Unveiling the Urban Divide

Novel Insights into Economic Segregation Using Fine-Grained Data

Javier San Millán Tejedor, Clémentine Cottineau & Maarten Van Ham
TU Delft

ECTQG'23

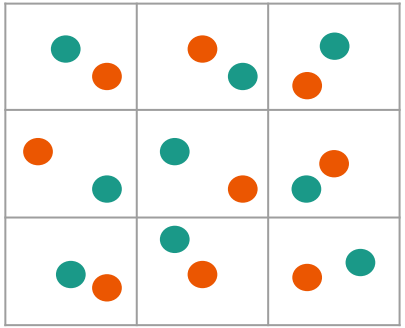
What is it?

Economic inequality

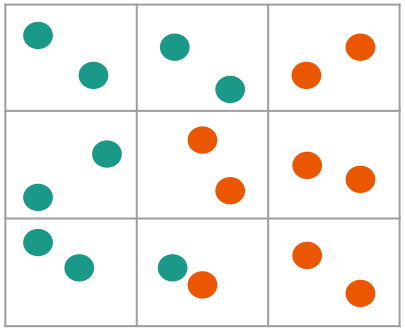
The uneven distribution of economic resources among the population

Urban economic segregation

The uneven distribution of population groups throughout the urban area on the basis of their economic status



Equally unequal, unequally segregated



Why do we care about urban economic segregation?

- Threatens social cohesion
- Exacerbates the regressive impact of contextual effects
- Amplifies and reinforces other forms of inequality

Why do we care now?

Inequality is (seemingly) rising

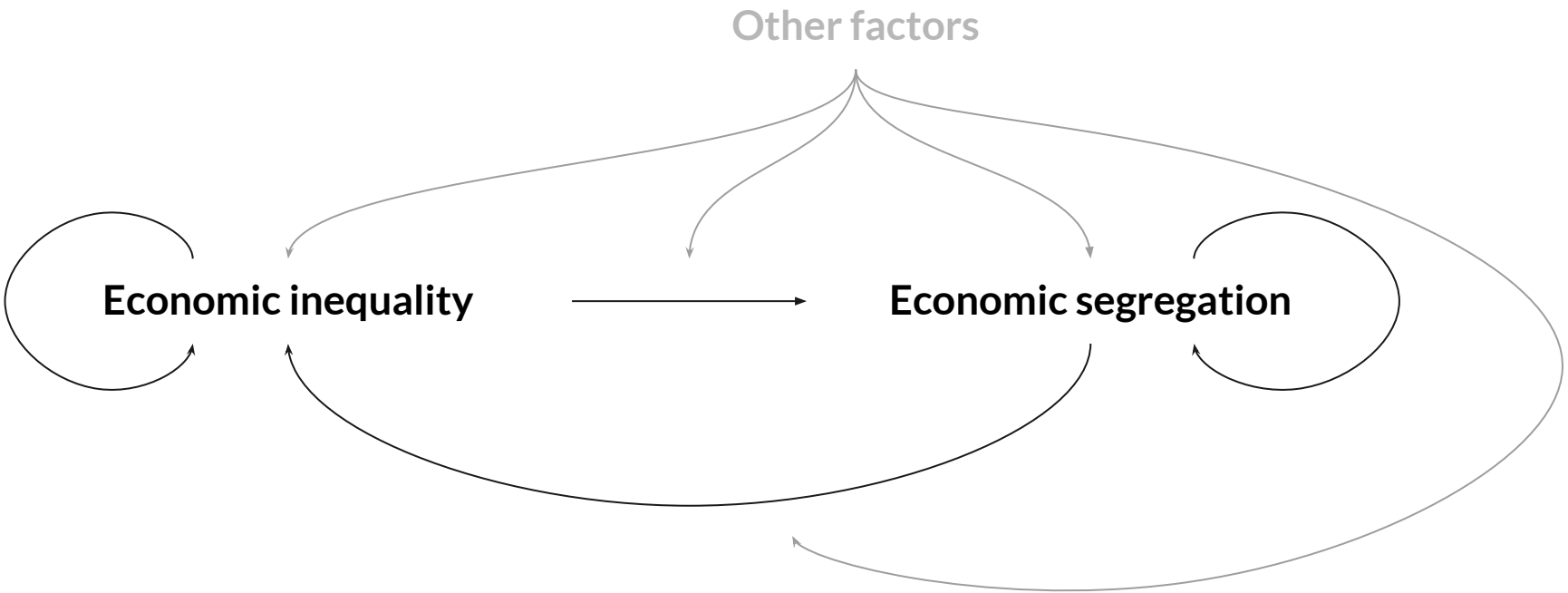
↓

Segregation is (seemingly) rising

Research puzzles:

