Supplemental appendix for online publication

Additional material and robustness checks for "Inference for the neighborhood inequality index"

Francesco Andreoli* Eugenio Peluso

University of Verona and LISER

June 2020

Contents

*Corresponding author. Andreoli and Peluso are both at the Department of Economics, University of Verona, Via Cantarane 24, 37129 Verona, Italy, and the Luxembourg Institute of Socio-Economic Research, LISER, MSH, 11 Porte des Sciences, L-4366 Esch-sur-Alzette/Belval Campus, Luxembourg. E-mails: francesco.andreoli@liser.lu and eugenio.peluso@liser.lu.

Reading note

This appendix to Andreoli and Peluso (2020) delivers additional results that are useful for implementing the Neighborhood Inequality index and its variance bounds empirically. This appendix refers to equations, tables and figures appearing in the reference paper with a "*", using the same numbering sequence appearing therein.

A Empirical estimators of the NI index and its variance bounds

Consider a sample of size n. Income realizations are denoted y_i, with $i = 1, \ldots, n$. The income vector $\mathbf{y} = (y_1, \ldots, y_n)$ is a draw from the spatial random process $\{Y_s : s \in \mathcal{S}\}$ distributed as $\mathcal{F}_S$, where a location s identifies a precise point on a map. Information about location (latitude and longitude) of an observation i is denoted by $s_i \in \mathcal{S}$. Distance measures between locations can be easily constructed based on the geodesic formula. Furthermore, observed incomes are associated with weights $w_i \geq 0$ and are indexed according to the sample units, with $w = \sum_i w_i$. It is often the case that the sample weights give the inverse probability of selection of an observation from the population. Sample weights may also incorporate complex survey design structure. In many applications, it is natural to assume $w_i = 1/n$.

The mean income within an individual neighborhood of size d, denoted μ_{id}, is estimated by $\widehat{\mu}_{id} = \sum_{j=1}^{n} \hat{w}_j y_j$ where

$$\hat{w}_j := \frac{w_j \cdot \mathbf{1}(||s_i - s_j|| \leq d)}{\sum_j w_j \cdot \mathbf{1}(||s_i - s_j|| \leq d)}$$

so that $\sum_j \hat{w}_j = 1$, and $\mathbf{1}(.)$ is the indicator function. The estimator of the average neighborhood mean income is instead $\hat{\mu}_d = \sum_{i=1}^{n} \frac{w_i}{w} \widehat{\mu}_{id}$. The estimator of the NI index, denoted $\hat{NI}(\mathbf{y}, d)$, is the sample weighted average of the mean absolute deviation of the income realization in location s_i from the income realization in any other location s_j such that $||s_i - s_j|| \leq d$. Formally

$$\hat{NI}(\mathbf{y}, d) = \sum_{i=1}^{n} \frac{w_i}{w} \frac{1}{2\hat{\mu}_{id}} \sum_{j=1}^{n} \hat{w}_j \, |y_i - y_j|,$$

where $\hat{w}_j$ is defined as above.

The estimation is conditional on d, which is a parameter under control of the researcher. The distance cutoff d is conventionally reported in miles and is meant to capture a continuous measure of the size of an individual neighborhood. In empirical applications, one can estimate as many values of d as there are pairs of observations in distinct locations on the maps. The value of d is generally bounded by the population under analysis: if the focus is on neighborhood inequality within a city, the definition of the urban area delimits the space where interactions take place, thus providing a bound for d. For computational reasons, however, the NI index and its SE are estimated only for a finite number of distance cutoffs, identifying intervals of fixed length. The maximum number of cutoffs indicates the point at which distance between observations is large enough that the NI index converges to the Gini index and its SE is constant. For a given neighborhood of size d, we partition the distance interval $[0, d]$, defining the size of the individual neighborhood, into K intervals $d_0, d_1, \ldots, d_K$ of equal size, with $d_0 = 0$. We always use d_k to denote the distance between any pair of observations i and j located at distance $d_{k-1} < ||s_i - s_j|| \leq d_k$ one from the other. The pairs $(d_k, \hat{NI}(\mathbf{y}, d_k))$ for any $k = 1, \ldots, K$ can be hence plotted on a graph. The curves resulting by linearly interpolating these points are the empirical equivalent of the neighborhood inequality curves.

A plug-in estimator for the asymptotic standard error of the NI indices can be derived under the assumptions listed in the previous sections. The SE estimator crucially depends on four components: (i) the consistent estimator for the average $\tilde{\mu}$, denoted $\hat{\mu}$, which

coincides with the sample average; (ii) the consistent estimator for variance σ^2, denoted $\hat{\sigma}^2$, which is given by the sample variance; (iii) the consistent estimator for the variogram; (iv) the estimator of the weighting schemes.

Empirical estimators $\hat{\mu}$ and $\hat{\sigma}^2$ are standard. The robust non-parametric estimator of the variogram proposed by Cressie and Hawkins (1980) can be used to assess the pattern of spatial dependency of georeferenced data on income realizations. The empirical variogram is defined for given distance ranges, meaning that it produces a measure of spatial dependence among observations that are located exactly at a given distance range one from the others. We use $b = 1, \ldots, B$ to partition the empirical distance range between any given pair of locations into equally spaced lags. Then, we estimate the variogram on each of these lags. This means that $2\gamma(b)$ refers to the correlation between incomes placed at distance lags of exactly b intervals, each of size d/B.

It is understood that the size of the sample is large compared to B, in the sense that the sampling rate per unit area remains constant when the partition into lags becomes finer. This assumption allows to estimate a non-parametric version of the variogram at every distance cutoff. The normality assumption is central for using variograms and linear predictions from the underlying model (see also Diggle, Tawn and Moyeed 1998). The assumption may be violated when dealing with income due to the presence of heavy tails and outliers. Following Cressie (1991), we use an empirical variogram estimator that is robust with respect to contamination from outliers and sensible to tails. We then use weighted least squares to fit a theoretical variogram model to the empirical variogram estimates we obtain from the data. The theoretical model is a continuous parametric function mapping distance into the corresponding variogram level. In the application, we adopt the spherical (semi)variogram model (see Cressie 1985), denoted $\gamma(h) = \alpha + \beta(3/2 \min\{h/a, 1\} - 0.5 \min\{h/a, 1\})$, where $\alpha,\ \beta$ are parameters to be estimated and a is the so-called range level: beyond distance a, the random variables Y_{s+h} and Y_s with $h > a$ are assumed to be spatially uncorrelated. The variogram satisfies the condition $\gamma(0) \to 0$ and $\gamma(a) = \sigma^2$. The maximum number of intervals B is set so that $d = 2a$. The estimated parameters are then used to draw predictions for the estimator $2\hat{\gamma}$ of the variogram at

each distance cutoff.[1] The predictions are then plugged into the SE estimator of the NI index.

Finally, SE estimation requires to produce reliable estimators of the weights ω. These are non-parametrically identified from the formulas provided above. In some cases, however, computation of the exact weights requires several iterations across observations. The overall computation time thus increases exponentially in the number of observations and the procedure becomes quickly unfeasible. We propose alternative, feasible estimators for these weights, denoted $\hat{\omega}$, that are expressed as linear averages. The computational time is, nevertheless, quadratic in the number of observations as it requires to construct a routine that first computes weights estimators for each observation separately, and then averages across all observations at given distance cutoffs.

We consider here only the weights that appear in the estimators $\hat{SE}_d$ in (5*) that cannot be directly inferred (i.e., are computationally unfeasible) from observed weights. For a given observation i, define $w(b,i) = \sum_{j \in d_{bi}} w_j$ for any ring $b = 1, \ldots, B_d, \ldots, B$ of radius d_b the weight associated with income realizations that are exactly located b lags away from i. Then, denote $w(d,i) = \sum_{j \in d_i} w_j = \sum_{b=1}^{B_d} w(b,i)$. Consider the following estimators

$$\text{for (9*)} \quad : \quad \hat{\omega}(m,b,b',d) = \sum_i \frac{w_i}{w} \frac{w(b,i)}{w(d,i)} \frac{w(m,i)}{w} \frac{w(m+b',i)}{w(m+d,i)};$$

$$\text{for (11*)} \quad : \quad \hat{\omega}(m,b',d) = \sum_i \frac{w_i}{w} \frac{w(m,i)}{w} \frac{w(b',i)}{w(d,i)}.$$

To compute these weights, each observation i has to be first assigned with the total weight $w(b,i)$ of those observations $j \neq i$ that are located exactly at distance b from i. Then, $\hat{\omega}(m,b,b',d)$ and $\hat{\omega}(m,b',d)$ are obtained by averaging these weights across i's. The key feature of these estimators is that weights of observations occurring at distance b' from an observation located at distance m from i are estimated by averaging across all observations the relative weight of observations at distance $m+b'$ from i.

[1] Cressie (1985) has shown that this methodology leads to consistent estimates of the true variogram function under the stationarity assumptions mentioned above.

