shortfall is positive in those neighborhoods where poverty is mostly concentrated, and increases if the proportion of the poor $\frac{P_i}{N_i}$ grows in some of the neighborhoods with $i \leq z$. The next axiom emphasizes that urban poverty indices should be written as normalized (weighted) averages of urban poverty shortfalls.

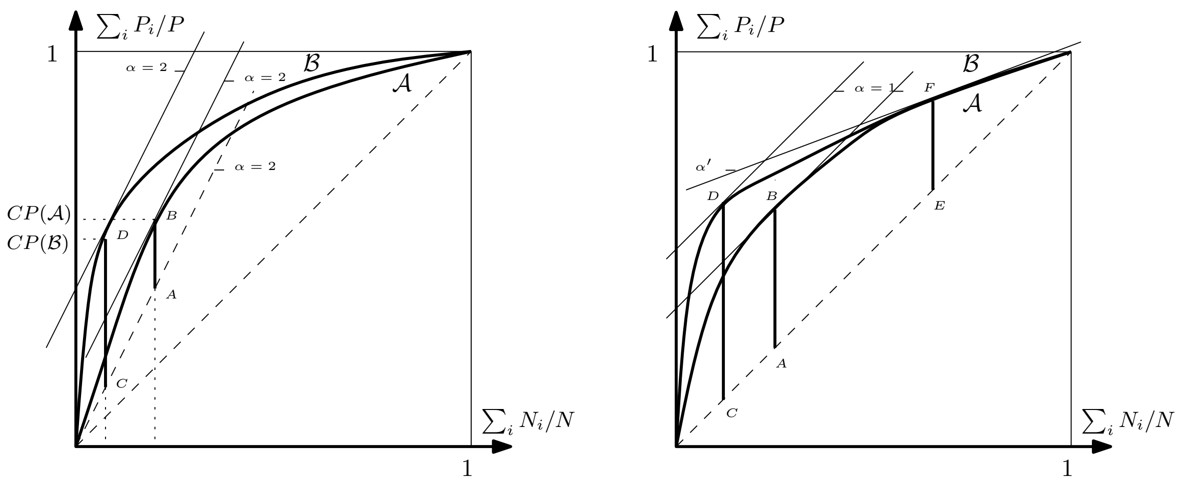
Axiom *A2* (Urban Poverty) *The urban poverty index for configuration $\mathcal{A}$ at relative urban poverty threshold α is:*

$$UP(\mathcal{A};\alpha) \quad := \quad A(\mathcal{A},\alpha) \sum_{i=1}^{z} \frac{N_i}{N} \left(\frac{P_i/N_i}{P_z/N_z} - 1 \right) \; w_i(\mathcal{A},\alpha), \tag{2}$$

with $A(\mathcal{A},\alpha)$ a normalization factor and $w_i(\mathcal{A},\alpha)$ are normative weights attached to the neighborhoods (and distinct from the population weights $\frac{N_i}{N}$).

Different urban poverty indicators obtain for specific choices of the normalization and weighting parameters. Let consider the case in which $A(\mathcal{A},\alpha) = \alpha$ and $w_i(\mathcal{A},\alpha) = 1$ for every neighborhood i, implying that the measure expresses exclusively concerns for the incidence of concentrated poverty, but not for the distribution of poor individuals across

Figure 2: Urban poverty curve and corrected concentrated poverty



(a) Corrected concentrated poverty

(b) A measurement issue

Note: The corrected concentrated poverty index CP^* corresponds to the vertical 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 $\mathcal{A}$. The corresponding CP^* indices at different poverty thresholds $\alpha = 1$ and $\alpha' < \alpha$ are also provided.

neighborhoods where poverty is more concentrated. Under these conditions we have that

$$\begin{aligned} UP(\mathcal{A};\alpha) &= \alpha \sum_{i=1}^{z} \frac{N_i}{N} \left(\frac{P_i/N_i}{P_z/N_z} - 1 \right) \\ &= CP(\mathcal{A};\alpha) - \alpha \sum_{i=1}^{z} \frac{N_i}{N} =: CP^*(\mathcal{A};\alpha). \end{aligned}$$

The result, which follows from (1), shows that the index $CP(.;\alpha)$ can be made consistent with Axioms A1 and A2 only if corrected by a term $\alpha \sum_{i=1}^{z} \frac{N_i}{N}$, measuring the expected degree of concentrated poverty among the z neighborhoods, under the assumption that the poor population is evenly spread out across the city neighborhoods. 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 corrected concentrated poverty indices $CP^*(\mathcal{A},\alpha)$ (segment AB) and $CP^*(\mathcal{B},\alpha)$ (segment CD).[4] Consistently with the ordering of configurations induces by

[4]These indices visually correspond to the distance computed at abscissa $\frac{N_z}{N}$ between the urban poverty curve and the line with slope α intersecting the origin.

the urban poverty curves in the graphs, the corrected concentrated poverty index ranks $CP^*(\mathcal{B};\alpha) > CP^*(\mathcal{A};\alpha)$. Since every urban poverty curve is concave and lies above the diagonal, the index $CP^*(.;\alpha)$ is always positive and bounded above by $CP(.;\alpha)$.

The corrected concentration index $CP^*(.;\alpha)$ might be regarded to as a natural reference measure for urban poverty assessments. It combines three aspects of poverty: A normative view about the identification of concentrated poverty (α), which reflects a policy target; the incidence of the burden of concentrated poverty across the population (denoted by the index H, the proportion of individuals residing in high-poverty neighborhoods); the intensity of poverty in the neighborhoods where poverty is concentrated (denoted by I, the neighborhood poverty gap). The index can be hence decomposed as follows:

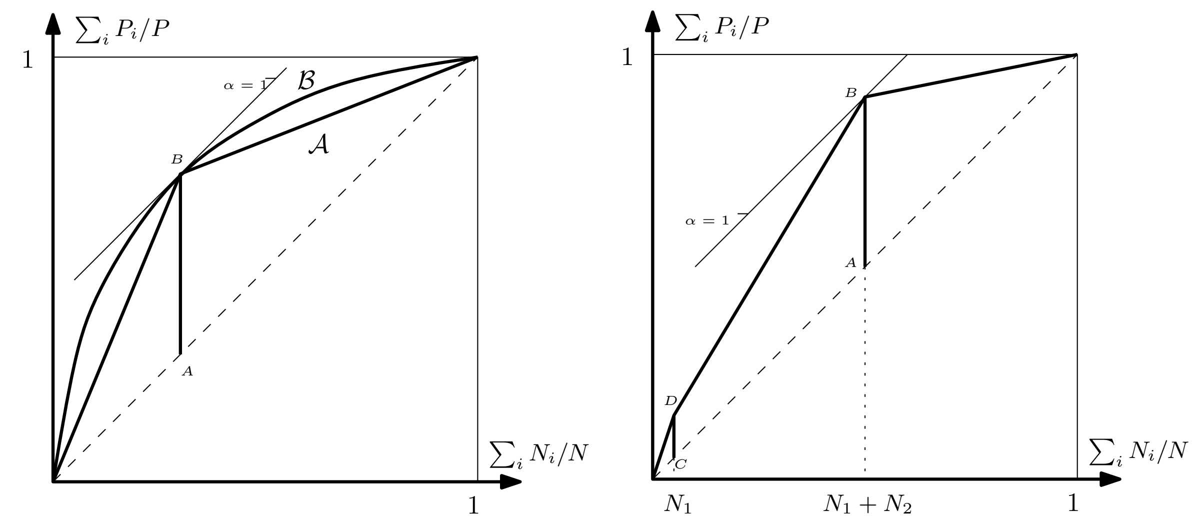
$$\begin{aligned} CP^*(\mathcal{A};\alpha) &= \alpha \left(\sum_{i=1}^{z} \frac{N_i}{N} \right) \sum_{i=1}^{z} \frac{N_i/N}{\sum_{i=1}^{z} N_i/N} \left(\frac{P_i/N_i}{P_z/N_z} - 1 \right) \\ &= \alpha\ H\ I. \end{aligned}$$

The index $CP^*(.;\alpha)$ is consistent with the ranking of configurations induces by non-intersecting urban poverty curves, but it is far from being an ideal measure of urban poverty, for at least two reasons. First, the index measures the degree of concentration of poverty by focusing on a particular coordinate of the urban poverty curve. Hence, the the index might not distinguish between two situations even if they are unambiguously ranked by the urban poverty curves. Panel (b) in Figure 2 reports one of these cases.[5]

The second critical aspect of $CP^*(.;\alpha)$ is that the index does not value heterogeneity in the concentration of poor individuals across the city's neighborhoods. There are two potential sources of heterogeneity. First, heterogeneity in $\frac{P_i}{N_i}$ ratios for any $i \leq z$. When these ratios are homogenous across neighborhoods where poverty is concentrated, i.e.,

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

Figure 3: Corrected concentrated poverty and neighborhood structure heterogeneity



(a) In neighborhood composition

(b) In neighborhood size

Note: Corrected concentrated poverty measures at poverty thresholds $\alpha = 1$ are given by solid line segments AB in both graphs.

$\frac{P_1}{N_1} = \ldots = \frac{P_z}{N_z} \leq \alpha \frac{P}{N}$, the $CP^*(.;\alpha)$ index is a sufficient statistic for urban poverty. If they are not, the index $CP^*(.;\alpha)$ might rank very dissimilar configurations as equivalent in terms of urban poverty. The 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 $CP^*(\mathcal{B};1) = CP^*(\mathcal{A};1)$.[6]

Another source of heterogeneity is in the demographic size of the neighborhoods, $\frac{N_i}{N}$. Evaluations based on the $CP^*(.;\alpha)$ index might not be robust to small changes in the relative poverty threshold α, implying that heterogeneity in neighborhoods composition drives urban poverty assessments. Panel (b) of Figure 3 reports a problematic case where poverty is concentrated in two neighborhoods ($z = 2$) when $\alpha = 1$, with $\frac{P_1}{N_1} \approx \frac{P_2}{N_2}$, but N_1 is substantially smaller than N_2.[7]

[6] In the case of configuration $\mathcal{A}$, the poor are evenly distributed in neighborhoods where poverty is concentrated (so that the urban poverty curve is piecewise linear). In configuration $\mathcal{B}$, instead, there is heterogeneity in the distribution of poverty across neighborhoods, implying dominance in urban poverty curves.

[7] In this case, urban poverty of configuration $\mathcal{A}$ is $CP^*(\mathcal{A}, 1)$, given by segment AB in the figure. If the poverty threshold is sized down to $\frac{P_2/N_2}{P/N}$, urban poverty does not change. If the poverty threshold is

We retain the corrected concentrated poverty as a proper benchmark for measuring urban poverty only for those configurations where neighborhoods where poverty is concentrated have homogeneous size and poverty is evenly distributed therein. This is formalized with a normalization axiom for the urban poverty index.

Axiom *A3 (Normalization)* *For any configuration $\mathcal{A}$ where $\frac{P_i}{N_i} = \frac{P^*}{N^*}$ and $\frac{N_i}{N} = \frac{N^*}{N}$ for all neighborhoods $i \leq z$ and P^* and N^* are constant, urban poverty can be normalized to $UP(\mathcal{A};\alpha) = CP^*(\mathcal{A};\alpha) = \alpha HI$.*

Our last axiom highlight the welfare consequences of urban poverty. There is increasing empirical evidence that the place, and even the neighborhood, experienced during youth and adulthood has strong implications for many social, economic and health outcomes, and definitely for individual welfare. The degree of concentration of poverty in the neighborhood is an important source of externalities that have potential negative implications for welfare of the individuals that are exposed to it, implying a double welfare burden of urban poverty. We assume that, other things being equal, the welfare of an individual living in a neighborhood i and exposed to a high proportion of poor residents is smaller than the welfare the same individual would achieve in any other neighborhood j with a smaller fraction of residents in poverty. Let denote by $W(., \frac{P_i}{N_i})$ the welfare of this individual. It depends, among other things, on percentage of poor individuals in his neighborhood. The next axiom conveys the idea that concentration of poverty in the neighborhood produce negative externalities on individual welfare.[8]

Axiom *A4 (Double burden of poverty on welfare)* *If $\frac{P_i}{N_i} \geq \frac{P_j}{N_j}$ then $W(., \frac{P_i}{N_i}) \leq$*

further reduced to any level α' just smaller than $\frac{P_2/N_2}{P/N}$, urban poverty jumps to the new level identified by the line segment CD in the figure, which is definitely smaller than $CP^*(\mathcal{A}, 1)$. Hence, the corrected concentrated poverty index might display substantial discontinuities with respect to the poverty threshold definition.