Several cities record diverging evolutions of inequality and segregation (Van Ham *et al.*, 2021)
Data is often incomplete, relies on approximate proxies and is based on decennial censuses
The link between inequality and segregation is probably mediated, multi-level and time-delayed

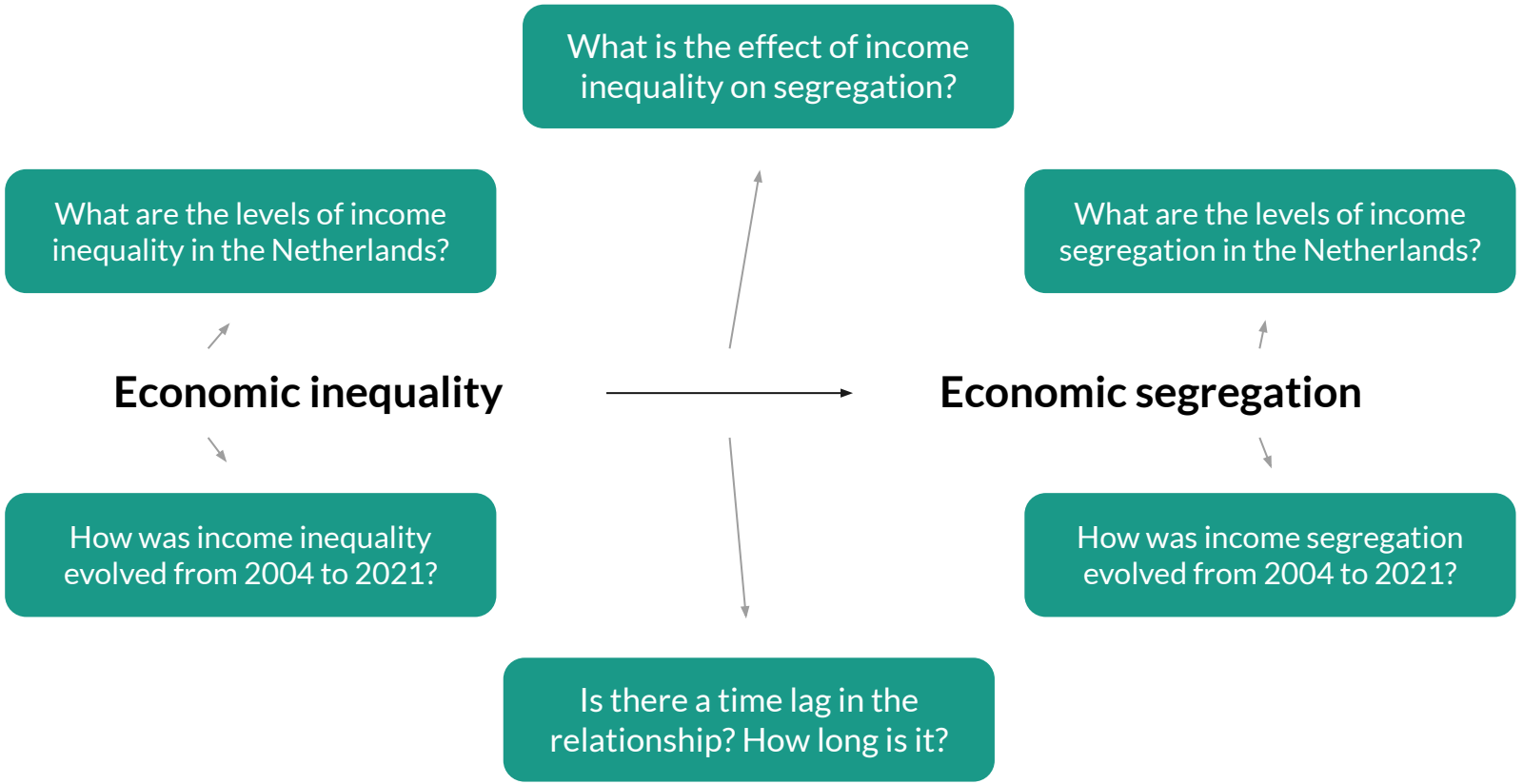
What do we aim to study?