B Monte Carlo study: Additional details

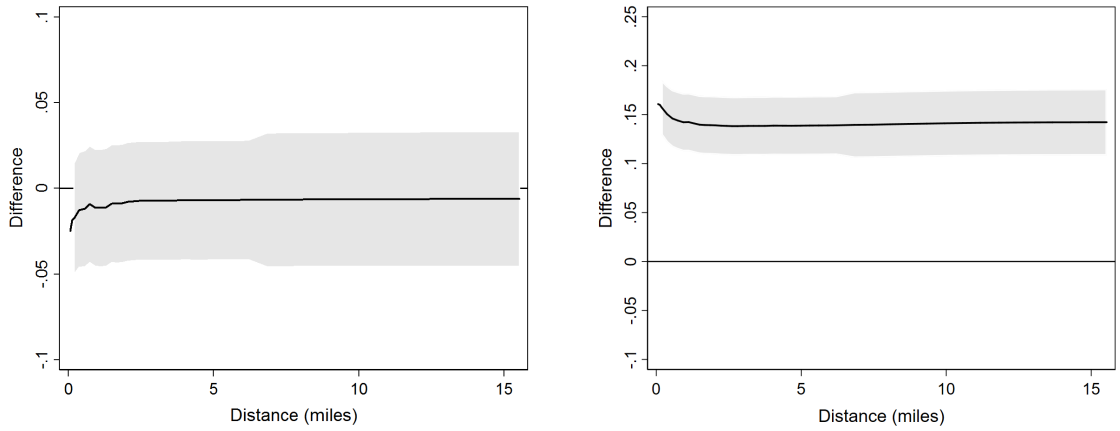
This section describes the underpinnings of the Monte Carlo study presented in the paper. The study reports simulated size and power for tests of differences of NI indices estimated at pre-determined distance cutoffs on samples of variable size n (n = 1000, 2000, 5000, 8000 observations), each drawn from distinct known distributions. The distributions are calibrated to represent the actual distribution of gross equivalent household income in Chicago IL in 2014, obtained from the Census Bureau's American Community Survey data, 2010-2014 module. We compare the actual distribution with counterfactual distributions obtained by applying suitable transformations to the actual ACS 2010-2014 data, so that these distributions can be clearly ordered in terms of NI curves dominance. Then, we use moments of these population distributions to identify moments of the income data generating processes adopted in the simulation study.

The first distribution $\mathcal{F}_0$ represents the spatial income distribution in Chicago, 2014. This distribution has mean income $\mu_0 = \$53,456$, standard deviation $\sigma_0 = \$55,310$ and spatial covariance structure $cov(s, v)$ across pairs of locations at distance h one from the other. The spatial covariance function, estimating spatial dependence in the population model, is characterize by the variogram $\gamma_0(.)$, so that $cov(s, v) = \sigma_0^2 - \gamma_0(h)$. Interpolating the spherical model for the (semi)variogram function to the data yields parameters estimates $\alpha_0 = -327,203$, $\beta_0 = 21.14$ and range level $a = 10$ miles.

We produce two counterfactual population distributions $\mathcal{F}_1$ and $\mathcal{F}_2$ from the same data. The distribution $\mathcal{F}_1$ is obtained by adding noise to $\mathcal{F}_0$, so that $y_1 = y_0 + \varepsilon$ for $y_1 \sim \mathcal{F}_1$, $y_0 \sim \mathcal{F}_0$ and $\varepsilon \sim N(0, 6118.44)$, where the variance term of idiosyncratic disturbances is half a million time smaller than σ_0^2. This counterfactual distribution displays similar patterns of neighborhood inequality as $\mathcal{F}_0$. The null hypothesis H_0^3: $NI(\mathcal{F}_0, d) = NI(\mathcal{F}_1, d)$ for at least some d, cannot be rejected, as shown in panel (a) of figure 3. This new population distribution has expectation $\mu_1 = \mu_0$, standard error $\sigma_1 = \$55,631 > \sigma_0$ and variogram $\gamma_1(.)$ with parameters $\alpha_1 = -69,660$ and $\beta_1 = 21.19$.

The second counterfactual distribution $\mathcal{F}_2$ is designed in a way that its NI curve lies always below that of $\mathcal{F}_0$. This distribution is obtained by simulating the effect of a

Figure 1: Neighborhood inequality in Chicago, IL, 2014, versus two counterfactual distributions



(a) $NI(\mathcal{F}_0, d) - NI(\mathcal{F}_1, d)$ (b) $NI(\mathcal{F}_0, d) - NI(\mathcal{F}_2, d)$

Note: Author analysis of US Census and ACS data. Confidence intervals are at 95% level.

redistributive linear income tax scheme applied to incomes distributed as $\mathcal{F}_0$. Andreoli and Peluso (2018) have demonstrated that only a basic income flat tax scheme guarantees that $\mathcal{F}_2$ dominates $\mathcal{F}_0$ in terms of NI curves. We hence use the transformation $y_2 = (1-t)y_0+m$, for $y_0 \sim \mathcal{F}_0$, a flat tax rate $t = 0.3$ and basic income $m = 0.3\mu_0$. This counterfactual distribution displays different patterns of neighborhood inequality compared to $\mathcal{F}_0$. The null hypothesis H_0^3: $NI(\mathcal{F}_0, d) = NI(\mathcal{F}_1, d)$ for at least some d is clearly rejected in favor of a dominance alternative, as shown in panel (b) of figure 3. This new population distribution has expectation $\mu_2 = \mu_0$, standard error $\sigma_2 = \$38,716 < \sigma_0$ and variogram $\gamma_2(.)$ with parameters $\alpha_2 = -158,424.5$ and $\beta_2 = 20.43$.

The simulation study is based on models for the income process, denoted $\mathbf{Y}_f^n$ for $f = 0, 1, 2$, that replicate the population distributions $\mathcal{F}_0$, $\mathcal{F}_1$ and $\mathcal{F}_2$, respectively. As before, the income process is a collection of random variables indexed by n, a parameter controlled within the experiment, and defined over the random field $\mathcal{S}_n$. The first concern is to replicate the spatial structure of the data and construct a random field $\mathcal{S}_n$ that is representative of the map of Chicago in terms of distance scale and population density. To do so, we draw a random field $\mathcal{S}_n$ (reporting information about latitude and longitude of

each of the n locations) directly from the ACS 2010-2014 map of Chicago, by sampling n locations without replacement. This procedure should guarantee that the structure of ACS data for Chicago is always reflected in the outcomes of the simulation. These sampled locations are stored in a separate file for each n and used throughout the simulation experiment. Results will be conditional to the random field $\mathcal{S}_n$.[2]

The second concern is to model the spatial income process $\mathbf{Y}_f^n$ so that it represents income variability and spatial association underlying the population distributions $\mathcal{F}_f$. Given the random field $\mathcal{S}_n$, we maintain the assumption that the spatial income process is stationary and the Gaussian hypothesis, implying that the process is fully characterized by known moments of the population distribution, so that $\mathbf{Y}_f^n \sim (\mu_f, \sigma_f^2, \gamma_f(.))$ for $f = 0, 1, 2$. The Monte Carlo experiment consists in randomly drawing realizations from $\mathbf{Y}_f^n$, each denoted $\mathbf{y}_{f,r}^n$ with $r = 1, \ldots, 200$, and assessing for each draw r if a certain null hypothesis about dominance in NI curves can or cannot be rejected, provided that the actual pattern of dominance in the populations is known. The SE approximations discussed in Section 3.1* are then adopted to conclude about acceptance/rejections of the relevant null hypothesis. The decision outcome is registered with an indicator, which is then averaged across the 200 replicas to simulate size and power of the tests.

Each draw $\mathbf{y}_{f,r}^n$ from the spatial income process $\mathbf{Y}_f^n$, $f = 0, 1, 2$, should be representative of the degree of spatial association in the underlying population distribution $\mathcal{F}_f$. Coherently with previous assumptions, the spatial association between any pair of locations s, v on the random field $\mathcal{S}_n$ at geographic distance h (in miles) is provided by the covariance term $c_{s,v} = \sigma_f^2 - \gamma_f(h)$. The empirical estimates of the moments $(\mu_f, \sigma_f^2, \gamma_f(.))$ from the population distribution $\mathcal{F}_f$ identify the covariance matrix $\mathbf{C}_f$ of the spatial income process, with $\mathbf{C}_f = \{c_{s,v}\}$, $s = 1, \ldots, n$, $v = 1, \ldots, n$ and $c_{s,s} = \sigma_f^2$. We use decomposition methods to factorize the covariance matrix as $\mathbf{C}_f = \mathbf{D}_f \cdot \mathbf{D}_f'$, where $\mathbf{D}_f$ is a lower triangular matrix of size $n \times n$. This matrix conveys the information about the spatial association and variability in the population.[3] Each replica r of a distribution f

[2]The extracted coordinates of the random fields, alongside parameter estimates and replication code for this Monte Carlo study, are made available on the author web page.

[3]In some cases, the covariance matrix $\mathbf{C}_f$ may not be positive semi-definite, implying that exact symmetric decompositions are not available. We use approximations based on the spectral theorem, as

(of size n) is then obtained as

$$\mathbf{y}_{f,r}^{n} = \mu_f \mathbf{e}_n \ + \ \mathbf{D}_f \cdot \boldsymbol{\nu}_r,$$

where $\mathbf{e}_n$ is a $n \times 1$ vector with all elements equal to one and $\boldsymbol{\nu}_r$ is a $n \times 1$ vector of standard normal distributed i.i.d. innovations. Throughout all replicas, values of the NI index and of the SE approximations can be meaningfully computed only at some distance cutoffs. For samples of size $n = 2000,\ 5000,\ 8000$, distance cutoffs are set at approximately a third of a mile distance range increments within the first 2 miles, and then at increasing increments within the next 12 miles (at 19 miles range the NI index converges to citywide inequality). For the sample of size $n = 1000$, distance thresholds within 1 and 19 miles are set by looking at increments of three quarters of a mile exclusively. $H_0^3(d)$ is tested at each distance cutoff. The null hypothesis of the type H_0^3 is tested instead by looking at all distance cutoffs.