[8]The unequal distribution of concentrated poverty across the city neighborhoods implies, in particular, an unequal exposure to the *double burden of poverty* for the poor individuals living therein: These people already have low welfare because poor and, additionally, they are subject to the negative external effects associated to living in a place where poverty is concentrated.

$W(., \frac{P_j}{N_j})$ for any admissible individual welfare function W.

A natural way to relate the measurement of urban poverty to Axiom A4, and to introduce the idea that poor individuals living in places where poverty is concentrated are exposed to a double welfare burden, is to assume that neighborhoods where poverty is more concentrated also receive the largest weights in urban poverty assessments. There are many weighting functions $w(.;\alpha)$ in (2) that are consistent with this view. We focus on those weights that depend exclusively on positional information given by the ranking that each individual occupies in the distribution of people across neighborhoods, ordered by the degree of poverty concentration they are exposed to in their neighborhood.

Axiom *A5* *(Rank weights)* *The weight $w_i(.,\alpha)$ associated to neighborhood i is given by the ranking in the distribution of welfare occupied by the individuals living in i.*

Since individual welfare is assumed monotonic in the proportion of poor in the neighborhood and that all individuals in the same neighborhood share the same proportion of concentrated poor, their rank is constant within the neighborhood. In any neighborhood $i \leq z$ there are N_i individuals, each weighted $1/N$. They are all sharing the same position in the welfare ranking. Consistently with axioms A4 and A5, we can hence express the weight of neighborhood i as follows:

$$w(.,\alpha) \quad = \quad \sum_{j=1}^{z} \frac{N_j}{N} - \sum_{j=1}^{i} \frac{N_j}{N} + \frac{N_i}{N} \tag{3}$$

There is only one urban poverty index that is consistent with axioms A1 and A2, that converges to $CP^*(.;\alpha)$ in specific cases and that accounts for heterogeneity in the distribution of concentrated poverty in a way that is consistent with the implications of concentrated poverty on individual welfare. The functional form characterized in the next lemma shows that any urban poverty index must depend exclusively on the relative poverty threshold and the data.

Lemma 1 *For any configuration $\mathcal{A}$ with a large number of neighborhoods, the unique urban poverty index that satisfies axioms A1-A5 is given by:*

$$UP(\alpha, \alpha) \;=\; \frac{\alpha z}{(z+1)} H \left[I \;+\; (I+1)G(\mathcal{A};\alpha) \;-1+\; \frac{2}{H^2}\sum_{i=1}^{z}\frac{N_i}{N}\sum_{j=1}^{i-1}\frac{N_j}{N} \right], \qquad (4)$$

where $G(\mathcal{A};\alpha)$ is related to the Gini index of the distribution of poor people proportions across the neighborhood where poverty is concentrated.

Proof. See supplemental appendix. ■

The urban poverty index reflects the implication of three aspects of the distribution of poverty across the city neighborhoods: the incidence, H, the intensity, I, and the degree of inequality in the distribution of poor people in those neighborhoods that display higher levels of concentrated poverty, $G(.;\alpha)$. Which aspect of urban poverty prevails depends on the full distribution of poverty across the neighborhoods where poverty is more concentrated.

The index $UP(.;\alpha)$ characterized in Lemma 1, however, has only an ordinal interpretation since its scale depends on the chosen relative urban poverty threshold and on the number and size of neighborhoods. Furthermore, the urban poverty index does not evaluate the extent of the distribution of poor in neighborhoods where the incidence of poverty is smaller than that implied by the urban poverty line.

Next, we propose axioms that overcome these limitations and we show that there is only one index consistent with the representation of urban poverty given in Lemma 1 and with these new axioms.

2.5 Main result: A unique urban poverty index

In empirical analysis of urban poverty it is desirable to use indices that have the form of (4) and that satisfy a minimum degree of cardinal comparability across configurations that differ in the number of neighborhoods. Comparability is achieved by scaling the $UP(.;\alpha)$ index characterized in Lemma 1 by a factor that depends upon the poverty threshold definition and the data, so that neighborhood z can be identified.

Axiom *A6* *(Cardinality)* *Urban poverty evaluations should not be affected by the number of neighborhoods. The urban poverty index $UP(.;\alpha)$ should be hence scaled by the factor $\frac{z+1}{z\alpha}$.*

Another concern for empirical urban poverty analysis rests on the implications of the size of the neighborhoods on poverty evaluations. We investigate the possibility of reshaping the size and number of neighborhoods by a particular operation, denoted the *neighborhood splitting.* An operation of neighborhood splitting applied to neighborhood i implies splitting the neighborhood territory into two new neighborhoods i' and i'' of smaller size, such that $\frac{P_i}{N_i} = \frac{P_{i'}}{N_{i'}} = \frac{P_{i''}}{N_{i''}}$ and $N_i = N_{i'} + N_{i''}$. Any sequence of splits of neighborhoods increases the number of neighborhoods while reshaping their size, without affecting the distribution of poverty across these neighborhoods. We postulate that this operation is a source of invariance for every urban poverty indicator.[9]

Axiom *A7* *(Invariance to neighborhood splitting)* *The $UP(.;\alpha)$ index is invariant to any sequence of neighborhood splitting operations.*

Lastly, we formalize the idea that urban poverty evaluations should be concerned with the distribution of poor people across the whole city, rather than being focused on the subset of neighborhoods of the city where poverty is more concentrated. By doing so, we explicitly account for the fact that urban concentration of poverty gives rise to double burden of poverty in those neighborhoods where the poor are over-represented, and a double welfare benefit for the residents living in neighborhoods where the poor are under-

[9]This axiom is related to the replication invariance property often adopted in inequality analysis (Atkinson 1970, Cowell 2000). It is also a basic property of segregation indices (Hutchens 1991, Frankel and Volij 2011, Andreoli and Zoli 2014). Furthermore, the urban poverty curve is preserved by effect of any sequence of neighborhood splitting operation

represented compared to the citywide average. We take a normative stance on this by requiring that $z = n$, a result which can be achieved by setting $\zeta = 0$.

Axiom *A8 (Focus on citywide urban poverty)* $\alpha \to 0_+$.

Theorem 1 *The urban poverty index* $U(.;\alpha)$ *satisfies Axioms A1-A8 if and only if it is the Gini index* $G(.)$.

Proof. See supplemental appendix. ■

Theorem 1 brings forward four contributions to the measurement of concentrated poverty. First, it shows that the simple, normatively appealing axiomatic model A1-A8 developed in the section characterizes exactly one measure of urban poverty, which does not depend on a urban poverty line (i.e., $UP(\mathcal{A}, 0) := UP(\mathcal{A})$), and which takes the specific functional form of the Gini inequality coefficient of the distribution of poverty shares across the city neighborhoods (i.e., $UP(\mathcal{A}) = G(\mathcal{A})$).

Second, the theorem highlights that urban poverty arises when the proportion of poor people in each neighborhood, $\frac{P_i}{N_i}$, is dissimilar from proportion of poor people in the city, $\frac{P}{N}$. Coherently with the intuitions in **?**, urban poverty ca be also interpreted as a form of segregation of poverty across the neighborhoods of a city.

Third, urban poverty evaluations that account for the welfare burden generated by exposure to concentrated poverty in the neighborhood should account for the distribution of poverty throughout the city. As a consequence, urban poverty comparisons across cities can be performed irrespectively of the choice of the underlying urban poverty line. Comparisons based on the $UP(.)$ index are always consistent with the ranking of configurations produced by non-intersecting urban poverty curves.

Fourth, the urban poverty index $UP(.)$ can be conveniently decomposed to keep track of changes in urban poverty that take into account the longitudinal dimension of the data. This aspect is relevant for the American case, where poverty concentration within the same census tract can be followed through time and its contribution to urban poverty at the level of the city can be then isolated. The next section investigates a decomposition of $UP(.)$ that is relevant for investigating trends of urban poverty in American cities.

3 Addressing changes in urban poverty

3.1 Decomposing changes in urban poverty

We focus now on changes in urban poverty between two period t and $t' > t$ within the same metro area. We address more specifically the American case, and we use the partition in census tracts provided by the Census Bureau to denote the city's neighborhoods. Urban poverty in an American city in period t is given by configuration $\mathcal{A}$ and in t' by $\mathcal{A}'$. We are interested in the difference

$$\Delta UP = UP(\mathcal{A}') - UP(\mathcal{A}) = G(\mathcal{A}') - G(\mathcal{A}).$$

In the same city, census tracts are assumed to be held fixed across time.[10] We hence observe P_i and N_i in any neighborhood i both in time t and t'. We exploit the longitudinal component of our data to decompose changes in poverty into three components.

The first component of changes in urban poverty captures the dynamic effect of changes in the *demographic weights* of the census tracts on urban poverty, and is denoted denoted 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 at least a census tract i. The change in the demographic weight of the census tract has non-trivial effects on urban poverty changes. If the demographic weight increases in those tracts that are more dissimilar in terms of poverty composition (i.e., there is a demographic expansion in tracts with relatively few or relatively many poor residents), then the contribution to urban poverty is positive ($W > 0$). Conversely, if the demographic growth is predominantly

[10] See crosswalk files XXXX.

concentrated in those tracts displaying a more proportionate distribution of the poor in relative terms (i.e., where $\frac{P_i}{N_i} \approx \frac{P}{N}$), then urban poverty decreases ($W < 0$). The element W captures the interplay between growth in proportions of poverty and change in absolute poverty (P_i). It allows to factor out the effect of population change from changes related to the distribution of poverty across the city's census tracts.

The second component of changes in urban poverty isolates *changes in incidence of poverty* in the city and is denoted C. It is a function of the growth rate c of the incidence of poverty in the city, defined as:

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

The element C measures the implication of a citywide increase in poverty incidence on urban poverty. This component allows to separate the component related to proportional growth in concentrated poverty across all neighborhoods (which coincides with the growth rate of citywide poverty P/N, i.e., $\frac{P_i^{\mathcal{A}'}}{N_i^{\mathcal{A}'}} = (1+c)\frac{P_i^{\mathcal{A}}}{N_i^{\mathcal{A}}}$ for every i) from the neighborhood-specific growth rates of poverty (that are heterogeneously distributed across the city's neighborhoods). By factoring out C, we can isolate the component of urban poverty change that is related to changes of poverty incidence in the city from other components that are related to changes in the distribution of poverty across census tracts.

The last component we consider captures the implications of disproportionate changes in tracts' poverty rates on changes in urban poverty. If tracts poverty rates *diverge*, we expect to observe larger panel growth rates of poverty in those census tracts where poverty is already concentrated in period t. This leads to increasing urban poverty. Poverty rates instead *converge* across census tracts if poverty rates grow faster in those tracts where poverty were least concentrated in t. The implications of convergence of concentrated poverty on changes in urban poverty are, nevertheless, ambiguous. If convergence is limited, urban poverty may decrease. This always happens when poverty incidence in each neighborhood is closer to the poverty incidence in the city in t' than it was in t. If, however, converge is strong enough then census tracts where poverty was highly concentrated in t become tracts with low poverty concentration in t' and, viceversa, census tracts displaying low concentrated poverty in t become tracts of high poverty concentration in t'. In this case, we may end up attributing a reduction or lack of changes in urban poverty to a *re-ranking* of the census tracts.