What do we aim to study?

Economic inequality → **Economic segregation**

What do we aim to study?

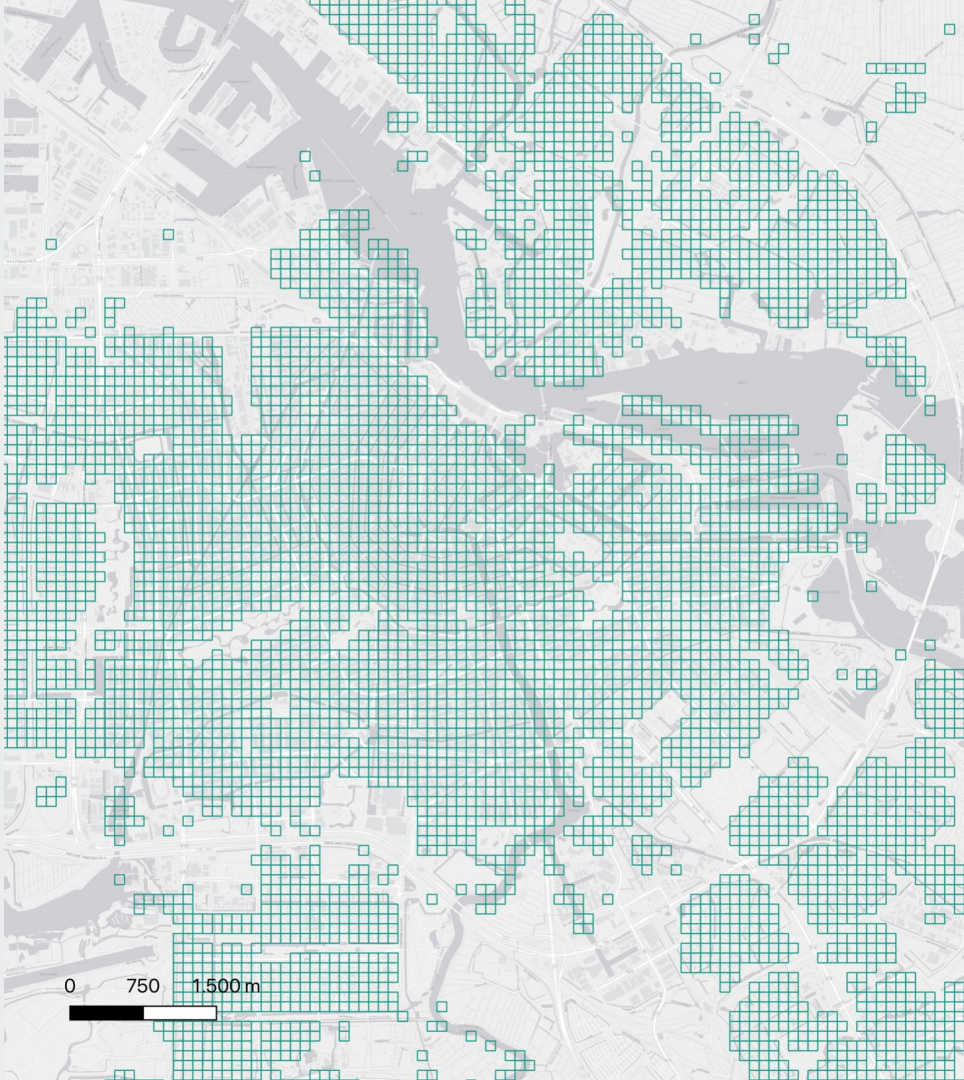


Data

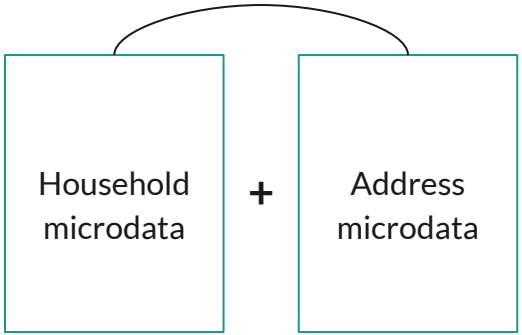
Household microdata from Statistics Netherlands (CBS)

Data on the annual income of households + residential location:

- From 2004 to 2021
- Covering the entire population of the Netherlands
- Geo-coded at a very high resolution (100m x 100m grid cells)



Methodology



Every household is assigned to a metropolitan area

35 Functional Urban Areas following Eurostat and the OECD classification:

- They capture the actual extent of urban areas
- Regardless of institutional boundaries

Income inequality estimations

Measured by the Gini coefficient

Segregation estimations

Fixed-effects regression

Accounting for metropolitan fixed-effects and time-varying characteristics

Measuring urban segregation

Rank-Ordered Information Theory Index

Ratio of within-unit (grid cell) income rank variation to overall (metropolitan area) income rank variation

- Adapted to a continuous/ordinal variable
- Does not rely on arbitrary thresholds and takes advantage of all income information
- Independent of actual income inequality

The Rank-Order Information Theory Index