C Robustness check 1: Normality

The variance bound estimator in Andreoli and Peluso (2020) is derived under the normality assumption. As noted in Cressie (1991) and Diggle, Zheng and Durr (2005), it is often (albeit sometime not explicitly) assumed in geostatistics literature that the data are obtained from a gaussian model:

$$Y_i = \mu + X(s_i) + \varepsilon_i,$$

where μ is non stochastic and may depend on observable characteristics of each unit $i = 1, ..., n$, $X(s_i)$ is a stationary Gaussian process with known means and variances and spatial covariances $Cov[X(s), X(s')]$ for locations $s \neq s'$, $s, s' \in \mathcal{S}$ that can be explic-

suggested in Bunch and Parlett (1971), to decompose $\mathbf{C}_f = \mathbf{X} \cdot diag(\mathbf{L}) \cdot \mathbf{X}'$, where $\mathbf{L}$ is a vector of eigenvalues and $\mathbf{X}$ collects the corresponding eigenvectors. We then set negative eigenvalues to zero to obtain $\mathbf{L}^*$ and produce an approximation of $\mathbf{C}_f$, denoted $\mathbf{C}_f^*$, such that $\mathbf{C}_f^* = \mathbf{X} \cdot \sqrt{diag(\mathbf{L}^*)} \cdot \sqrt{diag(\mathbf{L}^*)} \cdot \mathbf{X}'$. We finally apply a Q-R decomposition of $\mathbf{C}_f^*$ to obtain $\mathbf{C}_f^* = \mathbf{R}_f' \cdot \mathbf{R}_f$ where $\mathbf{D}_f = \mathbf{R}_f'$ is a lower-triangular matrix.

itly written as a function of the variogram, and where ε_i are i.i.d. zero-mean normally distributed disturbancies. Under this assumption, the spatial process is jointly normal conditional on the random field. This is the assumption of normality.

The assumption may be restrictive for describing the empirical patterns of some variables, such as equivalized housold income, which tend to have heavy tails and are reasonably bounded on the positive domain. Furthermore, the normality assumption may not hold in presence of outliers, inducing bias in the estimation of the variogram. Cressie and Hawkins (1980) propose estimators for the variogram that are roust with respect to these empirical issues. Such estimators are adopted in the main paper to produce robust estimates of the empirical variogram, which is then used to produce estimates of second order moments under the normality assumption.

The Gaussian assumption has no implications for the derivation of the variance bound in (5*), although it bears consequences for the empirical estimation of two of the SE components, namely $Var[\Delta_d]$ and $Cov[\Delta_d, \tilde{\mu}]$. Cressie (1991), p.81 and 137, suggests to use Gaussian methods after applying a marginal non-linear transformation to the data.[4] We follow this line and consider an alternative assumption to normality, that is that income data are log-normally distributed.

The log-normality assumption has been widely adopted in economics as a theoretical model for the distribution of income and consumption opportunities, which can be measured by income equivalized by family needs (as in the reference paper, see also Battistin, Blundell and Lewbel 2009).[5] Furthermore, there exists only one transformation, the exponential, that maps log-income data data that are normally distributed into dollar-income. This transformation can be used to produce unbiased empirical estimates of the variogram. Additionally, under log-normality, the transformed data are jointly normal, and relevant parameters can be simulated from the knowledge of first and second moments of the underlying distribution and of the variogram.

[4]Linear predictors, stemming from the normality assumption, are employed in spatial kriging. See also Diggle et al. (1998) for an alternative proposal using linear predictors in a more general distributional framework.

[5]The rationale for log-normality of consumption rests on the Gibrat's law, which states that income is the cumulation of idiosynratic shocks.

In this section, we assume that the dollar-income data follow the process $Z_i = exp\{Y_i\}$ and that Y_i, generating the log-income data, is normal conditional on the random field. Following Cressie (1991), we consider a log transformation of the data and then use log-transformed data to fit the appropriate variogram model. We implement the empirical procedure outlined in appendix A to these data to obtain estimates of the underlying empirical variogram. The variogram predictions are then used to estimate covariances of log-income realizations across locations, and hence estimate the SE of the NI index. Noticing that $ln(Z_i)$ is normal, one can use the first and second moments of this distributions to estimate μ, σ^2 and simulate the covariances $Cov[(Y_j, Y_i, Y_k, Y_\ell)]$ (that we use to estimate the term $Var[\Delta_d]$) and $Cov[(Y_j, Y_i, Y_\ell)]$ (that we use to estimate the term $Cov[\Delta_d, \tilde{\mu}]$).

Identification of the first component of (5*), $Var(\tilde{\mu})$, does not rely on normality assumption. The second component of (5*), $Var[\Delta_d]$, depends on two expectations terms that cannot be further simplified without assuming normality of the underlying spatial process. Assuming log-normality instead implies that the expectation $E[|exp\{Y_i\} - exp\{Y_j\}||exp\{Y_\ell\} - exp\{Y_k\}|]$ can be simulated from a large number S (with $S = 1,000$) of independent draws $(y_{1s}, y_{2s}, y_{3s}, y_{4s})$ with $s = 1, \ldots, S$, from the random vector (Y_j, Y_i, Y_k, Y_ℓ). The simulated expectation is a function of the variogram parameters m, b, b' and d and of σ^2. It is denoted $\theta(m, b, b', d, \sigma^2)$ and estimated as follows:

$$\theta(m, b, b', d, \sigma^2) = \frac{1}{S} \sum_{s=1}^{S} |exp\{y_{2s}\} - exp\{y_{1s}\}| \cdot |exp\{y_{4s}\} - exp\{y_{3s}\}|.$$

Furthermore, the expectation $E[|exp\{Y_i\} - exp\{Y_j\}|]$ can be simulated from the same set of independent draws. The simulated expectation is denoted $\theta(b, d, \sigma^2)$ and estimated as follows:

$$\theta(b, d, \sigma^2) = \frac{1}{S} \sum_{s=1}^{S} |exp\{y_{2s}\} - exp\{y_{1s}\}|.$$

Combining together, the estimator of the variance component of the NI standard error

writes:

$$Var[\Delta_d] = \sum_{m=1}^{B}\sum_{b=1}^{B_d}\sum_{b'=1}^{B_d}\omega(m,b,b',d)\theta(m,b,b',d,\sigma^2) - 4\left(\sum_{b}^{B_d}\omega(b,d)\theta(b,d,\sigma^2)\right)^2. \quad (1)$$

The wights are defined as in reference paper.

The third term of the SE formula, $Cov[\Delta_d, \tilde{\mu}]$, depends on the expectation $E[\min\{exp\{Y_j\}exp\{Y_\ell\} - exp\{Y_i\}exp\{Y_\ell\}, 0\}]$ under the log-normality assumption. It can also be estimated under the joint normality assumption of the vector (Y_i, Y_j, Y_ℓ) and simulated as in the main paper.

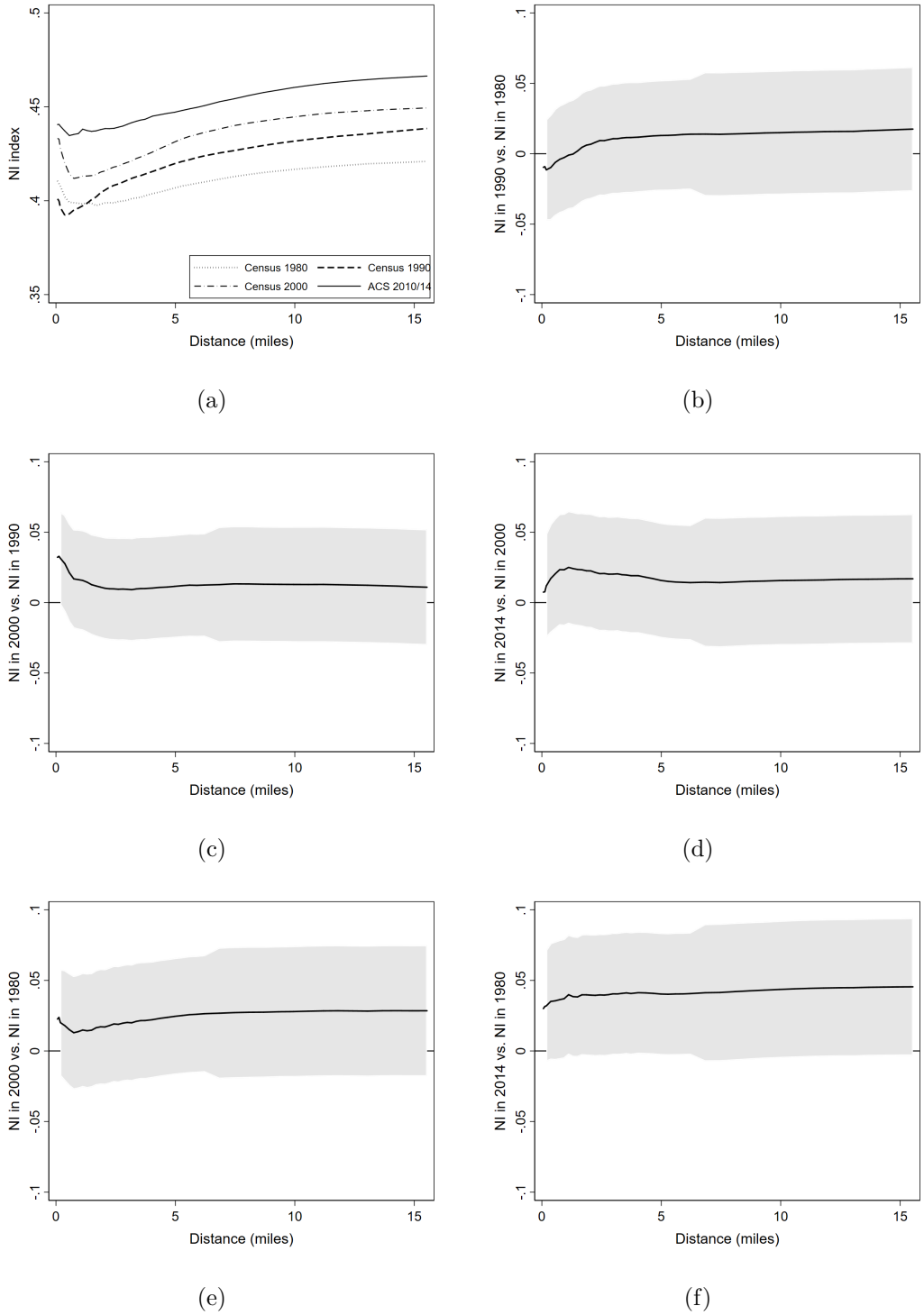
We apply the robust variogram estimator on dollar-income data to estimate (8*), while we use the estimators outlines above to produce estimates of $Var[\Delta_d]$ and of $Cov[\Delta_d, \tilde{\mu}]$, that we can use to draw SE bounds for the NI index at any distance d without relying on the normality assumption. Empirical counterparts are as in appendix A.