Borrowing the terminology from the literature studying the distributional effects of longitudinal income growth (Jenkins and Van Kerm 2016), we propose to isolate two components of convergence in poverty incidence across census tracts. The first component, denoted R, captures the pure effect of re-ranking of census tracts and is relevant to isolate situations where a given tract i changes position from t to t' in the ranking of tracts, but the overall distribution of concentrated poverty after the re-ranking remains the same. The second component, denoted E, captures instead the extent of divergence/convergence in concentrated poverty by comparing a census tract i in t with tract i' in t', such that i and i' occupy the same position in the tracts' ranking.

3.2 Result and discussion

Our first result is that the changes in concentrated poverty can be linearly decomposed into the four components illustrated above.[11]

Corollary 1 *The change in urban poverty ΔUP from configuration $\mathcal{A}$ in time t to $\mathcal{A}'$ in time t' for a urban poverty index satisfying axioms A1-A8 can be decomposed as follows:*

$$\Delta UP = G(\mathcal{A}') - G(\mathcal{A}) = W \ + \ R \ + \ C \cdot E,$$

[11] The proof of the corollary exploits a representation of the urban poverty index $G(.)$ that involves comparisons of pairs of neighborhoods. Some of these pairs display a change in weight, in their relative ranking (based on the intensity of concentrated poverty) and of poverty rates. These components are then re-organized in a way that can be expressed as the decomposition in the corollary.

where $C = 1/(1+c)$.

Proof. See supplemental appendix. ■

The interesting elements of the decomposition are E and R. The term $R + C \cdot E$ captures the degree of convergence or divergence once changes in population composition have been factored our. The component E measures changes relative to the citywide poverty incidence, it is positive in case of convergence, negative in case of divergence of poverty rates across census tracts. The component R, instead, is always non-negative: this term offsets the implications of strong forms of convergence (implying $E < 0$) that simply boil down to reversal in the ranks of the census tracts where convergence occurs.[12] The component R arises exclusively from census tracts that revert their ranking in the distribution. The component E, instead, arises comparing census tracts that do not change their relative position in the ranking, and by comparing concentrated poverty in those tracts exhibiting a change in ranking in t' with the level of concentrated poverty in tracts occupying the same position in t.

The decomposition in Corollary 1 has advantages for comparing the dynamics of urban poverty across cities when information on urban poverty is limited to the data defining configurations in $\mathcal{P}$.[13] First, by factoring out the effect of demographic changes, W, on urban poverty one can control for the differences in demographic growth across cities.

Second, the components R and $C \cdot E$ pick up specific aspects of changes in poverty concentration that cannot be inferred from the knowledge of ΔUP alone. For instance, consider two cities displaying no decennial changes in urban poverty, whereas $R = C \cdot E = 0$ for the first city, while $R = -C \cdot E > 0$ for the second. While the poor population is

[12]Oversimplifying, a city with two census tracts where all poor people are concentrated in neighborhood 1 in t would display high levels of convergence in urban poverty, if all poor people migrate to neighborhood 2 in period t'. This change, however, can be described as a swap in the name of the tracts, that arguably has no impact on the citywide level of urban poverty. In this case, the components C and R counterbalance each others.

[13]The American Census Summary Tape Files, for instance, only report information on population counts at different geographic scales and one can, at best, study the dynamics of these population counts within census tracts.

largely immobile in the first city, poverty concentration varies substantially in the second city, despite the change does not imply a neat form of convergence in the concentration of poverty, but rather a shift of poverty across the census tracts of the city. Another interesting example could be that in which urban poverty grows in both cities. However, in one city urban poverty might grow because the number of poor households grows more rapidly in places that are historically poor, implying a divergence in poverty concentration across the census tracts of the city. The component R would be small in this case. In the other case, instead, the map of poverty might be substantially re-designed, with traditionally poor neighborhoods experiencing substantial reductions in the share of poor residents, and middle- and lower-class neighborhoods witnessing a growth in concentrated poverty that is even more intense than the average. The component R and $C \cdot E$ would be both large in this case.

The distinctions highlighted by the decomposition are relevant for empirical analysis that aim at assessing and distinguishing the implications of different phenomena on urban poverty. In the example above, for instance, the growth in urban poverty for the first city might be well driven by changes in the income distribution (with the population getting poorer), and increasing concentration can be likely explained by the sorting behavior of the poor, who settle in more affordable tracts of the city where poverty is already concentrated. We expect to observe this pattern in cities that have an history of gentrification, and where sorting can be likely explained by local characteristics of the city, such as access to public goods, amenities, transportation and access to the job market. Conversely, the features of the change in urban poverty displayed by the second city are more coherent with the premises of recent waves of gentrification: poor households in historically poor census tracts are increasingly replaced by middle-class homeowners, and forced to move and concentrate in marginal areas of the city, offering lower cost of living. In these cities, empirical patterns of urban poverty can be best explained by gentrifying potential of the historically poor census tracts.

We use the decomposition described in Corollary 1 and the American Census and American Community Survey data to further characterize changes in urban poverty in American cities, and to describe more in depth the association with the Big Sort hypothesis. Before doing so, in the next section we further decompose changes in urban poverty that are clustered at the geographic level, from those that occur randomly throughout the city. It is important to distinguish empirically these two cases to better understand the implications of the drivers of urban poverty: while clustering of concentrated poverty highlight that features of the place where tracts are located (along with access to jobs, or belonging to the same school district) matters, changes in urban poverty driven by spatially unrelated tracts imply that local characteristics of the housing market might be predominant.

3.3 Spatial components of urban poverty

The decomposition illustrated in Corollary 1 does not take into account the urban geography of the urban poverty changes. The measures ΔUP, W, R and E ignore information about the location and proximity of the census tracts. Exploiting this information, we can further separate elements of ΔUP, W, R and E that are related to a "neighborhood component" and a "non-neighborhood component" of urban poverty concentration (Rey and Smith 2013). The former component single out the sources of changes in concentrated poverty that stem from census tracts that are geographically clustered. The latter component, instead, focuses on the contribution to urban poverty of tracts that are arguably spatially uncorrelated.

We obtain the spatial decomposition from a given proximity matrix $\mathbf{N}$, its element $n_{ij} \in [0, 1]$ indicating the extent of proximity between census tracts i and j according to some underling criterion. The matrix $\mathbf{N}$ can be constructed from the data and is assumed fixed throughout the comparisons, but is specific to the metro area. Spatial dependence of concentrated poverty is accounted for by looking at the spatial proximity of the census tracts. In this case, row i of $\mathbf{N}$ would indicate the probability that any tract j is contiguous to tract i[14]

We now show that the linear decomposition advocated in Corollary 1 is preserved even when changes in urban poverty and its components are further decomposed into changes occurring among spatially close census tracts (denoted with a "N" subscript) and tracts that are likely spatially unrelated (denoted with a "nN" subscript).

Corollary 2 *The change in urban poverty ΔUP from configuration $\mathcal{A}$ in time t to $\mathcal{A}'$ in time t' for a urban poverty index satisfying axioms A1-A8 can be decomposed as follows:*

$$\begin{aligned} \Delta UP = G(\mathcal{A}') - G(\mathcal{A}) &= (G_N(\mathcal{A}') + G_N(\mathcal{A}')) - (G_{nN}(\mathcal{A}) + G_{nN}(\mathcal{A})) \\ &= (W_N + W_{nN}) + (R_N + R_{nN}) + C\,(E_N + E_{nN}). \end{aligned}$$

Proof. See supplemental appendix. ■

The decomposition offered by Corollary 2 are useful to trace out the spatial component of urban poverty changes. First, the corollary shows that the urban poverty index $G(.)$ is linearly decomposed into urban poverty that is generated from census tracts that display a neighboring structure. When G_N is large relative to G, most of the heterogeneity in urban poverty occurs in census tracts that are located nearby in space. Conversely, when G_N is small, neighboring census tracts display similar levels of concentrated poverty, thus

[14] One other alternative is to measure proximity with respect to reference points on the city map. In this case, row i of matrix $\mathbf{W}$ could measure the extent of proximity of i to these points. These reference points could be assigned exploiting exclusively geographical information. Interesting cases are those in which proximity is expressed with respect to the neighborhoods making up the central business district, recognized by the literature as the most likely place where concentrated poverty arise and stagnate in American metro areas. Alternatively, the reference points could be identified on the basis of historical trends of concentrated poverty, i.e. on those neighborhoods that historically have displayed high levels of concentrated poverty. These different decompositions would allow to capture different aspects of spatial proximity. In the latter case, for instance, one would be able to account for the attractive power of historically distressed neighborhoods on actual and future concentrated poverty.

providing evidence of spatial *clustering* of the poor.[15]

The clustering component of concentrated poverty is relevant for policy analysis for at least two reasons. First, a large N component is evidence of a stronger double burden on welfare of the poor that us due to concentrated poverty: not only a disproportionate majority of poor people live in poor neighborhoods, but these neighborhoods are spatially concentrated. Clustering might hence be symptoms of lack of access to transportation, to the job market, to high-quality supply of public goods and definitely to economic and social opportunities offered by cities. When poverty clustering overlaps with administrative divisions of the territory, such as counties (most of the largest American metropolitan areas include more than five 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 intergenerational mobility prospects of the children living therein. Political participation and voting decisions might be as well hampered by the implications of living in strict contact with poor residents.

The second reason is technical: the decomposition is additive and its elements can be separated within and across periods. Simply differentiating these terms allows to picture the neighborhood and non-neighborhood dynamics of concentrated poverty. Studying the N and nN components of the index G, as well as their evolution in time, provide summary information about the extent at which poverty concentrate geographically in a city, and whether the poverty clustering increases or decreases in time both in absolute as well as in relative terms in a way consistent with the underlying axiomatic model A1-A8.

Corollary 2 further shows that the components W, R and $C \cdot E$ are also additively decomposable into a neighborhood and a non-neighborhood component. This decomposition is relevant, for instance, for separating implications of gentrification on clustering. Consider the case in which urban poverty grows on aggregate in the city over the period

[15]The decomposition is close in spirit to a variance decomposition in a spatial context: clusters are identified by their ability of to explain overall variance in the observed data.

t to t', although the component G_N, small, does not vary. There is evidence of clustering (some places are characterized by high poverty incidence, some others by low poverty incidence), although the intensity of clustering does not change across time. This pattern is consistent with two separate yet relevant cases. In one case, it may be that poverty movements in t' occur within the clusterings already existing in t, implying substantial re-ranking of neighboring census tracts (since re-ranking is likely among clustered census tracts display similar levels of concentrated poverty) while clusters remain substantially unchanged: there is little spatial immobility of poverty across spatial clusters of poverty, although there are changes in poverty within the cluster. We expect the component R_N to be large but close in level to $C \cdot E_N$ and R_{nN} is relatively small, while residual changes in urban poverty are driven by E_{nN}. In a second case, instead, there might be substantial changes in poverty clustering across the city, where some clusters disappear and some other appear between period t and t', making G_N small in both periods. In this case, poverty incidence changes across the city and, more importantly, across pre-existing clusters. We expect the component R_{nN} to be large in this case because re-ranking mainly occurs across clusters, while R_N can be small if clusters displaying hight/low poverty concentration in t tend to display low/high poverty concentration in t' (implying that clusters of high and low poverty likely reflect administrative partitions of the urban space), or it can be large if clusters in t' overlap clusters existing in t.

Other examples can be constructed to show that different elements of the decompositions in Corollaries 1 and 2 reflect different aspects of the spatial organization of urban poverty and of the changes in poverty concentration along neighborhoods of the city. We will analyze cross-cities differences in the geographic organization of urban poverty as well as of its changes, making use of information on G_N, G_{nN}, R_N, R_{nN}, E_N and G_{nN} and their associations.

4 Patterns, trends and causes of urban poverty in American cities: 1980-2015

4.1 Data

We assess spatial inequality based on information on incomes distributions within U.S. cities over four decades, drawing on the census files of the U.S. Census Bureau for 1980, 1990 and 2000. Information about population counts, income levels and family composition at a very fine spatial grid is taken from the decennial census Summary Tape File 3A.[16] Due to anonymization issues, the STF 3A data are given in the form of statistical tables representative at the block group level, the finest available statistical partition of the American territory. 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 the 2010-2014 5-years Estimates ACS module. Sampling rates in ACS vary independently at the census block level according to 2010 census population counts, covering on average 2% of the U.S. population over the 2010/14 period. To our knowledge, ACS 2010/14 wave has not yet been used for empirical analysis of urban inequality.