Reardon et al. (2006) describe the rank-order information theory index in detail; we summarize its key features here. First, let p denote income percentile ranks (scaled to range from zero to one) in a given income distribution (i.e., $p = F(Y)$, where Y measures income and F is the cumulative income density function). Now, for any given value of p, we can dichotomize the income distribution at p and compute the residential (pairwise) segregation between those with income ranks less than p and those with income ranks greater than or equal to p. Let $H(p)$ denote the value of the traditional information theory index (Theil and Finezza 1971; Theil 1972; Zoloth 1976; James and Taeuber 1985) of segregation computed between the two groups so defined. Likewise, let $E(p)$ denote the entropy of the population when divided into these two groups (Theil and Finezza 1971; Theil 1972; Pielou 1977). That is,

$$E(p) = p \log_2 \frac{1}{p} + (1 - p) \log_2 \frac{1}{1 - p} \quad (1)$$

and

$$H(p) = 1 - \sum_j \frac{t_j E_j(p)}{TE(p)}, \quad (2)$$

where T is the population of the metropolitan area and t_j is the population of neighborhood j. Then the rank-order information theory index (H^R) can be written as

$$H^R = 2 \ln(2) \int_0^1 E(p) H(p) dp. \quad (3)$$

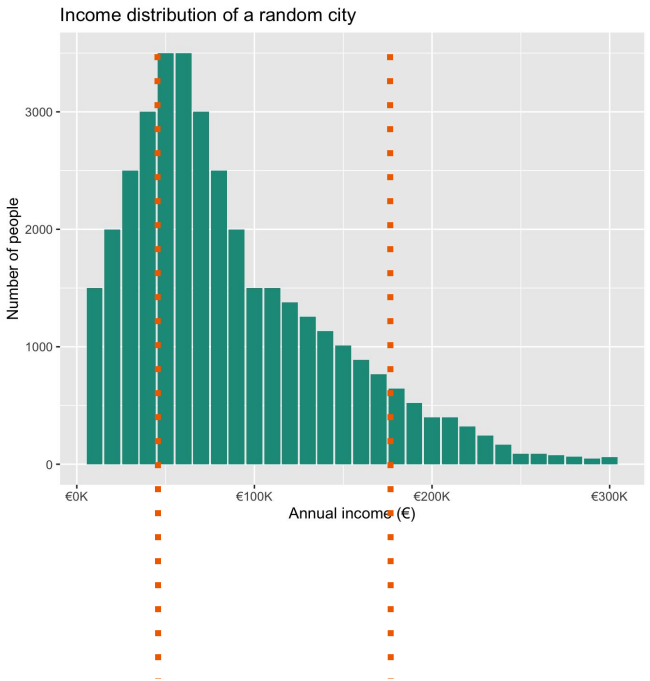
Extracted from Reardon & Bischoff (2011)