C.1 Empirical evidence

We maintain the log-normality assumption and use the newly derived SE estimators to infer about changes in neighborhood inequality in Chicago, IL over the period 1980-2014. The empirical application is based on the same data presented in section 5* of the main paper.

Compared to the results obtained under normality (Figure 2*), the new SE bounds for the NI index identified under log-normality are larger, suggesting a broader acceptance region for the null hypothesis that difference in NI curves are equal to zero. As Figure 2 shows, comparisons of the evolution of neighbrohood inequality in Chicago across decades reveal rising trends, albeit differences are not statistically significant at 95% confidence level. The cumulative effect of growth in neighborhood inequality over from 1980 to 2014 is positive albeit not statistically different from zero at 95% confidence level (panel f of Figure 2). The cumulative gap in NI curves turns out significant at 90% confidence level,

Figure 2: Trends in neighborhood inequality in Chicago, IL



Note: Author analysis of US Census and ACS data. Confidence intervals are at 95% level. SE estimates based on the log-normal distributional assumption.

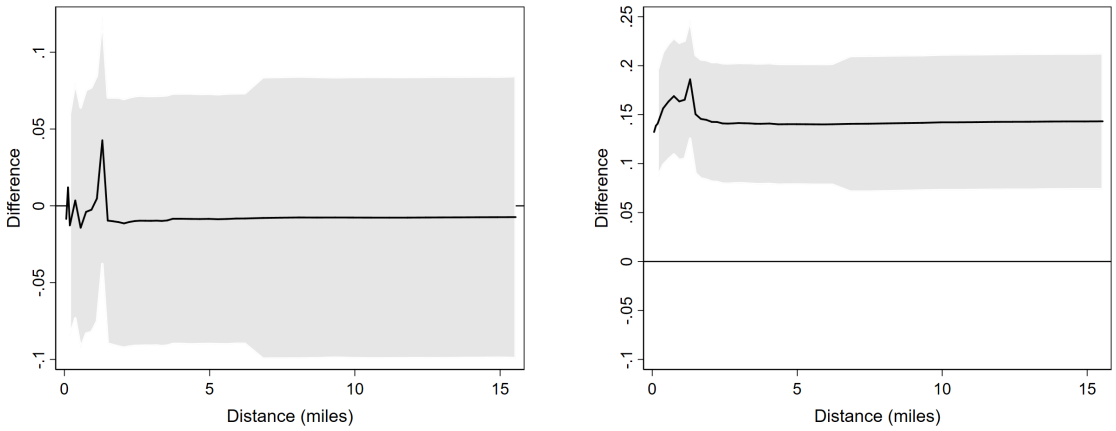
Years	Distance d in miles					
	0.4	1	2	3	5	12
Panel A: p-values for H_0^1						
1980	0.034	0.022	0.024	0.037	0.083	0.383
1990	0.000	0.000	0.000	0.000	0.001	0.054
2000	0.000	0.000	0.000	0.000	0.003	0.098
2014	0.001	0.001	0.002	0.004	0.012	0.188
Panel B: p-values for H_0^2						
1980	.	0.439	0.443	0.485	0.413	0.238
	0.000	*-0.003*	*-0.003*	*-0.001*	*0.005*	*0.017*
1990	.	0.411	0.213	0.123	0.052	0.010
	0.000	*0.004*	*0.013*	*0.020*	*0.027*	*0.042*
2000	.	0.344	0.416	0.475	0.275	0.092
	0.000	*-0.008*	*-0.004*	*0.001*	*0.011*	*0.027*
2014	.	0.474	0.475	0.417	0.318	0.125
	0.000	*-0.001*	*0.001*	*0.004*	*0.010*	*0.026*

Table 1: P-values (computed under the log-normality assumption) for null hypothesis of the type $H_0^1: \ NI(\mathbf{y}_t, d) = G(\mathbf{y}_t)$ and $H_0^2: \ NI(\mathbf{y}_t, d) = NI(\mathbf{y}_t, 0.4)$, with $t =$ 1980, 1990, 2000, 2014 and $G(\mathbf{y}_{1980}) = 0.434$, $G(\mathbf{y}_{1990}) = 0.461$, $G(\mathbf{y}_{2000}) = 0.473$, $G(\mathbf{y}_{2014}) = 0.486$. Differences in levels of the NI index are in italic.

reproducing the outcomes in Figure 2*.

Table 1 reports p-values (obtained under the log-normality assumption) for null hypothesis concerning the shape of the NI curves. Tests for the null hypothesis H_0^1 are in panel A. In line with results in Table 2*, we find evidence that the extent of neighborhood inequality is indistinguishable from the citywide Gini index only when individual neighborhoods are set to a relatively large size, above 5 miles radius. The results confirm that the normality assumption, compared to log-normality, is not driving the conclusions that neighborhood inequality is a relevant and growing phenomenon that is distinct from citywide inequality at small individual neighborhood scale. In panel B of the same table, we also report p-values for H_0^2 obtained under the log-normality assumption. Conclusions of the paper are unaffected when replacing normality with the log-normality assumption: at almost all distance thresholds below 3 miles considered in Table 1 we cannot reject that neighborhood inequality is equal to the level of neighborhood inequality observed in very small (less than half a mile) individual neighborhoods. NI estimates based on individual

Figure 3: Neighborhood inequality in Chicago, IL, 2014, versus two counterfactual distributions



(a) $NI(\mathcal{F}_0, d) - NI(\mathcal{F}_1, d)$ (b) $NI(\mathcal{F}_0, d) - NI(\mathcal{F}_2, d)$

Note: Author analysis of US Census and ACS data. Confidence intervals are at 95% level. SE estimates based on the log-normal distributional assumption.

neighborhoods of large size are statistically different. Compared to Table 2* in the main paper, Table 1 displays generally larger p-values, as a consequence of the more stringent rejection regions estimated under log-normality assumption.

C.2 Monte Carlo experiment

We check the robustness of the normality assumption under alternative distributional configurations. Building on appendix B (and adopting similar notation), we produce a Monte Carlo study using distributions $\mathcal{F}_0$ to represent the spatial distribution of income in Chicago, IL in 2014, and counterfactual distributions $\mathcal{F}_1$ and $\mathcal{F}_2$ as in the main paper. These distributions are assumed log-normal for two reasons: first, the assumption guarantees a good fit to income data and is related to normality with known transformations; second, the simulation routine developed under the normality assumption can be applied straightforwardly to the log-income data and then convert the simulated samples into dollar-incomes.

An empirical investigations of the available data for $\mathcal{F}_0$, $\mathcal{F}_1$ and $\mathcal{F}_2$ reveals patterns of dominance in NI curves that are similar to those reported in Figure 1* in the main

paper. In particular, Figure 3 shows evidence that $\mathcal{F}_1$ and $\mathcal{F}_0$ are indistinguishable from a neighborhood inequality perspective (i.e. H_0^3 cannot be rejected), while equality of NI curves is rejected when comparing distributions $\mathcal{F}_0$ and $\mathcal{F}_2$. Confidence intervals in both panels of Figure 3 are obtained under the log-normality assumption. As highlighted before, confidence interval bounds obtained under the log-normal assumption identify a larger acceptance region compared to what we obtain under the normality assumption.

The Monte Carlo experiment evaluates the size and power of the tests for dominance in NI curves that are based on the normality assumption (as we do in section 4* of the main paper), knowing that the data are generated from a joint log-normal distribution (differently from what we do in the paper, where data are also assumed jointly normal). Comparing size and power of these tests with those reported in Table 1* is informative about the impact of the normality assumption on the patters of acceptance and rejections of relevant null hypothesis on the spatial income inequality pattern.

Given $Z_f \sim \mathcal{F}_f$, $f = 0, 1, 2$, the random variable $Y_f = ln(Z_f)$ is normally distributed. We consider a population model $\mathbf{Y}_f^n = (Y_f^1, \ldots, Y_f^n)$ with $\mathbf{Y}_f^n \sim (\mu_f, \sigma_f^2, \gamma_f(.))$ for $f = 0, 1, 2$, where μ_f and σ_f^2 are the first and second moments and γ_f the variogram function of the normally distributed process (in log-incomes terms). The Monte Carlo experiment consists in drawing realizations from $\mathbf{Y}_f^n$, each denoted $\mathbf{y}_{f,r}^n$ with $r = 1, \ldots, 200$, that are representative of the extent of spatial association that we find in the population. The spatial association between any pair of locations s, v on the random field $\mathcal{S}_n$ at geographic distance h (in miles) is provided by the covariance term $c_{s,v} = \sigma_f^2 - \gamma_f(h)$. We proceed as in the main paper and use mean and sample variance of the log-income distribution of income Chicago 2014 to obtain estimates of μ_f and σ_f^2, all expressed in log-incomes. Furthermore, we use log-transformed data to fit variograms models and obtain estimators for the spatial covariance matrix $\mathbf{C}_f$. From these data we draw 200 replica samples $\mathbf{y}_{f,r}^n$ of size n and then convert simulated log-incomes into dollar-incomes applying the transformation $z_{f,r,i}^n = exp\{z_{f,r,i}^n\}$ for $i = 1, \ldots, n$, which gives the simulated sample $\mathbf{z}_{f,r}^n$. We consider the same parametrization of the simulation exercise as in the paper ($n = 1000$, 2000, 5000, 8000; distance cutoffs are set at approximately a third of a

n	Distance cutoffs (miles)									# Rej.	Rej.	Weak	Strong
	0.4	0.7	1	1.4	1.7	2	3	5	12				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Panel A : Size comparisons for the true null H_0^3 : $NI(\mathbf{z}_{0,r}^n, d) = NI(\mathbf{z}_{1,r}^n, d)$ for some d													
1000	.	0.00	.	0.00	.	0.00	0.00	0.00	0.00	0.0	0.00	.	.
2000	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0.00	.	.
5000	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.1	0.03	0.00	0.00
8000	0.00	0.00	0.03	0.00	0.04	0.00	0.00	0.00	0.03	0.4	0.11	0.03	0.00
Panel B : Power comparisons for the true alternative H_a^3 : $NI(\mathbf{z}_{0,r}^n, d) \geq NI(\mathbf{z}_{2,r}^n, d)$, $\forall d$													
1000	.	0.00	.	0.01	.	0.19	0.59	0.79	0.95	19.0	0.98	1.00	0.00
2000	0.00	0.00	0.62	0.69	0.99	0.94	1.00	1.00	1.00	32.2	1.00	1.00	0.00
5000	0.00	0.67	0.81	0.81	0.81	0.81	0.81	0.81	0.81	27.5	0.81	1.00	0.00
8000	0.31	0.91	0.93	0.93	0.93	0.93	0.93	0.93	0.93	31.7	0.93	1.00	0.00