The units of analysis are households with one or more income recipients. The focus is on the gross household income distribution. There are two available sources of information that can be used to model the income distribution at the block group level. The first set of tables display aggregate income at the block group level. The second set of tables show instead counts of households per income interval at the block group level.[17] There are 17

[16]The Census STF 3A provides cross-sectional data for all U.S. States and their subareas in hierarchical sequence down to the block group level (the finest urban space partition available in the census). The geography of the block group partition changes over the decades to keep track with demographic changes within the Counties of each State.

[17]The ACS estimates of population counts should be interpreted as average measures across the 2010-2014 time frame. The survey runs over a five years period to guarantee the representativeness of income and demographic estimates at the block group level.

income intervals in the census 1980, 25 in the census 1990 and 16 in the census 2000 and in the ACS. In all cases, the highest income bracket is not top-coded. We use a methodology based on Pareto distribution fitting as in Nielsen and Alderson (1997), to convert tables of household counts across income intervals into a vector of representative incomes for each income interval, along with the associated vector of households frequencies corresponding to these incomes.[18] Estimates of incomes and household frequencies vary across block groups, implying strong heterogeneity within the city in block-group specific household gross income distributions.

The STF 3A files and the ACS also provide tables of household counts by size (scoring from 1 to 7 or more household members) for each block group. To draw conclusions about the distribution of income across block groups that differ in households demographics, we construct equivalence scales that are representative at the block group level (the square root of average household composition in the block group level, obtained from households counts information). We can hence convert the representative incomes at the block group level into the corresponding equivalized incomes by scaling the estimated reference income values by the block group-specific equivalence scale.

Income reference levels, population frequencies associated with these levels and equivalence scales are estimated separately for each block group of a city in each census and ACS years. All block groups are georeferenced, and measures of distance between the block groups centroids can therefore be constructed. All income observations within the

[18]The procedure consists in fitting a Pareto distribution to the grouped data (population shares and income thresholds) and then estimating references incomes within each interval. For income intervals below the median, the estimated reference income is the midpoint of the interval. For other intervals, estimates are obtained under the constraint that estimated average income at the block-group level should coincide with the observed average income in the data. Estimated medians for top income intervals are used as reference incomes, and empirical population counts as weights. Fitting methods consist in GMM (preferred) and quantile estimation as in Quandt (1966). Alternative estimation methods draw instead from the log-normality assumption, as in Wheeler and La Jeunesse (2008). Incomes estimates based on the preferred method display an MSA-year level average correlation of 95.2% with quantile fitting income estimates (MSA-years population weighted correlations range between $min = 76\%$ and $max = 98.9\%$, with 95% of the correlations larger than 89.3%), and 90.4% average correlation with log-normal fitting income estimates at the block group level (MSA-years population weighted correlations range between $min = 45.6\%$ and $max = 97.1\%$, with 95% of the correlations larger than 85.1%).

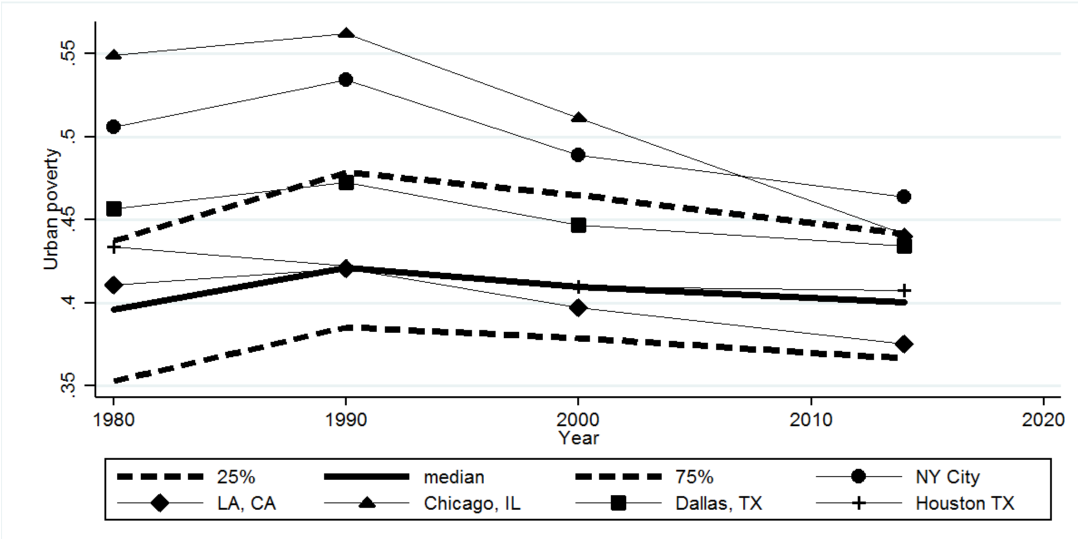
same block group are assumed to occur on its centroid. To identify the relevant urban space, defining the extension of a city, we resort to the Census definition of a Metropolitan Statistical Area (MSA) based on the 1980 Census definition.[19] For each city-year pair we therefore obtain an income database consisting of strings of incomes and frequency weights at each geocoded location on the map. Thus, weighted variants of the GINI index estimators can be used to evaluate facts about spatial inequality at various distance scales.

4.2 Patterns and trends

The distribution of the Gini indices calculated for the cites in our sample in each year is considered, and the quartiles of the distribution (first quartile, median and third quartile) are used to briefly describe each distribution. Figure 4 shows the quartiles of the distributions in years 1980, 1990, 2000 and 2014, together with the urban poverty concentration measured for the five largest US cities in the same years. Urban poverty concentration overall increased from 1980 to 1990, since the quartiles increased over this period. This increase in urban poverty concentration also occurred in the largest cities, except Huston where urban poverty concentration shows a decreasing trend. Urban poverty concentration generally decreased between 1990 and 2014, since the quartiles of the distribution in 2014 were lower than both the quartiles in 2000 and in 1990. New York, Chicago and Dallas were among the top 25 percent of the cites in terms of urban poverty concentration in 1980. Urban poverty concentration considerably decreased in these cities after 1990, especially between 1990 and 2000, with Chicago and Dallas having urban poverty concentrations below the third quartile in 2014. New York remained in the top 25 percent of cities in 2014, however the gap between its urban poverty concentration and the third

[19]The U.S. counties defining the MSAa in 1980 can be found at this link: *http://www.census.gov/population/metro/files/lists/historical/80mfips.txt*. The 1980 Census definition of MSA guarantees comparability of estimates across urban areas that are expanding or shrinking over the 35 years considered in this study.

Figure 4: Urban poverty

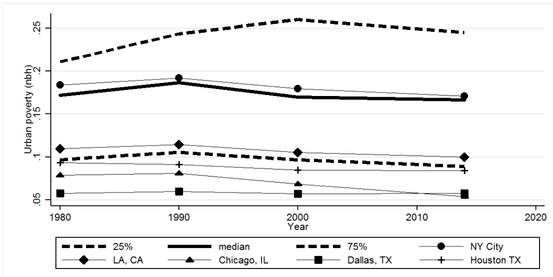


Note: Trends of urban poverty for median, top and bottom quartile cities in the sample, and selected cities.

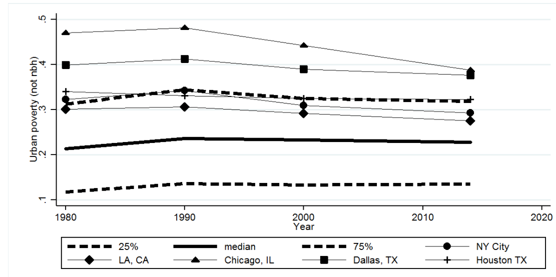
quartile was far larger in 1980 than in 2014.

We now examine the neighborhood and non-neighborhood components of urban poverty concentration. Figure 5 shows the neighborhood component of urban poverty concentration over the 1980-2014 period, while figure **??** shows the non-neighborhood component. The comparison of the two figures indicates that the non-neighborhood component is predominant over the neighborhood component in the largest cities. Since the largest cities have urban poverty concentrations above the median, the non-neighborhood component plays a major role in urban poverty concentration in the most populous cities, especially in those with higher urban poverty concentration (e.g., Dallas, Chicago). New York shows a slightly different trend, since the gap between the non-neighborhood and neighborhood components is not very large, even though the the former overcomes the latter. The trend of the third quartile in figure 5 indicates an increase in the neighborhood component over time, while the median and first quartile slightly changed their values. Unlike, the trends of the quartiles for the non-neighborhood component of urban poverty concentration are quite similar.

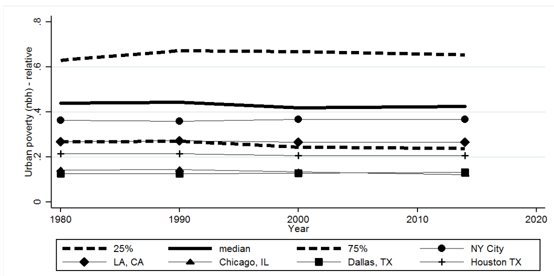
Figure 5: Urban poverty



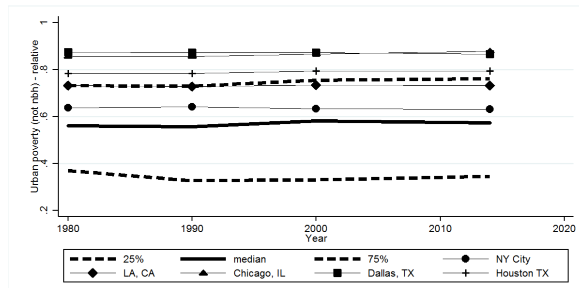
(a) Neighborhood



(b) Non-neighborhood



(c) Neighborhood, relative



(d) Non-neighborhood, relative

Note: Trends of components in urban poverty for median, top and bottom quartile cities in the sample, and selected cities: Neighborhood component G_N, non-neighborhood component G_{nN}, in absolute ($G_N + G_{nN} = G$) and relative (G_N/G, G_{nN}/G) terms.

We now focus on the components of the changes in urban poverty concentration in the five largest cities over the period considered. Table 1 shows the decomposition results for three sub-periods (1980-1990, 1990-2000 and 2000-2014). Relative disparities between neighborhood poverty incidences decreased in the three sub-periods in each of the five cities since E is always negative, with both E_N and E_{nN} negative. This indicates that disparities between poverty incidences reduced both among neighboring census tracts and among non-neighboring census tracts over the period considered. However, the contribution of E was partially offset by the re-ranking contribution, reducing the equalizing effect of the change in disparities between neighborhood poverty incidences. The re-ranking effect totally offset the effect of E during the 1980-1990 sub-period in all cities except Houston, increasing urban poverty concentration over that sub-period. The contribution of W was generally less important than those of the other components.