Measuring urban segregation

Rank-Ordered Information Theory Index

Ratio of within-unit (grid cell) income rank variation to overall (metropolitan area) income rank variation

- Adapted to a continuous/ordinal variable
- Does not rely on arbitrary thresholds and takes advantage of all income information
- Independent of actual income inequality



Bottom 20% **Top 20%**

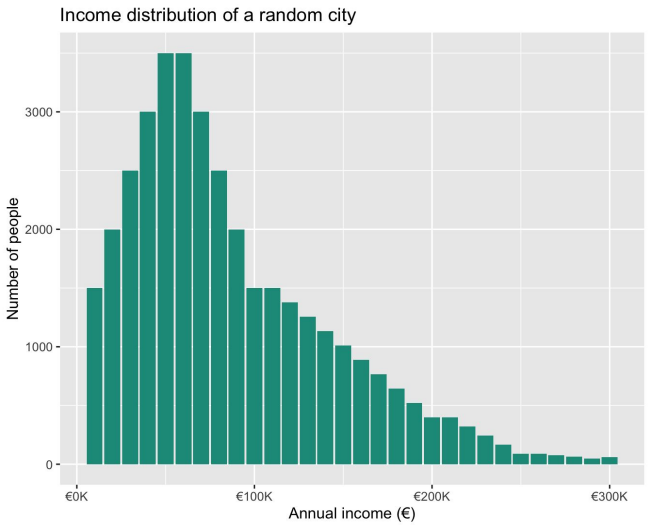
e.g. Dissimilarity Index

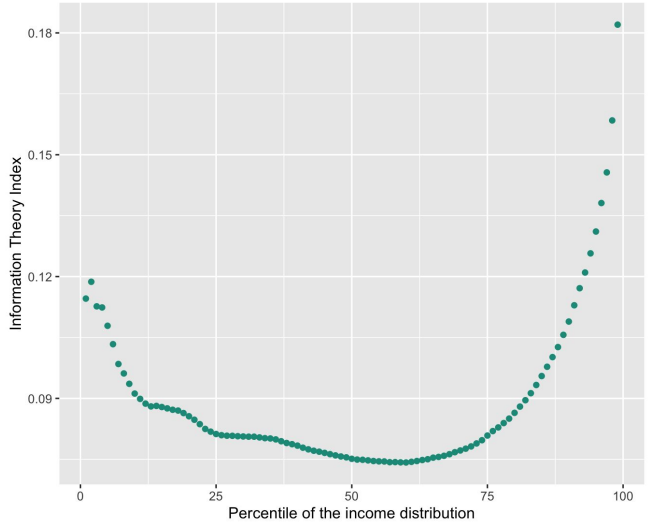
Measuring urban segregation

Rank-Ordered Information Theory Index

Ratio of within-unit (grid cell) income rank variation to overall (metropolitan area) income rank variation