Table 2: Monte Carlo simulations of the size and power of dominance tests for NI curves that are based on the NI index SE approximations. Data are simulated based on the log-normal joint distribution. SE are obtained under the normality assumption.

mile distance range increments within the first 2 miles, and then at increasing increments within the next 12 miles; $H_0^3(d)$ is tested at each distance cutoff; H_0^3 is tested instead by looking at all distance cutoffs).

Table 2 reports the size and power of the tests for $H_0^3(d)$ under different configurations and for different distance thresholds (d). Interpretation of the entries of the table are as for Table 1* in the main paper. In both Table 2 and Table 1* we produce results using SE for the NI index that employ the normality assumption. In Table 2, however, the underlying spatial process is log-normal. Comparing the two tables is informative about the relative performances of the SE bounds under normality assumption when data are non-normal.

Panel A of Table 2 reveals the estimated size for $H_0^3(d)$ at fixed distance thresholds, knowing that the null is true in the population. The size of the tests is virtually zero for samples of size smaller than 5000 units, irrespectively of the distance threshold considered. When the population spatial process is log-normal, the SE bounds estimated under the normality assumption are larger than the nominal size, implying a thinner rejection region for the null H_0^3: in small samples, the SE bounds obtained under normality produce test statistics that are over selective. The size of tests for $H_0^3(d)$ approaches the nominal 5% values only for sample sizes larger than 8000 units, whereas the upper bound size of joint tests for H_0^3 (product of columns (11) and (12)) are of about 3.5%. Test sizes in Table 1*,

where SE bounds are obtained under normality and the spatial data generating process is Gaussian, are generally larger than those in Table 2, both for small and larger sample sizes. Under reasonable distributional assumptions, the SE bounds based on the normal approximation produce conservative tests statistics for the relevant null hypothesis. In empirical application, we are bound to accept null hypothesis about equality in NI curves even in cases where there is evidence of dominance in the population, compared to what we would do if the population data were normally distributed. The severity of this problem seems to substantially disappear in larger (8000 units) sample sizes when the population data are ssumed jointly normal. In panel B of the table, we report power levels of the tests, which is informative about the ability of the test to discriminate the alternatives. For samples of 5000 units or above, power is above 80%, at all distance ranges. These figures are uniformly larger than those estimated in Table 1*, indicating larger discriminatory power related to SE bounds under the Gaussian hypothesis.

D Robustness check 2: Edge effects on NI estimates and inference

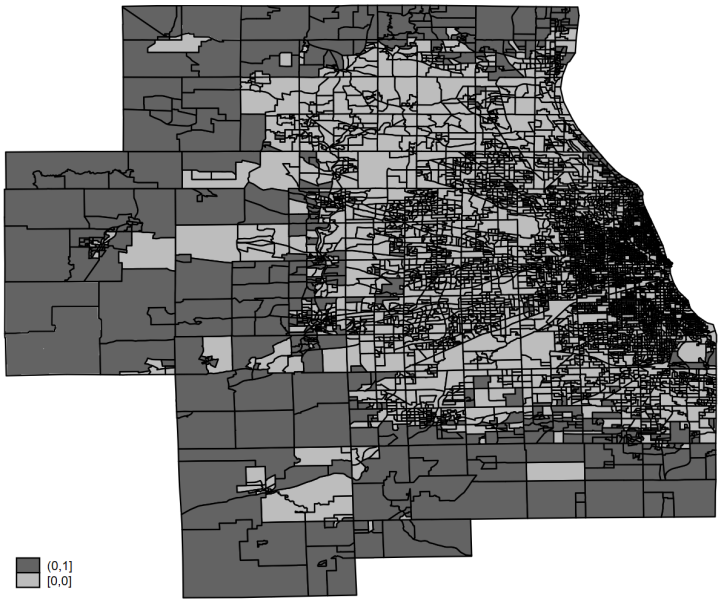
The second robustness check aims at assessing the implications of *edge effects* on the NI index estimator and on the SE bounds. Edge effects arise because a spatially continuous process is empirically constrained by a spatial organization in areal units, with well defined bounds. Edge effects attributable to administrative division of the urban territory, most likely occurring in MSA spanning across different counties and census tracts, may give rise to discontinuities in spatial association and generate biases in the estimation of spatial association measures, such as the variogram. Griffith (1985) and subsequent literature have produced techniques to mitigate the bias in estimation of spatial correlation measures. Griffith and Csillag (1993) discuss implications of edge effects for the variogram estimation. The considerations developed in these contributions can be adapted to our setting to produce consistent estimators for the NI index and for the SE bounds (based on the variogram). Furthermore, edge effects arising from arbitrary partitioning the urban

space do not seem to produce majors concern for the empirical evaluations carried out in the main paper, which is constrained by availability of empirical data at block group level. In fact, block groups do not define relevant administrative division for local taxation or supply of public goods that might affect differentially the location choices of rich and poor individuals. The empirical evaluations and simulations we produce in this section assume continuity of the spatial process.

Edge effects may also arise from the administrative restrictions imposed on the size and shape of the urban area considered for producing neighborhood inequality estimates. In neighbrohood inequality analysis, for instance, the relevant size of the urban environment may be defined on the basis of exogenous administrative criteria, such as the division in MSA or Commuting Zones operated by the Census Bureau. Albeit such definition of urban space is meaningful from a variety of perspectives, it defines a clear discontinuity at the border in the spatial distribution of incomes which may have an impact on the estimation of the NI index and its SE. The presence of administrative borders constraining the sample of analysis may have implications for the estimation of the NI index (and the variogram) at both large and small scale. In fact, boundary areas of the city define the distance threshold over which, by construction, all units are observed. In correspondence of such distance threshold, the NI index converges to the citywide Gini index by construction. Furthermore, observations located at the edges of the relevant urban area contribute to bias estimates of neighborhood inequality, insofar individual neighborhoods of these units only partially overlap with the territory of the city and indices Δ_i for people residing in these areas are biased. If these areas are densely populated, the effects of such limitations may induce relevant bias in NI estimates, as well as on the empirical variogram (and hence of the SE) even at small distance range. Finally, the presence of edges to the spatial sample distribution is inconsistent with the assumption that the variogram is isotropic (an assumption we make use of for identifying the SE bounds), insofar individual neighborhoods cannot span evenly across all directions.

The presence of edges in the urban area definition has an impact as well on the simulation study, which takes the structure of the random field as given. Xu and Dowd

Figure 4: The MSA of Chicago, IL. Edge block groups are in darker gray.



(2012) have supported, in the context of simulation studies, the possibility of using buffer zones (or guard areas) as a mean to mitigate the implications of the edge effects (see also Griffith 1983). We follow this line and consider a buffer zone in the estimation of the NI index and of the variogram. Figure 4 reports the 2014 map of the MSA of Chicago, IL and its block group division. We use such map to define borders of the metropolitan area under study in both the empirical evaluation and the Monte Carlo study. The buffer zone for Chicago is shaded in gray in Figure 4.[6]

D.1 Empirical assessment

Let $\mathcal{E}$ define the buffer zone of the city distribution $\mathcal{S}$. Edge effects arise from the contribution to the NI index of inequality estimated for income units located in $\mathcal{E}$. We consider producing estimates of the NI index and of its SE focusing only on observations located

[6]To identify the buffer area in a non ad-hoc fascion, we have first organized block groups by percentiles groups based on their position on the latitude and longitude scale (computed from their centroid coordinates). Each cell of the partition on the latitude or longitude scale gathers about 60 block groups, arranged north-south for fixed latitude and east-west for fixed longitude. Then, we have trimmed 3% of the locations at the extremes of these distributions to identify locations at the extremes of the map. These locations define the buffer area.