Table 1: Urban poverty and decomposition of its changes in largest American cities: 1980-2014

component	G_{2014}	G_{1980}	ΔG	W	R	E	C	D
New York								
N	0.17019	0.18415	-0.01396	-0.00130	0.03023	0.04660	0.92052	0.04289
nN	0.29364	0.32194	-0.02829	-0.00478	0.04915	0.07894	0.92052	0.07267
total	0.46383	0.50609	-0.04226	-0.00608	0.07938	0.12554	0.92052	0.11556
Los Angeles								
component	G_{2014}	G_{1980}	ΔG	W	R	E	C	D
N	0.10111	0.10991	-0.00880	-0.00145	0.02405	0.04374	0.71796	0.03140
nN	0.27417	0.30090	-0.02674	-0.00314	0.05852	0.11438	0.71796	0.08212
total	0.37527	0.41082	-0.03554	-0.00459	0.08257	0.15812	0.71796	0.11352
Chicago								
component	G_{2014}	G_{1980}	ΔG	W	R	E	C	D
N	0.05433	0.07874	-0.02441	-0.00430	0.01296	0.05498	0.60145	0.03306
nN	0.38709	0.47047	-0.08338	0.00208	0.08551	0.28427	0.60145	0.17097
total	0.44143	0.54921	-0.10779	-0.00222	0.09847	0.33924	0.60145	0.20404
Huston								
component	G_{2014}	G_{1980}	ΔG	W	R	E	C	D
N	0.08204	0.09308	-0.01104	0.00782	0.03574	0.11300	0.48318	0.05460
nN	0.32547	0.34059	-0.01513	0.04381	0.13542	0.40225	0.48318	0.19436
total	0.40751	0.43368	-0.02617	0.05163	0.17115	0.51525	0.48318	0.24896
Dallas								
component	G_{2014}	G_{1980}	ΔG	W	R	E	C	D
N	0.05800	0.05770	0.00029	0.00666	0.02058	0.05310	0.50754	0.02695
nN	0.37657	0.39889	-0.02232	0.02838	0.13518	0.36623	0.50754	0.18588
total	0.43457	0.45660	-0.02203	0.03504	0.15576	0.41933	0.50754	0.21283

To do list:

- picture trends of median, 25 and 75 in each year.
- picture trends of large cities
- picture trends of each component
- picture patterns of different components
- picture patterns of components vs poverty incidence vs inequality vs pop size vs poverty changes and poverty in initial period
- picture patterns of poverty concentration and changes by looking at gentrification.

- classify cities by components. Assign a dummy to those whose G grows more than the median city in 2000 and over the following decade,then assign a dummy to each component N and nN of growth: RN RnN EN EnN. These dummies can be assinged wrt to the median city in a given year, or with respect to the relative weight of each component (Rn/R¿¡1?) to indicate types of cities. Each city is an observation and the dummies characterize a patterns of urban poverty and its change in that city. We can hence construct groups. We then analyse how charcteristics of the cities in 2000 affect the probability that a city belongs to one or the other group (hence displaying a specific patterns of urban poverty)

- Use regression methods to assess the implications of gentrification on urban poverty and on the components of urban poverty changes: model 1) studies the effect of variables and past gentrification on 2000 urban poverty. Model 2) studies how the same variables have affected changes in urban poverty over the decade.

4.3 Causes of urban poverty and the Big Sort hypothesis

TO BE COMPLETED

5 Conclusions

TO BE COMPLETED

References

Andreoli, F. and Peluso, E. (2017). So close yet so unequal: Spatial inequality in american cities, *LISER, 2017, Working Papers n°2017-11* .

Andreoli, F. and Zoli, C. (2014). Measuring dissimilarity, *Working Papers Series, Department of Economics, Univeristy of Verona, WP23* .

Atkinson, T. 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.

Chetty, R. and Hendren, N. (2016). The impacts of neighborhoods on intergenerational mobility i: Childhood exposure effects. mimeo.

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.

Chetty, R., Hendren, N., Kline, P. and Saez, E. (2014). Where is the land of opportunity? The geography of intergenerational mobility in the United States, *The Quarterly Journal of Economics* **129**(4): 1553–1623.

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.

Frankel, D. M. and Volij, O. (2011). Measuring school segregation, *Journal of Economic Theory* **146**(1): 1 – 38.

Hutchens, R. M. (1991). Segregation curves, Lorenz curves, and inequality in the distribution of people across occupations, *Mathematical Social Sciences* **21**(1): 31 – 51.

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.

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.

Marx, I., Nolan, B. and Olivera, J. (2015). Chapter 23 - the welfare state and antipoverty policy in rich countries, *in* A. B. Atkinson and F. Bourguignon (eds), *Handbook of Income Distribution*, Vol. 2 of *Handbook of Income Distribution*, Elsevier, pp. 2063 – 2139.

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. and Grossi, L. (2015). Decomposing changes in CO_2 emission inequality over time: the roles of re-ranking and changes in per capita CO_2 emission disparities, *Energy Economics* **49**: 274–281.

Nielsen, F. and Alderson, A. S. (1997). The Kuznets curve and the great u-turn: Income inequality in U.S. counties, 1970 to 1990, *American Sociological Review* **62**(1): 12–33.

Quandt, R. (1966). Old and new methods of estimation and the Pareto distribution, *Metrika* **10**: 55–82.

Rey, S. J. and Smith, R. J. (2013). A spatial decomposition of the Gini coefficient, *Letters in Spatial and Resource Sciences* **6**: 55–70.

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.

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.*

Supplemental appendix

A Proofs

A.1 Proof of Lemma 1

Proof. Consider first the case in which $\frac{P_1}{N_1} = \ldots = \frac{P_z}{N_z} = \frac{P^*}{N^*}$ and $N_1 = \ldots = N_z = N^*$ with P^* and N^* two natural numbers such that $\frac{P^*}{N^*} \leq \zeta$. Under axioms A1 and A2 we write:

$$\begin{aligned} UP(.;\alpha) &= A(.;\alpha)\sum_{i=1}^{z}\frac{N^*}{N}\left(\frac{P^*/N^*}{\zeta}-1\right)w_i(.;\alpha) \\ &= A(.;\alpha)\frac{N^*}{N}\left(\frac{P^*/N^*}{\zeta}-1\right)\sum_{i=1}^{z}w_i(.;\alpha). \end{aligned} \tag{5}$$

Axioms A4 and A5 imply that (5) can be written as follows:

$$\begin{aligned} UP(.;\alpha) &= A(.;\alpha)\frac{N^*}{N}\left(\frac{P^*/N^*}{\zeta}-1\right)\sum_{i=1}^{z}\left(\sum_{j=1}^{z}\frac{N^*}{N}-\sum_{j=1}^{i}\frac{N^*}{N}+\frac{N^*}{N}\right) \\ &= A(.;\alpha)\frac{N^*}{N}\left(\frac{P^*/N^*}{\zeta}-1\right)\frac{N^*}{N}\sum_{i=1}^{z}(z-i+1) \\ &= A(.;\alpha)\frac{N^*}{N}\left(\frac{P^*/N^*}{\zeta}-1\right)\frac{N^*}{N}\frac{z(z+1)}{2}. \end{aligned} \tag{6}$$

According to axiom A3, the index $UP(.;\alpha)$ can be also written as follows:

$$\begin{aligned} UP(.;\alpha) &= \alpha HI = \alpha\sum_{i}^{z}\frac{N^*}{N}\left(\frac{P^*/N^*}{\zeta}-1\right) \\ &= \alpha\frac{N^*}{N}\left(\frac{P^*/N^*}{\zeta}-1\right)z. \end{aligned} \tag{7}$$

Equating (6) to (7) and solving for $A(.;\alpha)$ we obtain the following specification for the scaling coefficient:

$$\begin{aligned} A(.;\alpha) &= \frac{2\alpha}{z+1}\frac{N}{N^*} \\ &= \frac{2\alpha z}{z+1}\frac{1}{H}, \end{aligned} \tag{8}$$

where (8) follows from the fact that $N^* = \sum_{i=1}^{z}\frac{N_i}{z}$ and from the definition of H.

Using the definition of rank-dependent weights consistent with axioms A4 and A5, and substituting for (8), we can write:

$$
\begin{aligned}
UP(.;\alpha) &= \frac{2\alpha z}{z+1}\frac{1}{H}\sum_{i=1}^{z}\frac{N_i}{N}\left(\frac{P_i/N_i}{\zeta}-1\right)\left(\sum_{j=1}^{z}\frac{N_j}{N}-\sum_{j=1}^{i}\frac{N_j}{N}+\frac{N_i}{N}\right) \\
&= \frac{2\alpha z}{z+1}\frac{1}{H}\left[\frac{1}{\zeta}\sum_{i=1}^{z}\frac{N_i}{N}\frac{P_i}{N_i}\left(\sum_{j=1}^{z}\frac{N_j}{N}-\sum_{j=1}^{i}\frac{N_j}{N}+\frac{N_i}{N}\right)-\sum_{i=1}^{z}\frac{N_i}{N}\left(\sum_{j=1}^{z}\frac{N_j}{N}-\sum_{j=1}^{i}\frac{N_j}{N}+\frac{N_i}{N}\right)\right] \quad (9)
\end{aligned}
$$

We show now that the first term within square brackets in (9) can be expressed as a function of known elements and of the Gini index $G(.;\alpha)$, measuring the unequal distribution of poverty shares (P_i/N_i) across the neighborhood where poor people are mostly concentrated.

Let m_α denote the average incidence of poverty among the neighborhoods in which poverty is more concentrated given poverty threshold defined by α, so that

$$
m_\alpha = \sum_{i=1}^{z}\frac{1}{\sum_{i=1}^{z}N_i/N}\frac{N_i}{N}\frac{P_i}{N_i}. \quad (10)
$$

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

$$
\begin{aligned}
G(.;\alpha) &= \frac{1}{2m_\alpha\left(\sum_{i=1}^{z}N_j/N\right)^2}\sum_{i=1}^{z}\sum_{j=1}^{z}\frac{N_i}{N}\frac{N_j}{N}\left|\frac{P_i}{N_i}-\frac{P_j}{N_j}\right| \\
&= \frac{1}{2m_\alpha\left(\sum_{i=1}^{z}N_j/N\right)^2}\sum_{i=1}^{z}\sum_{j=1}^{z}\frac{N_i}{N}\frac{N_j}{N}\left[2\max\left\{\frac{P_i}{N_i},\frac{P_j}{N_j}\right\}-\frac{P_i}{N_i}-\frac{P_j}{N_j}\right] \\
&= \frac{1}{2m_\alpha\left(\sum_{i=1}^{z}N_j/N\right)^2}\left[\sum_{i=1}^{z}\sum_{j=1}^{z}\frac{N_i}{N}\frac{N_j}{N}2\max\left\{\frac{P_i}{N_i},\frac{P_j}{N_j}\right\}-2\sum_{i=1}^{z}\frac{N_i}{N}\sum_{i=1}^{z}\frac{N_i}{N}\frac{P_i}{N_i}\right] \quad (11)
\end{aligned}
$$

We now work out the first term appearing in squared brackets in (11), denoted *max* in short-hand notation, to show that it can written as a function of the rank weights. First, let develop the double summations term as follows:

$$\begin{aligned}
max &= \sum_{i=1}^{z}\sum_{j=1}^{z}\frac{N_i}{N}\frac{N_j}{N}\max\left\{\frac{P_i}{N_i},\frac{P_j}{N_j}\right\}\\
&= \frac{N_1}{N}\frac{N_1}{N}\frac{P_1}{N_1}+\left(\frac{N_1}{N}\frac{N_2}{N}\frac{P_1}{N_1}+\ldots+\frac{N_1}{N}\frac{N_z}{N}\frac{P_1}{N_1}\right)+\\
&\quad+\frac{N_2}{N}\frac{N_1}{N}\frac{P_1}{N_1}+\frac{N_2}{N}\frac{N_2}{N}\frac{P_2}{N_2}+\left(\frac{N_2}{N}\frac{N_3}{N}\frac{P_2}{N_2}+\ldots+\frac{N_2}{N}\frac{N_z}{N}\frac{P_2}{N_2}\right)+\\
&\quad+\frac{N_3}{N}\frac{N_1}{N}\frac{P_1}{N_1}+\frac{N_3}{N}\frac{N_2}{N}\frac{P_2}{N_2}+\frac{N_3}{N}\frac{N_3}{N}\frac{P_3}{N_3}+\left(\frac{N_2}{N}\frac{N_4}{N}\frac{P_3}{N_3}+\ldots+\frac{N_3}{N}\frac{N_z}{N}\frac{P_3}{N_3}\right)+\\
&\quad\ldots+\frac{N_{z-1}}{N}\frac{N_1}{N}\frac{P_1}{N_1}+\ldots+\frac{N_{z-1}}{N}\frac{N_{z-1}}{N}\frac{P_{z-1}}{N_{z-1}}+\frac{N_{z-1}}{N}\frac{N_z}{N}\frac{P_{z-1}}{N_{z-1}}+\\
&\quad+\frac{N_z}{N}\frac{N_1}{N}\frac{P_1}{N_1}+\ldots+\frac{N_z}{N}\frac{N_z}{N}\frac{P_z}{N_z}.
\end{aligned}$$