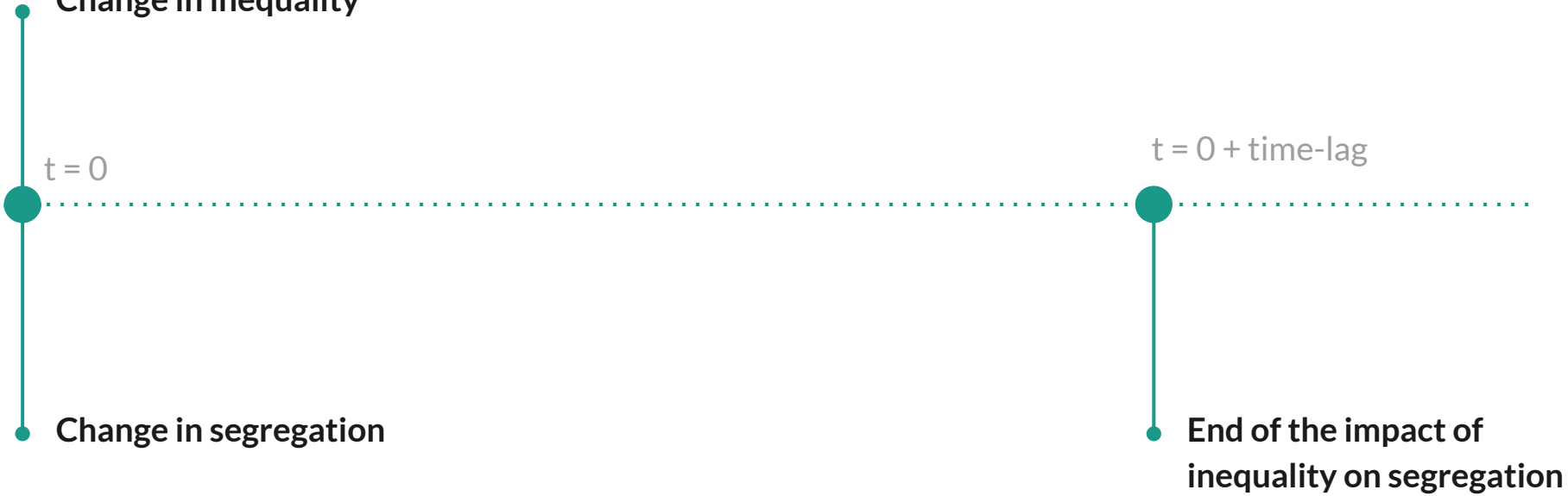
- Adapted to a continuous/ordinal variable
- Does not rely on arbitrary thresholds and takes advantage of all income information
- Independent of actual income inequality