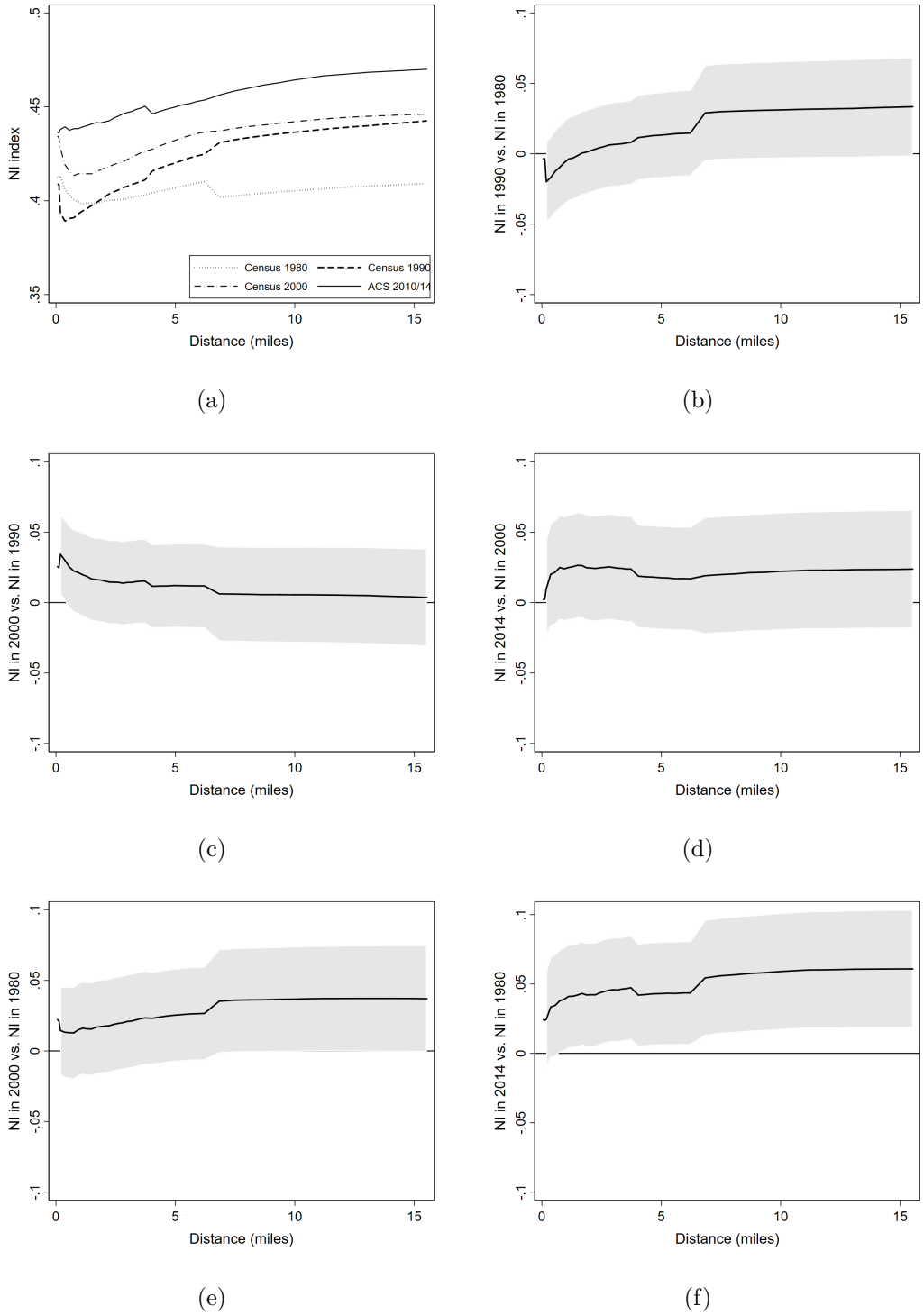
within the inner city area $\mathcal{S}/\mathcal{E}$. This is tantamount to redesign the shape of the relevant administrative boundaries to those of the MSA excluding $\mathcal{E}$. In the case of Chicago in Figure 4, this area corresponds to the light gray area. The procedure we use discriminates in an objective manner (which can be easily re-parametrized) the new sample n' in empirical evaluations as well as the new random field we consider in the simulation exercise.

The NI index and the variogram are calculated over the sample $\mathcal{S}/\mathcal{E}$ (with n' replacing n in the relative formulas). Besides, information on all units in the city (all locations in $\mathcal{S}$) are used to construct individual neighborhoods and estimate Δ_i for all $i \in \mathcal{S}/\mathcal{E}$, which may hence include individuals located in the $\mathcal{E}$ buffer zone. In the empirical exercises reported in this section, we disregard individual neighborhoods and inequality for all income units located in the dark gray area in Figure 4, but we use these incomes to estimate individual neighborhoods of individuals located in the light gray area in the same figure. The new empirical income distributions, obtained using the same buffer area across income years, is denoted ${}^{E}\mathbf{y}_t$.

We held locations in the buffer area identified in Figure 4 as fixed in the empirical assessment, and employ the normality assumption to derive the SE bounds of the NI index. Figure 5 displays the NI curves for years 1980, 1990, 2000 and 2014 (panel a) and year-by-year differences in NI curves with their confidence internal implied by the SE bounds estimated when considering the buffer zone. Results in the figure are remarkably similar to those obtained in the main text without controlling for edge effects adopting a buffer zone. Table 3 reports p-values for H_0^1 (comparisons of the NI index with the citywide Gini index) and H_0^2 (shape of the NI curve). Results are virtually indistinguishable from those in Table 2* in the main paper.

The empirical analysis does not provide evidence of edge effects, insofar the size of the NI index and of its SE, as well as the conclusions about a variety of null hypothesis concerning the shape and the temporal differences in NI curves, do not vary when we consider a buffer zone that is supposed to mitigate the implications of edge effects. We evaluate more carefully the implications of edge effects in a Monte Carlo study.

Figure 5: Trends in neighborhood inequality in Chicago, IL



Note: Author analysis of US Census and ACS data. Confidence intervals are at 95% level. Buffer zone defined as in Figure 4. SE bounds estimates based on the normality assumption.

Years	Distance d in miles					
	0.4	1	2	3	5	12
Panel A: p-values for H_0^1						
1980	0.016	0.003	0.004	0.007	0.024	0.047
1990	0.000	0.000	0.000	0.000	0.000	0.048
2000	0.000	0.000	0.000	0.000	0.000	0.030
2014	0.001	0.001	0.003	0.010	0.013	0.257
Panel B: p-values for H_0^2						
1980	.	0.347	0.357	0.403	0.480	0.479
	0.000	*-0.007*	*-0.006*	*-0.004*	*0.001*	*0.001*
1990	.	0.381	0.160	0.068	0.010	0.000
	0.000	*0.004*	*0.013*	*0.020*	*0.031*	*0.050*
2000	.	0.386	0.458	0.410	0.216	0.079
	0.000	*-0.005*	*-0.002*	*0.004*	*0.013*	*0.025*
2014	.	0.482	0.448	0.347	0.306	0.105
	0.000	*-0.001*	*0.003*	*0.008*	*0.010*	*0.028*

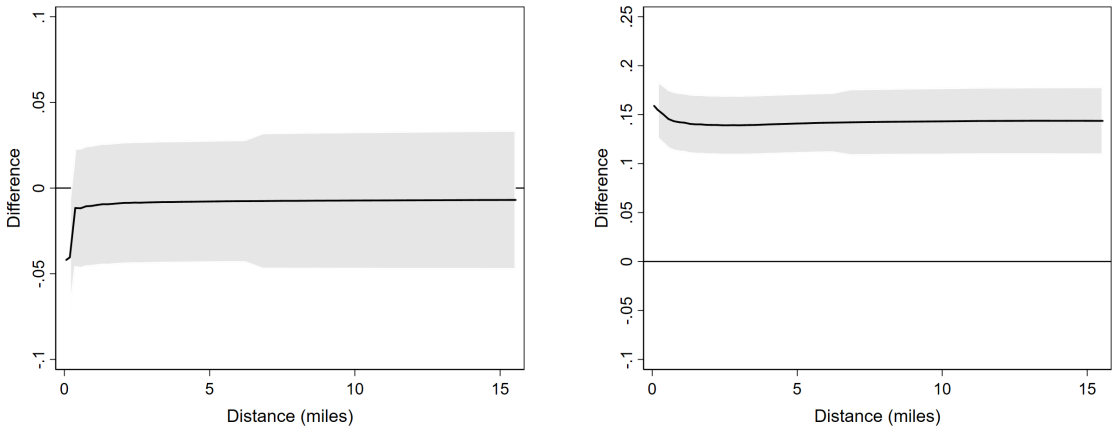
Table 3: P-values for null hypothesis of the type $H_0^1: \ NI({}^E\mathbf{y}_t, d) = G(\mathbf{y}_t)$ and $H_0^2: NI({}^E\mathbf{y}_t, d) = NI({}^E\mathbf{y}_t, 0.4)$, with $t = 1980,\ 1990,\ 2000,\ 2014$ and $G(\mathbf{y}_{1980}) = 0.434$, $G(\mathbf{y}_{1990}) = 0.461$, $G(\mathbf{y}_{2000}) = 0.473$, $G(\mathbf{y}_{2014}) = 0.486$. Differences in levels of the NI index are in italic.

D.2 Monte Carlo experiment

The Monte Carlo experiment replicates the setting described in the main paper. We consider the 2014 distribution of incomes in Chicago, which has population distribution $\mathcal{F}_0$, and obtain counterfactual distributions $\mathcal{F}_1$ (adding noise) and $\mathcal{F}_2$ (simulating the effects of an income redistribution scheme). We further consider the buffer zone $\mathcal{E}$ described above to estimate the underlying variograms and implied spatial correlations, considering only spatial units located in $\mathcal{S}/\mathcal{E}$. The distributions characterized by first and second moments of $\mathcal{F}_f$ and by the variograms estimates based on the buffer zone are denoted ${}^E\mathcal{F}_f,\ f = 0, 1, 2$ to idicate that the underlying spatial association already accounts for pontential edge effects.

We proceed as in the main paper and use first and second moments estimates together with spatial covariance matrix $\mathbf{C}_f$ to draw samples ${}^E\mathbf{y}_{f,r}^n$ (replica $r \leq 200$) from a collection of n (simulated sample size $n = 1000, 2000, 5000, 8000$) random variables that are jointly normally distributed, with means, averages and spatial association implied by ${}^E\mathcal{F}_f,\ f =$

Figure 6: Neighborhood inequality in Chicago, IL, 2014, versus two counterfactual distributions



(a) $NI({}^{E}\mathcal{F}_0, d) - NI({}^{E}\mathcal{F}_1, d)$ (b) $NI({}^{E}\mathcal{F}_0, d) - NI({}^{E}\mathcal{F}_2, d)$

Note: Author analysis of US Census and ACS data. Confidence intervals are at 95% level. SE bounds obtained under the normality assumption

$1, 2, 3$.[7] We derive NI index estimates and SE bounds using the procedure described in the previous section. We always produce SE bounds estimates under the normality assumption.