Rearranging the terms in the summation, this quantity can be equivalently written as:

$$\begin{aligned}
max &= \frac{N_1}{N}\frac{P_1}{N_1}\left(\sum_{j=1}^{z}\frac{N_j}{N}+\sum_{j=2}^{z}\frac{N_j}{N}\right)+\frac{N_2}{N}\frac{P_2}{N_2}\left(\sum_{j=2}^{z}\frac{N_j}{N}+\sum_{j=3}^{z}\frac{N_j}{N}\right)+\\
&\quad\ldots+\frac{N_{z-1}}{N}\frac{P_{z-1}}{N_{z-1}}\left(\sum_{j=z-1}^{z}\frac{N_j}{N}+\frac{N_z}{N}\right)+\frac{N_z}{N}\frac{P_z}{N_z}\frac{N_z}{N}\\
&= \sum_{i=1}^{z}\frac{N_i}{N}\frac{P_i}{N_i}\left(\frac{N_i}{N}+2\sum_{j=i+1}^{z}\frac{N_j}{N}\right) \qquad (12)
\end{aligned}$$

After adding and subtracting the quantity $\sum_{i=1}^{z}\frac{N_i}{N}\frac{P_i}{N_i}\frac{N_i}{N}$, we obtain:

$$\begin{aligned}
max &= \sum_{i=1}^{z}\frac{N_i}{N}\frac{P_i}{N_i}\left(2\frac{N_i}{N}+2\sum_{j=i+1}^{z}\frac{N_j}{N}\right)-\sum_{i=1}^{z}\frac{N_i}{N}\frac{P_i}{N_i}\frac{N_i}{N}\\
&= \sum_{i=1}^{z}\frac{N_i}{N}\frac{P_i}{N_i}\left(2\frac{N_i}{N}+2\left(\sum_{j=1}^{z}\frac{N_j}{N}-\sum_{j=1}^{i}\frac{N_j}{N}\right)\right)-\sum_{i=1}^{z}\frac{N_i}{N}\frac{P_i}{N_i}\frac{N_i}{N}\\
&= 2\sum_{i=1}^{z}\frac{N_i}{N}\frac{P_i}{N_i}\left(\sum_{j=1}^{z}\frac{N_j}{N}-\sum_{j=1}^{i}\frac{N_j}{N}+\frac{N_i}{N}\right)-\sum_{i=1}^{z}\frac{N_i}{N}\frac{P_i}{N_i}\frac{N_i}{N}, \qquad (13)
\end{aligned}$$

where the term in parenthesis in (13) coincide with the rank weights identified by axioms A4 and A5. We can now substitute the term max in (11) with the expression (13). Using the explicit formula for m_α, we obtain:

$$\begin{aligned} G(.;\alpha) &= \frac{1}{\left(\sum_{i=1}^{z} N_j/N\right)^2 \sum_{i=1}^{z} \frac{1}{\sum_{i=1}^{z} N_i/N} \frac{N_i}{N}\frac{P_i}{N_i}} \left[2\sum_{i=1}^{z} \frac{N_i}{N}\frac{P_i}{N_i}\left(\sum_{j=1}^{z}\frac{N_j}{N} - \sum_{j=1}^{i}\frac{N_j}{N} + \frac{N_i}{N}\right) - \sum_{i=1}^{z}\frac{N_i}{N}\frac{P_i}{N_i}\frac{N_i}{N}\right] - \\ &\quad - \frac{1}{\left(\sum_{i=1}^{z} N_j/N\right)^2 \sum_{i=1}^{z} \frac{1}{\sum_{i=1}^{z} N_i/N} \frac{N_i}{N}\frac{P_i}{N_i}} 2\left(\sum_{i=1}^{z}\frac{N_i}{N}\sum_{i=1}^{z}\frac{N_i}{N}\frac{P_i}{N_i}\right) \\ &= \frac{2}{\sum_{i=1}^{z}\frac{N_i}{N}\sum_{i=1}^{z}\frac{N_i}{N}\frac{P_i}{N_i}} \sum_{i=1}^{z}\frac{N_i}{N}\frac{P_i}{N_i}\left(\sum_{j=1}^{z}\frac{N_j}{N} - \sum_{j=1}^{i}\frac{N_j}{N} + \frac{N_i}{N}\right) - \\ &\quad - \frac{1}{\sum_{i=1}^{z}\frac{N_i}{N}\sum_{i=1}^{z}\frac{N_i}{N}\frac{P_i}{N_i}} \sum_{i=1}^{z}\frac{N_i}{N}\frac{N_i}{N}\frac{P_i}{N_i} - 1 \end{aligned} \tag{14}$$

The second term of (14) is a function of z. If the number of neighborhood is large enough, and the neighbrohood are small enough in size, this term converges to zero at a rate that is quadratic in the demographic size of the neighborhood. Hereafter we maintain that the number of neighborhoods is large, so that the rank weights in (14) can be approximated as follows:

$$\sum_{i=1}^{z}\frac{N_i}{N}\frac{P_i}{N_i}\left(\sum_{j=1}^{z}\frac{N_j}{N} - \sum_{j=1}^{i}\frac{N_j}{N} + \frac{N_i}{N}\right) \approx \frac{1}{2}(G(.;\alpha)+1)\sum_{i=1}^{z}\frac{N_i}{N}\sum_{i=1}^{z}\frac{N_i}{N}\frac{P_i}{N_i}. \tag{15}$$

Substituting (15) into (9) and using the fact that

$$\begin{aligned} \sum_{i=1}^{z}\frac{N_i}{N}\left(\sum_{j=1}^{z}\frac{N_j}{N} - \sum_{j=1}^{i}\frac{N_j}{N} + \frac{N_i}{N}\right) &= \sum_{i=1}^{z}\frac{N_i}{N}\left(\sum_{j=1}^{z}\frac{N_j}{N} - \sum_{j=1}^{i-1}\frac{N_j}{N}\right) \\ &= H^2 - \sum_{i=1}^{z}\frac{N_i}{N}\sum_{j=1}^{i-1}\frac{N_j}{N}, \end{aligned}$$

we get:

$$UP(.;\alpha) = \frac{2\alpha z}{z+1}\frac{1}{H}\left[\frac{1}{2}(G(.;\alpha)+1)\sum_{i=1}^{z}\frac{N_i}{N}\frac{1}{\zeta}\sum_{i=1}^{z}\frac{N_i}{N}\frac{P_i}{N_i} - H^2 + \sum_{i=1}^{z}\frac{N_i}{N}\sum_{j=1}^{i-1}\frac{N_j}{N}\right] \tag{16}$$

Adding and subtracting the term $\frac{1}{2}(G_\alpha+1)\left(\sum_{i=1}^{z} N_i/N\right)^2$ within square brackets in (16) we obtain:

$$\begin{aligned}UP(.;\alpha) &= \frac{2\alpha z}{z+1}\frac{1}{H}\left[\frac{1}{2}(G(.;\alpha)+1)H^2 I + \frac{1}{2}(G(.;\alpha)+1)H^2 - H^2 + \sum_{i=1}^{z}\frac{N_i}{N}\sum_{j=1}^{i-1}\frac{N_j}{N}\right]\\ &= \frac{\alpha z}{z+1}\frac{1}{H}\left[(G(.;\alpha)+1)H^2 I + \frac{1}{2}(G(.;\alpha)+1)H^2 - 2H^2 + 2\sum_{i=1}^{z}\frac{N_i}{N}\sum_{j=1}^{i-1}\frac{N_j}{N}\right]\\ &= \frac{\alpha z}{z+1}H\left[I + (I+1)G(.;\alpha) - 1 + \frac{2}{H^2}\sum_{i=1}^{z}\frac{N_i}{N}\sum_{j=1}^{i-1}\frac{N_j}{N}\right],\end{aligned}$$

which concludes the proof. ■

A.2 Proof of Theorem 1

Proof. Axioms A1-A5 are equivalent to (4). For any given configuration $\mathcal{A}$ with n neighborhoods, consider now an alternative configuration $\mathcal{A}'$ with $n' > n$ neighborhoods obtained from $\mathcal{A}$ by operations of splitting of neighborhoods, so that $\left(\frac{N_1^{\mathcal{A}}}{N}, \ldots, \frac{N_n^{\mathcal{A}}}{N}\right) \rightarrow \left(\frac{N_1^{\mathcal{A}'}}{N}, \ldots, \frac{N_{n'}^{\mathcal{A}'}}{N}\right)$ and $\frac{N_i^{\mathcal{A}'}}{N} = \frac{1}{n'}$ for any $i = 1, \ldots, n'$. Let z' be the poverty line at given α and for $\frac{P^{\mathcal{A}'}}{N^{\mathcal{A}'}} = \frac{P^{\mathcal{A}'}}{N^{\mathcal{A}'}}$. We can hence write for $UP(\mathcal{A}';\alpha)$:

$$\begin{aligned}2\sum_{i=1}^{z'}\frac{N_i^{\mathcal{A}'}}{N^{\mathcal{A}'}}\sum_{j=1}^{i-1}\frac{N_j^{\mathcal{A}'}}{N^{\mathcal{A}'}} &= 2\sum_{i=1}^{z'}\frac{1}{n'}\sum_{j=1}^{i-1}\frac{1}{n'} = \frac{2}{n'^2}\sum_{i=1}^{n'}(i-1)\\ &= \frac{2}{n'^2}\left(\frac{n'(n'+1)}{2} - n'\right) = \frac{n'-1}{n'} \approx 1, \qquad (17)\end{aligned}$$

when the number of neighborhood n' is large. From Axiom A8 it follows that $z \rightarrow n$ and $H \rightarrow 1$. Axiom A7 along with the fact that n is large imply that there always exists a neighborhood z such that $\frac{P_z}{N_z} \approx \alpha\frac{P}{N}$. Axiom A8 would then give:

$$\lim_{n\to\infty} I = \lim_{n\to\infty}\sum_{i=1}^{z}\frac{N_i/N}{\sum_{i=1}^{z}N_i/N}\left(\frac{P_i/N_i}{P_z/N_z} - 1\right) = \frac{1}{\alpha}\sum_{i=1}^{n}\frac{N_i}{N}\frac{P_i/N_i}{P/N} - 1 = 0.$$

Axiom A6, along with the result in (17) and the fact that $H = 1$ and $I = 0$ under axioms A7 and A8, give that $UP(\mathcal{A},\alpha) = G(\mathcal{A})$. ■

A.3 Proof of Corollary 1

Proof. Let neighborhood i be the neighborhood having rank i when neighborhoods are sorted in decreasing order of neighborhood poverty incidence. To simplify notation, 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 $\mathbf{p} = (p_1, \ldots, p_n)^T$ be the $n \times 1$ vector of neighborhood poverty incidences sorted in decreasing order and $\mathbf{s} = (s_1, \ldots, s_n)^T$ be the $n \times 1$ vector of the corresponding population shares. A 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:

$$\mathbf{P} = \frac{1}{\bar{p}} \left(\mathbf{1}_n \mathbf{p}^T - \mathbf{p} \mathbf{1}_n^T \right) = \begin{bmatrix} \frac{p_1 - p_1}{\bar{p}} & \cdots & \frac{p_n - p_1}{\bar{p}} \\ \vdots & \ddots & \vdots \\ \frac{p_1 - p_n}{\bar{p}} & \cdots & \frac{p_n - p_n}{\bar{p}} \end{bmatrix}, \tag{18}$$

where $\bar{p}$ is the overall poverty incidence 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 $\mathbf{S} = diag\left\{\mathbf{s}\right\}$ 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:

$$G(\mathbf{s}, \mathbf{p}) = \frac{1}{2} tr \left(\tilde{\mathbf{G}} \mathbf{P}^T \right), \tag{19}$$

where the matrix $\tilde{\mathbf{G}} = \mathbf{SGS}$ 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. Suppose that neighborhood poverty incidences and population shares are observed in times t and t'. Let $\mathbf{p}_t$ be the $n \times 1$ vector of the t poverty incidences sorted in decreasing order and $\mathbf{s}_t$ be the $n \times 1$ vector of the corresponding population shares. Let $\mathbf{p}_{t'}$ be the $n \times 1$ vector of the t' poverty incidences sorted in decreasing order and $\mathbf{s}_{t'}$ be the $n \times 1$ vector of the corresponding population shares. The change in urban poverty concentration from t to t' is measured by the difference between the Gini index in t' and the Gini index in t:

$$\Delta UP = G(\mathbf{s}_{t'}, \mathbf{p}_{t'}) - G(\mathbf{s}_t, \mathbf{p}_t) = \frac{1}{2} tr \left(\tilde{\mathbf{G}}_{t'} \mathbf{P}_{t'}^T \right) - \frac{1}{2} tr \left(\tilde{\mathbf{G}}_t \mathbf{P}_t^T \right). \tag{20}$$