Figure 6 illustrates that distributions ${}^{E}\mathcal{F}_0$ and ${}^{E}\mathcal{F}_1$ are statistically indistinguishable in terms of neighborhood inequality they display (H_0^3 cannot be rejected at conventional significance levels), whereas distribution ${}^{E}\mathcal{F}_0$ displays significantly more neighborhood inequality than ${}^{E}\mathcal{F}_2$ at any distance range (H_0^3 is rejected). Table 4 reports size and power for tests based on the SE bounds (after accounting for the edge effects) for the null hypothesis $H_0^3(d)$, derived at selected distance thresholds d. By comparing the size and power estimates in Table 4 with those reported in Table 1*, we are able to draw conclusions about the statistical relevance of edge effects (accounted for in Table 4 by limiting estimates to the set of locations $\mathcal{S}/\mathcal{E}$).

Panel A of Table 4 displays the size of the tests using $H_0^3(d)$ under the null, when we know that $H_0^3(d)$ is true in the population. For simulated samples of size 1000 or 2000, the

[7]Each replica sample is based on the orginal map of Chicago. Hence, some predicted incomes will fall in the buffer area $\mathcal{E}$.

n	Distance cutoffs (miles)									# Rej.	Rej.	Weak	Strong
	0.4	0.7	1	1.4	1.7	2	3	5	12				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Panel A : Size comparisons for the true null H_0^3 **:** $NI(^E\mathbf{y}^n_{0,r}, d) = NI(^E\mathbf{y}^n_{1,r}, d)$, *for some* d													
1000	.	0.00	.	0.35	.	0.38	0.09	0.05	0.00	1.6	0.64	0.30	0.00
2000	0.00	0.00	0.34	0.26	0.22	0.13	0.08	0.02	0.00	1.8	0.65	0.61	0.00
5000	0.00	0.01	0.16	0.07	0.08	0.07	0.05	0.01	0.00	0.7	0.34	0.50	0.00
8000	0.00	0.00	0.01	0.06	0.01	0.05	0.01	0.03	0.01	0.5	0.14	0.55	0.00
Panel B : Power comparisons for the true alternative H_a^3 **:** $NI(^E\mathbf{y}^n_{0,r}, d) \geq NI(^E\mathbf{y}^n_{2,r}, d)$, $\forall d$													
1000	.	0.00	.	0.22	.	0.23	0.12	0.04	0.00	1.1	0.46	0.65	0.00
2000	0.00	0.00	0.33	0.27	0.32	0.14	0.08	0.03	0.00	1.9	0.64	0.95	0.00
5000	0.00	0.01	0.24	0.10	0.19	0.08	0.06	0.01	0.00	1.1	0.38	0.99	0.00
8000	0.00	0.09	0.58	0.34	0.53	0.36	0.32	0.32	0.24	7.0	0.52	0.98	0.00

Table 4: Monte Carlo simulations of the size and power of dominance tests for NI curves that are based on the NI index SE approximations.

size estimates in Table 4 are larger than those in the reference paper, uniformly across all distance ranges. In simulated samples of size 5000, the estimated size in Table 4 is closer to the nominal 5% value and improves the estimates in the reference paper (in particular for NI estimates based on individual neighborhoods of size smaller than 3 miles). For samples of size 8000, accounting for edge effects does not seem to produce improvements for the estimated sizes. The size for H_0^3 is estimated of about 7% (product of columns (11) ans (12)) for the largest samples. The simulation exercise suggests that edge effects have implications for the size of the tests of hypothesis based on the SE bounds in relatively small sample sizes. Accounting for a buffer zone in the simulation drastically reduces to test size to nominal values in samples of size 5000, and generally reduces the sizes of the tests in samples smaller than 2000 units (which remains nonetheless larger than the nominal size). In small samples edge effects are found to rise the probability of mistakenly rejecting equality in NI curves (against unrestricted alternative), thus making more likely statements about changes in neighborhood inequality. The simulation study also reveals that upper bounds for the null H_0^3 (obtained by the product of figures in columns (11) and (12)) are smaller than those in Table 1*. Size estimates converge to nominal 5% values (with upper bounds of 7%) when evaluations about neighborhood inequality are performed on larger samples of at least 8000 units, which is generally smaller than those available for assessing neighborhood inequality in American MSA.

Comparing panel B in Table 4 with the respective panel in Table 1*, we detect a reduction in the power of the tests for the relevant nulls when edge effects are accounted for. In the largest samples, tests power for H_0^3 is about 50%, i.e. 30% smaller than values registered in Table 1*. The NI index SE bounds seem to produce less discriminating tests for true alternatives in presence of edge effects at the boundaries of the city, the difference with Table 1* being stronger in small samples.

E Robustness check 3: Addressing empirical restrictions on the data

One last concern arises from the nature of the income data used in the application. The Census reports income data at the block group level (the finest urban space partition), thus splitting the urban space into small areal units. The centroid of each block group is genocoded and in the empirical application we make the (possibly restrictive) assumption that all income units observed in each areal units are located at the centroid. This assumption is convenient to map the Census publicly available data (STF3A files and ACS data tables), organized into areal units, into data that can be assumed generated by a spatially continuous process. This is a convenient simplification. Indeed, the population distribution of income from where the census and ACS are drawn is spatially continuous in nature. Different data sources, such as register data from Sweden, collect data at the exact location. Papers analyzing these income data assume (as we maintain in the paper) that the underlying income process is continuously defined in space (see for instance Östh, Clark and Malmberg 2015, Türk and Östh 2019).

As a robustness check for our application, we investigate whether introducing transformations of the original data that try to map the areal unit structure into a more realistic spatially continuous structure bear consequences for our estimates of NI and of its SE. We consider simulated distributions where observations are scattered uniformly within each block group areal unit. Based on this assumption, we have constructed artificial distributions of incomes ${}^{S}\mathbf{y}_t$ for $t = 1980, 1990, 2000, 2014$ by assuming that all income

Figure 7: Map of Chicago, IL, reporting the distribution of income locations

(a) At block group centroid (b) Scattered

observations (and their population weights) observed in each block group centroid are uniformly scattered within the block's boundaries.[8] We compute NI and its SE and assess hypothesis over the index using this newly obtained income dataset. Figure 7 displays the actual distribution of locations (block groups centroids) overlying the areal units partition of Chicago that we use in the rest of the paper (panel a) and the artificial distribution obtained after scattering the income observations evenly within block groups (panel b).

Figure 8 reports NI curves for the years considered, alongside differences in these curves across years and the confidence interval for these differences. Results obtained with the artificial distribution perfectly match those reported in the paper: neighborhood inequality rises over the period considered, albeit only the cumulative gap over 1980-2014 is significant at 95% level. Differently from the main paper, the NI curves obtained from the scattered data tend to display larger levels of neighborhood inequality even for individual neighborhoods of size less than 1 mile. Table 5 displays p-values of various hypothesis about the shapes of the NI curves depicted in Figure 8. Panel A) of the

[8]To do so, we have considered maxima and minima of latitude and longitude of each block group's edges (denoted $\overline{x}, \underline{x}$ for $x = lat., lon.$), assigned two distinct random number in $[0, 1]$ to each income unit within each block group (denoted $u_{lat}, u_{lon} \sim U(0, 1)$), and then associated a latitude and longitude to that unit proportionally (by the respective random numbers) to the maximal difference in latitude and longitude observed on the block group where this unit is located (that is, the new latitude and longitude of the artificial distributions are $\tilde{x} = \underline{x} + u_x(\overline{x} - \underline{x})$ for $x = lat., lon.$).

Years	Distance d in miles					
	0.4	1	2	3	5	12
Panel A: p-values for H_0^1						
1980	0.168	0.392	0.275	0.277	0.235	0.310
1990	0.005	0.040	0.077	0.073	0.117	0.150
2000	0.004	0.075	0.061	0.068	0.023	0.076
2014	0.002	0.056	0.083	0.080	0.124	0.218
Panel B: p-values for H_0^2						
1980	.	0.357	0.421	0.421	0.446	0.434
	0.000	*0.006*	*0.003*	*0.003*	*0.002*	*0.003*
1990	.	0.301	0.234	0.241	0.199	0.336
	0.000	*0.007*	*0.010*	*0.009*	*0.011*	*0.007*
2000	.	0.215	0.237	0.225	0.320	0.264
	0.000	*0.013*	*0.012*	*0.012*	*0.008*	*0.011*
2014	.	0.214	0.178	0.181	0.146	0.137
	0.000	*0.013*	*0.015*	*0.015*	*0.018*	*0.020*

Table 5: P-values for null hypothesis of the type $H_0^1: \ NI({}^S\mathbf{y}_t, d) = G(\mathbf{y}_t)$ and $H_0^2: NI({}^S\mathbf{y}_t, d) = NI({}^S\mathbf{y}_t, 0.4)$, with $t = 1980, 1990, 2000, 2014$ and $G(\mathbf{y}_{1980}) = 0.434$, $G(\mathbf{y}_{1990}) = 0.461$, $G(\mathbf{y}_{2000}) = 0.473$, $G(\mathbf{y}_{2014}) = 0.486$. Differences in levels of the NI index are in italic.

table shows that H_0^1, that is that the NI curve is equal in levels to the citywide Gini index at some distance thresholds, cannot be rejected at small distance ranges (below 2 miles) characterizing individual neighborhoods. This pattern reflects the increase in local income heterogeneity (at very narrow distance ranges) introduced in the artificial distribution, which is now allowed to recover neighborhood inequality even within the block group. Panel B of the table reports instead p-values for H_0^2. Across all years, we find that inequality measured in individual neighborhoods of size smaller than 0.4 miles is statistically indistinguishable from inequality computed with larger neighborhoods, even when radius of the neighborhoods spans over 12 miles (i.e., each neighborhood almost cover the entire Chicago MSA size).