Equation (20) 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 disparity of neighborhood poverty incidences. 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 incidences. Let $\lambda = \bar{p}_{t'} / \bar{p}_{t'|t}$ be the ratio of the actual t' overall poverty incidence to the fictitious t' overall poverty incidence which is the weighted average of t' poverty incidences where the weights are the corresponding population shares in t. After defining $\mathbf{S}_{t|t'} = diag\left\{\mathbf{s}_{t|t'}\right\}$, the Gini index of t' neighborhood poverty incidences calculated by using the t neighborhood population shares is

$$\begin{aligned} G\left(\mathbf{s}_{t|t'}, \mathbf{p}_{t'}\right) &= \frac{1}{2} tr \left(\mathbf{S}_{t|t'} \mathbf{G} \mathbf{S}_{t|t'} \lambda \mathbf{P}_{t'}^T \right) \\ &= \frac{1}{2} tr \left(\tilde{\mathbf{G}}_{t|t'} \lambda \mathbf{P}_{t'}^T \right) \end{aligned} \tag{21}$$

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 (21), 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 $G\left(\mathbf{s}_{t|t'}, \mathbf{p}_{t'}\right)$ in equation (20), the contribution to ΔUP due to changes in neighborhood population shares can be separated from that attributable to changes in disparities between neighborhood poverty incidences:

$$\begin{aligned}\Delta UP &= \left[\frac{1}{2}tr\left(\tilde{\mathbf{G}}_{t'}\mathbf{P}_{t'}^T\right) - \frac{1}{2}tr\left(\tilde{\mathbf{G}}_{t|t'}\lambda\mathbf{P}_{t'}^T\right)\right] + \left[\frac{1}{2}tr\left(\tilde{\mathbf{G}}_{t|t'}\lambda\mathbf{P}_{t'}^T\right) - \frac{1}{2}tr\left(\tilde{\mathbf{G}}_{t}\mathbf{P}_{t}^T\right)\right] \\ &= \frac{1}{2}tr\left(\mathbf{W}\mathbf{P}_{t'}^T\right) + \left[\frac{1}{2}tr\left(\tilde{\mathbf{G}}_{t|t'}\lambda\mathbf{P}_{t'}^T\right) - \frac{1}{2}tr\left(\tilde{\mathbf{G}}_{t}\mathbf{P}_{t}^T\right)\right] \\ &= W + \left[\frac{1}{2}tr\left(\tilde{\mathbf{G}}_{t|t'}\lambda\mathbf{P}_{t'}^T\right) - \frac{1}{2}tr\left(\tilde{\mathbf{G}}_{t}\mathbf{P}_{t}^T\right)\right],\end{aligned} \tag{22}$$

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 concentration 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 concentration from t to t'.

The difference enclosed within square brackets on the right-hand side of equation (22) can be additively split into two components: one component measuring the re-ranking of neighborhoods, a second component measuring the change in disparity of neighborhood poverty incidences. Let $\mathbf{p}_{t'|t}$ be the $n \times 1$ vector of t' neighborhood poverty incidences sorted in decreasing order of the respective t neighborhood poverty incidences, 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 $\mathbf{P}_{t'|t} = \left(1/\bar{p}_{t'|t}\right)\left(\mathbf{1}_n\mathbf{p}_{t'|t}^T - \mathbf{p}_{t'|t}\mathbf{1}_n^T\right)$ contains the n^2 relative pairwise differences between the neighborhood poverty incidences as arranged in $\mathbf{p}_{t'|t}$. The concentration index of the t' poverty incidences sorted by the t poverty incidences, calculated by using the t population shares, is defined as follows:

$$C\left(\mathbf{s}_t, \mathbf{p}_{t'|t}\right) = \frac{1}{2}tr\left(\tilde{\mathbf{G}}_t\mathbf{P}_{t'|t}^T\right). \tag{23}$$

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 $\mathbf{P}_{t'|t} = \mathbf{B}\lambda\mathbf{P}_{t'}\mathbf{B}^T$, the concentration index $C\left(\mathbf{s}_t, \mathbf{p}_{t'|t}\right)$ expressed as a function of $\mathbf{P}_{t'}$ becomes

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

By adding $C\left(\mathbf{s}_t, \mathbf{p}_{t'|t}\right)$ as expressed in (23) and subtracting it as expressed in (24) to the difference enclosed within square brackets on the right-hand side of equation (22), we obtain

$$\begin{aligned}
\frac{1}{2}tr\left(\tilde{\mathbf{G}}_{t|t'}\lambda\mathbf{P}_{t'}^T\right) - \frac{1}{2}tr\left(\tilde{\mathbf{G}}_t\mathbf{P}_t^T\right) &= \left[\frac{1}{2}tr\left(\tilde{\mathbf{G}}_{t|t'}\lambda\mathbf{P}_{t'}^T\right) - \frac{1}{2}tr\left(\mathbf{B}^T\tilde{\mathbf{G}}_t\mathbf{B}\lambda\mathbf{P}_{t'}^T\right)\right] \\
&+ \left[\frac{1}{2}tr\left(\tilde{\mathbf{G}}_t\mathbf{P}_{t'|t}^T\right) - \frac{1}{2}tr\left(\tilde{\mathbf{G}}_t\mathbf{P}_t^T\right)\right] \\
&= \frac{1}{2}tr\left[\left(\tilde{\mathbf{G}}_{t|t'} - \mathbf{B}^T\tilde{\mathbf{G}}_t\mathbf{B}\right)\lambda\mathbf{P}_{t'}^T\right] \\
&+ \frac{1}{2}tr\left[\tilde{\mathbf{G}}_t\left(\mathbf{P}_{t'|t}^T - \mathbf{P}_t^T\right)\right] \\
&= \frac{1}{2}tr\left(\mathbf{R}\lambda\mathbf{P}_{t'}^T\right) + \frac{1}{2}tr\left(\tilde{\mathbf{G}}_t\mathbf{D}^T\right) \\
&= R + D,
\end{aligned} \tag{25}$$

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 concentration 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 change between neighborhood poverty incidences. The generic (i, j)-th element of $\mathbf{D}$ compares the relative difference between the t poverty incidences 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 negative value of D means that relative disparities in neighborhood poverty incidences have overall decreased from t to t', reducing urban poverty concentration. A positive value of D indicates that relative disparities in neighborhood poverty incidences have overall increased from t to t', increasing urban poverty concentration. If all neighborhood poverty incidences have changed by the same proportion from t to t', then $D = 0$.

Given equations (22) and (25), a three-term decomposition of ΔUP is obtained:

$$\Delta UP = \frac{1}{2}tr\left(\mathbf{W}\mathbf{P}_{t'}^T\right) + \frac{1}{2}tr\left(\mathbf{R}\lambda\mathbf{P}_{t'}^T\right) + \frac{1}{2}tr\left(\tilde{\mathbf{G}}_t\mathbf{D}^T\right) = W + R + D. \tag{26}$$

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

$$c = \frac{\bar{p}_{t'|t} - \bar{p}_t}{\bar{p}_t}. \tag{27}$$

Let $\mathbf{p}_{t'|t}^c = \mathbf{p}_t + c\mathbf{p}_t$ be the vector of neighborhood poverty incidences we would observe in t' if every neighborhood poverty incidence changed by proportion c. This implies that

$\bar{p}^c_{t'|t} = \bar{p}_{t'|t}$. Vector $\mathbf{p}_{t'|t}$ can be expressed as

$$\mathbf{p}_{t'|t} = \mathbf{p}^c_{t'|t} + \mathbf{p}^\delta_{t'|t},$$

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

$$\begin{aligned} \mathbf{p}_{t'|t} &= \underbrace{\mathbf{p}_t + \mathbf{p}^\delta_{t'|t}}_{\mathbf{p}^e_{t'|t}} + c\mathbf{p}_t \\ &= \mathbf{p}^e_{t'|t} + c\mathbf{p}_t, \end{aligned} \tag{28}$$

where the elements of $\mathbf{p}^e_{t'|t}$ account for disproportionate changes in neighborhood poverty incidences from t to t', as $\mathbf{p}^e_{t'|t}$ would equal $\mathbf{p}_t$ if there were no disproportionate changes in neighborhood poverty incidences. Given equations (27) and (29), matrix $\mathbf{P}_{t'|t}$ can be written as

$$\begin{aligned} \mathbf{P}_{t'|t} &= \left(1/\bar{p}_{t'|t}\right)\left(\mathbf{1}_n \mathbf{p}^T_{t'|t} - \mathbf{p}_{t'|t}\mathbf{1}^T_n\right) \\ &= \frac{1}{1+c}\begin{bmatrix} \frac{p^e_{1,t'|t} - p^e_{1,t'|t}}{\bar{p}_t} & \cdots & \frac{p^e_{n,t'|t} - p^e_{1,t'|t}}{\bar{p}_t} \\ \vdots & \ddots & \vdots \\ \frac{p^e_{1,t'|t} - p^e_{n,t'|t}}{\bar{p}_t} & \cdots & \frac{p^e_{n,t'|t} - p^e_{n,t'|t}}{\bar{p}_t} \end{bmatrix} + \frac{c}{1+c}\begin{bmatrix} \frac{p_{1,t} - p_{1,t}}{\bar{p}_t} & \cdots & \frac{p_{n,t} - p_{1,t}}{\bar{p}_t} \\ \vdots & \ddots & \vdots \\ \frac{p_{1,t} - p_{n,t}}{\bar{p}_t} & \cdots & \frac{p_{n,t} - p_{n,t}}{\bar{p}_t} \end{bmatrix} \\ &= \frac{1}{1+c}\mathbf{P}^e_{t'|t} + \frac{c}{1+c}\mathbf{P}_t. \end{aligned} \tag{29}$$

Since matrix $\mathbf{D}$ in equation (26) is obtained by subtracting $\mathbf{P}_t$ from $\mathbf{P}_{t'|t}$, $\mathbf{D}$ can be re-written as

$$\begin{aligned} \mathbf{D} &= \mathbf{P}_{t'|t} - \mathbf{P}_t \\ &= \frac{1}{1+c}\mathbf{P}^e_{t'|t} + \frac{c}{1+c}\mathbf{P}_t - \mathbf{P}_t \\ &= \underbrace{\left(\frac{1}{1+c}\right)}_{C}\underbrace{\left(\mathbf{P}^e_{t'|t} - \mathbf{P}_t\right)}_{\mathbf{E}} \\ &= C\mathbf{E} \end{aligned} \tag{30}$$

By replacing $\mathbf{D}$ in equation (26) with its expression in equation (30), the decomposition of the change in urban poverty concentration becomes

$$\Delta UP = \frac{1}{2}tr\left(\mathbf{W}\mathbf{P}^T_{t'}\right) + \frac{1}{2}tr\left(\mathbf{R}\lambda\mathbf{P}^T_{t'}\right) + C\frac{1}{2}tr\left(\tilde{\mathbf{G}}_t\mathbf{E}^T\right) = W + R + CE. \tag{31}$$

■

A.4 Proof of Corollary 2

Proof. 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 incidences of two neighboring neighborhoods, otherwise the (i, j)-th element of $\mathbf{N}_t$ is 0. Using the Hadamard product,[20] the relative pairwise differences between the poverty incidences of neighboring neighborhoods can be selected from $\mathbf{P}_t$:

$$\mathbf{P}_{N,t} = \mathbf{N}_t \odot \mathbf{P}_t. \tag{32}$$

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

$$\mathbf{P}^e_{N,t'|t} = \mathbf{N}_t \odot \mathbf{P}^e_{t'|t}. \tag{33}$$

Since $\mathbf{E} = \mathbf{P}^e_{t'|t} - \mathbf{P}_t$, 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 neighboring neighborhoods:

$$\mathbf{E}_N = \mathbf{P}^e_{N,t'|t} - \mathbf{P}_{N,t} = \mathbf{N}_t \odot \left(\mathbf{P}^e_{t'|t} - \mathbf{P}_t\right) = \mathbf{N}_t \odot \mathbf{E}. \tag{34}$$

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 incidences of two neighboring neighborhoods, 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

$$\mathbf{P}_{N,t'} = \mathbf{N}_{t'} \odot \mathbf{P}_{t'}. \tag{35}$$

The nonzero elements of $\mathbf{P}_{N,t'}$ are the relative pairwise differences between the t' poverty incidences of neighboring neighborhoods.