F Discussion

Section C shows that the normality assumption has no implications for estimation of NI index, but is has for the identification of its SE bounds. Assuming that the spatial

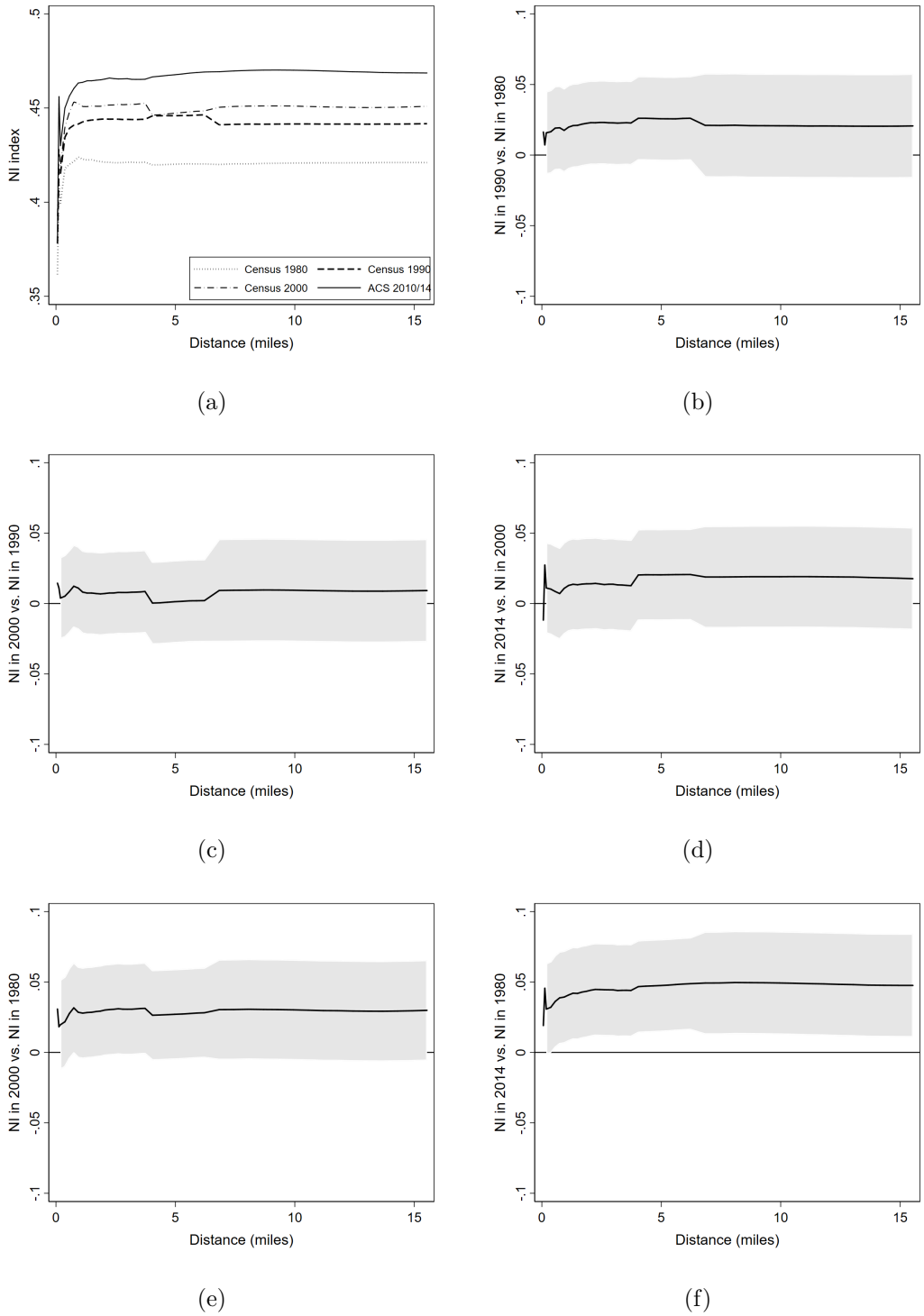
distribution of income is drawn from log-normals (a widely used hypothesis in income distribution analysis), we derive new SE bounds. The empirical estimates of such bounds, based on the Census and ACS data, are marginally larger than those obtained under normality. Empirical evidence reported in the paper is nonetheless verified at conventional (90%) significance level. The simulation exercise assumes log-normality and verifies the power and size of tests of hypothesis based on the normality assumption. We find that SE bounds based on the Gaussian hypothesis induce tests that are conservative against the null (low size) but highly discriminatory (high power) when applied to non-gaussian data generating processes.

Edge effects generated by the administrative division of the urban territory play a minor role on the NI index and its SE bounds estimates (block groups are conventional divisions of the Census for data collection purposes, not for local administration purposes), and only emerges because of the nature of the data we use in the empirical application. Furthermore, there exists techniques to control for these edge effects in calculating spatial association measures, such as the variogram. In section D we focus instead on the second source of edge effects, namely those arising from the choice of boundaries of the urban areas. Setting boundaries for the urban area has implication for setting the individual neighborhoods size (hence the converging values of the NI index) and for defining bounds on local variability of the data. Following relevant literature, we have devised an algorithm that identifies boundary areas to the Chicago MSA, we have designed a buffer zone, and considered units in the buffer zone when constructing individual neighborhood of variable size of units outside the buffer zone (in core Chicago area) while we have disregarded in the calculation of the NI index the contribution of the individual neighborhood inequality estimates attributable to units located in the buffer zone (weights have been adjusted accordingly). We have proceeded in a similar way to construct relevant variograms under the normality assumption. We do not detect any clear pattern in NI and its SE bounds estimates that differs from what reported in the main paper. The Monte Carlo study also employs this buffer zone when simulating potential income distributions. Simulated size of tests based on the SE bounds (under normality) are comparable to those reported

in the paper in small samples, and smaller than those in Table 1* for larger samples of 8000 units. When accounting for edge effects, the power of the NI index decreases, with upper bounds above 50%. The simulation study shows that edge effects related to boundaries of the urban area may play a role in small samples, the effect being stronger when tests of neighborhood inequality are performed on individual neighborhoods of small size. Drawing samples of size larger than 8000 units may help mitigating the problem.

In section E we address the effects of a change in support observed in the empirical application. The robustness check based on the artificial spatial distribution (obtained by evenly spreading income units across each block group are) reveals that patterns in neighborhood inequality observed in the data reflect those obtained with the artificial distribution ${}^{S}\mathbf{y}_t$, the evidence being robust across years. We also find statistical support. Albeit arbitrary, uniformly scattering income units within the block group boundaries reveals that NI estimated based on original data are underestimating neighborhood inequality when individual neighborhoods are of very small size. The pattern of underestimation seems, however, stable across time. Confidence intervals (hence rejection regions for H_0^3) do not vary in size when adding artificial heterogeneity in the local distribution of locations. We stress that such effect is driven from the way data are constructed and distributed (they are organized in spatial units, we assume all income units are located at the block groups centroids) rather than by a feature of the NI index or of its SE bounds, which always assume a continuous spatial pattern of the underlying data.

Figure 8: Neighborhood inequality in Chicago, IL - data scattered uniformely at the neighborhood level



Note: Author analysis of US Census and ACS data. Confidence intervals are at 95% level. SE bounds calculated under the normality assumption.

References

Andreoli, F. and Peluso, E. (2018). So close yet so unequal: Neighborhood inequality in American cities, *ECINEQ Working paper 2018-477* .

Andreoli, F. and Peluso, E. (2020). Inference for the neighborhood inequality index, *mimeo* .

Battistin, E., Blundell, R. and Lewbel, A. (2009). Why is consumption more log normal than income? gibrat's law revisited, *Journal of Political Economy* **117**(6): 1140–1154.

Bunch, J. and Parlett, B. (1971). Direct methods for solving symmetric indefinite systems of linear equations, *SIAM Journal on Numerical Analysis* **8**(4): 639–655.

Cressie, N. (1985). Fitting variogram models by weighted least squares, *Journal of the International Association for Mathematical Geology* **17**(5): 563–586.

Cressie, N. A. C. (1991). *Statistics for Spatal Data*, John Wiley & Sons, New York.

Cressie, N. and Hawkins, D. M. (1980). Robust estimation of the variogram: I, *Journal of the International Association for Mathematical Geology* **12**(2): 115–125.

Diggle, P. J., Tawn, J. A. and Moyeed, R. A. (1998). Model-based geostatistics, *Journal of the Royal Statistical Society: Series C (Applied Statistics)* **47**(3): 299–350.

Diggle, P., Zheng, P. and Durr, P. (2005). Nonparametric estimation of spatial segregation in a multivariate point process: Bovine tuberculosis in cornwall, uk, *Journal of the Royal Statistical Society. Series C (Applied Statistics)* **54**(3): pp. 645–658.

Griffith, D. A. (1983). The boundary value problem in spatial statistical analysis*, *Journal of Regional Science* **23**(3): 377–387.

Griffith, D. A. (1985). An evaluation of correction techniques for boundary effects in spatial statistical analysis: Contemporary methods, *Geographical Analysis* **17**(1): 81–88.

Griffith, D. A. and Csillag, F. (1993). Exploring relationships between semi-variogram and spatial autoregressive models, *Papers in Regional Science* **72**(3): 283–295.

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

Türk, U. and Östh, J. (2019). How much does geography contribute? Measuring inequality of opportunities using a bespoke neighbourhood approach, *Journal of Geographical Systems* **21**(2): 295–318.

Xu, C. and Dowd, P. A. (2012). *The Edge Effect in Geostatistical Simulations*, Springer Netherlands, Dordrecht, pp. 115–127.