The decomposition of the change in the neighbor component of urban poverty concentration is obtained by replacing $\mathbf{P}_{t'}$ and $\mathbf{E}$ in equation (31) with $\mathbf{P}_{N,t'}$ and $\mathbf{E}_N$ respectively:

$$\Delta UP_N = \frac{1}{2} tr\left(\mathbf{W}\mathbf{P}^T_{N,t'}\right) + \frac{1}{2} tr\left(\mathbf{R}\lambda\mathbf{P}^T_{N,t'}\right) + C\frac{1}{2} tr\left(\tilde{\mathbf{G}}_t\mathbf{E}^T_N\right) = W_N + R_N + CE_N. \tag{36}$$

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 incidences of non-neighboring neighborhoods is

$$\mathbf{P}_{nN,t'} = (\mathbf{J}_n - \mathbf{N}_{t'}) \odot \mathbf{P}_{t'}. \tag{37}$$

The matrix selecting the elements of $\mathbf{E}$ pertaining to the pairs of non-neighboring neigh-

[20] Let $\mathbf{X}$ 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 $x_{ij}y_{ij}$.

borhoods is

$$\mathbf{E}_{nN} = (\mathbf{J}_n - \mathbf{N}_t) \odot \mathbf{E}. \tag{38}$$

The decomposition of the change in the non-neighbor component of urban poverty concentration is obtained by replacing $\mathbf{P}_{t'}$ and $\mathbf{E}$ in equation (31) with $\mathbf{P}_{nN,t'}$ and $\mathbf{E}_{nN}$, respectively:

$$\Delta UP_{nN} = \frac{1}{2}tr\left(\mathbf{W}\mathbf{P}^T_{nN,t'}\right) + \frac{1}{2}tr\left(\mathbf{R}\lambda\mathbf{P}^T_{nN,t'}\right) + C\frac{1}{2}tr\left(\tilde{\mathbf{G}}_t\mathbf{E}^T_{nN}\right) = W_{nN} + R_{nN} + CE_{nN}. \tag{39}$$

Given equations (39) and (36), the spatial decomposition of the change in urban poverty concentration is

$$\Delta UP = W_N + W_{nN} + R_N + R_{nN} + C\left(E_N + E_{nN}\right). \tag{40}$$

■

B Testing for changes in urban poverty and its components

The significance of the decomposition in Corollary 2 is conditional on the statistical significance of the spatial component of urban poverty. To verify this point, it is sufficient to test the hypothesis that urban poverty is randomly distributes across the neighborhoods of the city, versus an unrestricted alternative of spatial association in poverty rates distributions. If the null hypothesis is rejected in each configuration (in year t and t'), the spatial decomposition of changes in urban poverty changes should bear meaningful information about spatial patterns of concentrated poverty.

We use the Rey and Smith nonparametric test (Rey and Smith 2013) for detecting spatial association in urban poverty exploiting the result in Theorem 1. For a given contingency matrix $\mathbf{N}$, we decompose total urban poverty in configuration $\mathcal{A}$ ($UP(\mathcal{A} = G(\mathcal{A}))$ into a component accounting for urban poverty among close neighborhoods ($G_N(\mathcal{A})$) and a component accounting for urban poverty among neighborhoods not related by spatial proximity ($G_{nN}(\mathcal{A})$). This decomposition gives $UP(\mathcal{A}) = G_N(\mathcal{A}) + G_{nN}(\mathcal{A})$. The test statistic is the ratio (Rey and Smith 2013):

$$SG = \frac{G_{nN}(\mathcal{A})}{G(\mathcal{A})},$$

which is the share of urban poverty explained by the disparities between non-neighboring census tracts. Inference on SG under the null hypothesis that the spatial configuration of neighborhood is irrelevant for assessing urban inequality changes is based on random spatial permutations of the observed poverty rates with their population shares across actual neighborhoods in order to simulate spatial randomness. Each permutation randomly re-assigns poverty rates with their population shares to census tracts. Thus, for M random spatial permutations, M random maps are obtained. For each map m, with $m = 1, ..., M$, the statistic SG_m is calculated. Define $\hat{SG}$ as the observed value of the test statistic SG, we define the indicator $\hat{M} = \sum_{m=1}^{M} \mathbf{1}[SG_m \geq \hat{SG}]$ be the number of the M random permutations producing SG values that exceed or are equal to the observed value $\hat{AG}$. A one-tailed pseudo significance level for the observed test statistic is:

$$p(SG) = \frac{1 + \hat{M}}{1 + M}.$$

The pseudo p-value $p(SG)$ is obtained by comparing $\hat{SG}$ to the distribution constructed under the null hypothesis that poverty rates with their population shares are randomly distributed between census tracts.

We use this statistics to verify the significance of the spatial decomposition in each of the base years used to assess urban poverty changes. Observing pseudo p-values less than 0.05 for a reasonably large number of permutations for at least one year would signal rejection of the null hypothesis that the spatial configuration of concentrated poverty does not affect urban poverty. This would call for considering the spatial decomposition of urban poverty changes in Corollary 2.

C Additional results

Table 2: decomposition of changes in urban poverty concentration.

period	component	$G_{t'}$	G_t	ΔG	W	R	E	C	D
Los Angeles									
period	component	$G_{t'}$	G_t	ΔG	W	R	E	C	D
80-90	N	0.11461	0.10991	0.00470	-0.00071	0.01189	-0.00706	0.91718	-0.00648
	nN	0.30602	0.30090	0.00512	-0.00619	0.02616	-0.01619	0.91718	-0.01485
	total	0.42063	0.41082	0.00982	-0.00691	0.03805	-0.02325	0.91718	-0.02133
90-00	N	0.10632	0.11337	-0.00705	-0.00157	0.01109	-0.01974	0.83985	-0.01658
	nN	0.29075	0.30726	-0.01651	0.00050	0.02513	-0.05018	0.83985	-0.04215
	total	0.39707	0.42063	-0.02357	-0.00106	0.03622	-0.06992	0.83985	-0.05873
00-14	N	0.10010	0.10557	-0.00546	0.00017	0.01355	-0.02064	0.92940	-0.01918
	nN	0.27517	0.29150	-0.01633	0.00331	0.03147	-0.05499	0.92940	-0.05111
	total	0.37527	0.39707	-0.02179	0.00348	0.04502	-0.07563	0.92940	-0.07029
Chicago									
period	component	$G_{t'}$	G_t	ΔG	W	R	E	C	D
80-90	N	0.08086	0.07874	0.00211	-0.00063	0.00522	-0.00286	0.86459	-0.00247
	nN	0.48125	0.47047	0.01078	-0.01077	0.03242	-0.01257	0.86459	-0.01087
	total	0.56211	0.54921	0.01290	-0.01140	0.03764	-0.01542	0.86459	-0.01334
90-00	N	0.06969	0.07902	-0.00933	-0.00169	0.00523	-0.01299	0.99123	-0.01288
	nN	0.44149	0.48309	-0.04160	0.00096	0.02789	-0.07107	0.99123	-0.07045
	total	0.51118	0.56211	-0.05093	-0.00072	0.03312	-0.08406	0.99123	-0.08332
00-14	N	0.05382	0.06860	-0.01477	-0.00292	0.00796	-0.02905	0.68242	-0.01982
	nN	0.38760	0.44258	-0.05498	0.00641	0.05099	-0.16468	0.68242	-0.11238
	total	0.44143	0.51118	-0.06976	0.00349	0.05895	-0.19373	0.68242	-0.13220
Dallas									
period	component	$G_{t'}$	G_t	ΔG	W	R	E	C	D
80-90	N	0.05988	0.05770	0.00218	0.00114	0.00845	-0.01034	0.71674	-0.00741
	nN	0.41259	0.39889	0.01369	0.01110	0.05103	-0.06758	0.71674	-0.04844
	total	0.47247	0.45660	0.01587	0.01224	0.05948	-0.07792	0.71674	-0.05585
90-00	N	0.06033	0.06307	-0.00274	0.00161	0.00734	-0.01156	1.01135	-0.01169
	nN	0.38628	0.40939	-0.02311	0.01266	0.03814	-0.07309	1.01135	-0.07391
	total	0.44661	0.47247	-0.02586	0.01427	0.04548	-0.08465	1.01135	-0.08561
00-14	N	0.05764	0.05721	0.00043	0.00295	0.01005	-0.01949	0.64462	-0.01256
	nN	0.37693	0.38941	-0.01247	0.01984	0.05592	-0.13688	0.64462	-0.08824
	total	0.43457	0.44661	-0.01204	0.02279	0.06597	-0.15637	0.64462	-0.10080
Houston									
period	component	$G_{t'}$	G_t	ΔG	W	R	E	C	D
80-90	N	0.09102	0.09329	-0.00227	0.00483	0.01445	-0.03713	0.58055	-0.02155
	nN	0.33113	0.34032	-0.00919	0.01653	0.05303	-0.13565	0.58055	-0.07875
	total	0.42215	0.43361	-0.01146	0.02136	0.06748	-0.17278	0.58055	-0.10031
90-00	N	0.08649	0.09079	-0.00430	0.00336	0.00738	-0.01466	1.02616	-0.01505
	nN	0.32326	0.33135	-0.00809	0.01366	0.02881	-0.04928	1.02616	-0.05057
	total	0.40975	0.42215	-0.01239	0.01703	0.03619	-0.06394	1.02616	-0.06561
00-14	N	0.08401	0.08469	-0.00068	0.00485	0.01335	-0.02528	0.74680	-0.01888
	nN	0.32334	0.32506	-0.00172	0.01445	0.05133	-0.09039	0.74680	-0.06750
	total	0.40735	0.40975	-0.00240	0.01930	0.06468	-0.11567	0.74680	-0.08638
New York									
80-90	N	0.19171	0.18353	0.00818	-0.00209	0.01386	-0.00323	1.10718	-0.00358
	nN	0.34283	0.32255	0.02028	-0.00028	0.02365	-0.00279	1.10718	-0.00309
	total	0.53454	0.50608	0.02846	-0.00237	0.03750	-0.00602	1.10718	-0.00667
90-00	N	0.18495	0.20161	-0.01666	-0.00081	0.01270	-0.03258	0.87642	-0.02855
	nN	0.30423	0.33293	-0.02870	-0.00323	0.02022	-0.05213	0.87642	-0.04569
	total	0.48918	0.53454	-0.04536	-0.00404	0.03292	-0.08471	0.87642	-0.07424
00-14	N	0.17059	0.17928	-0.00869	0.00028	0.01806	-0.02843	0.95085	-0.02703
	nN	0.29324	0.30990	-0.01666	0.00101	0.03086	-0.05104	0.95085	-0.04854
	total	0.46382	0.48918	-0.02535	0.00129	0.04893	-0.07947	0.95085	-0.07557