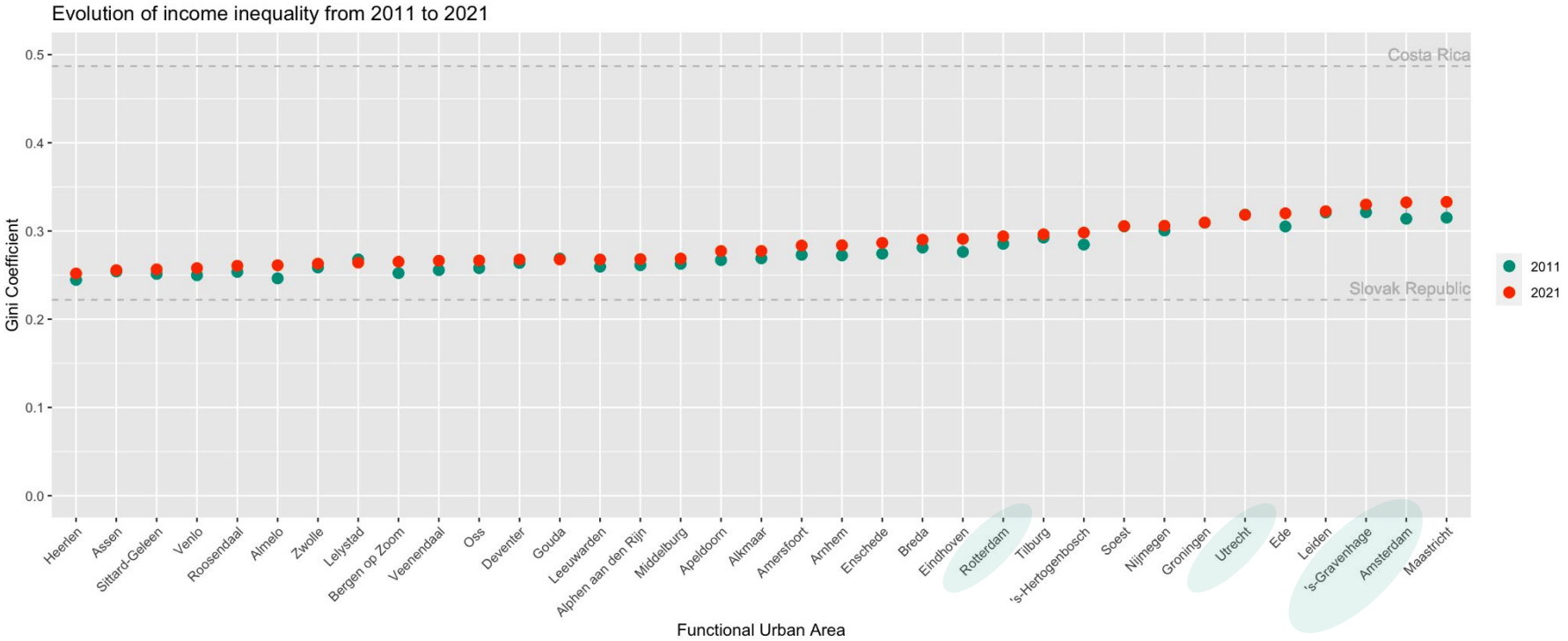


Time-delayed effect

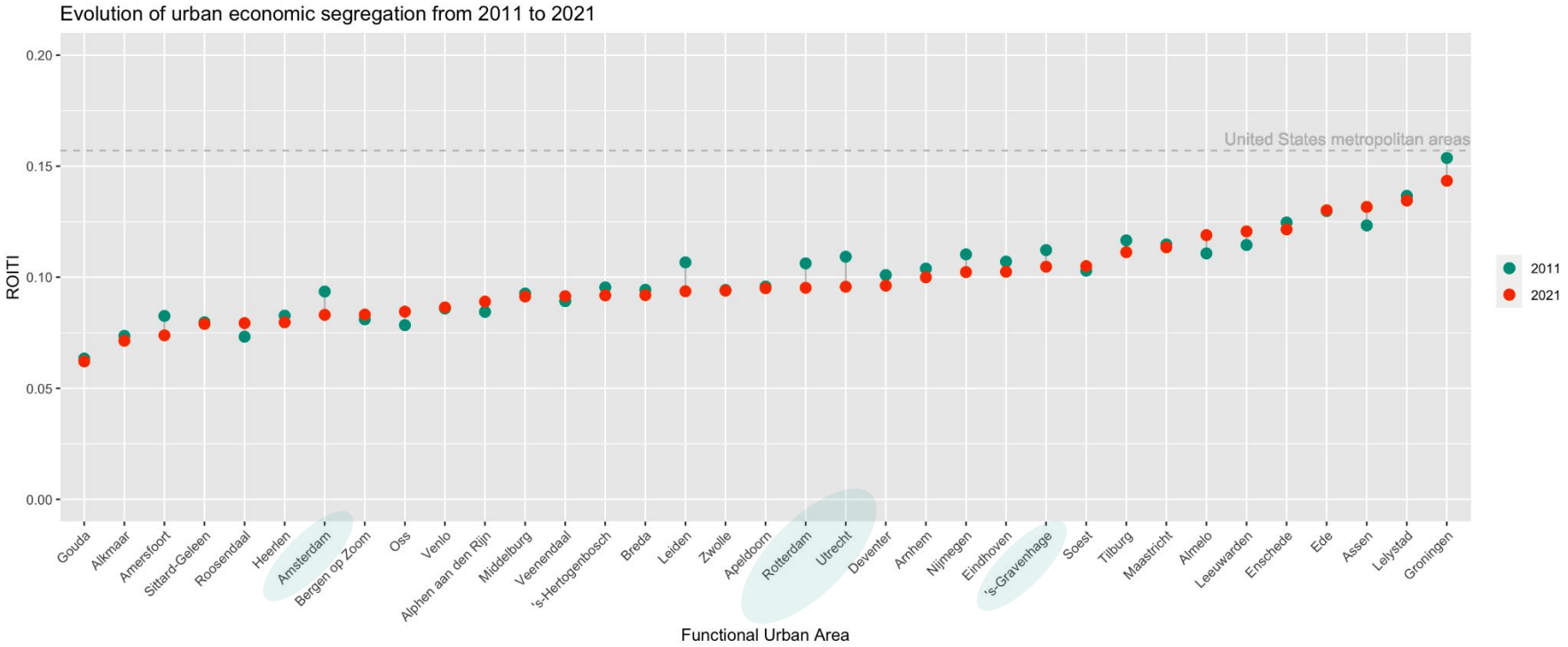
Part of the impact of inequality takes time to get translated into space



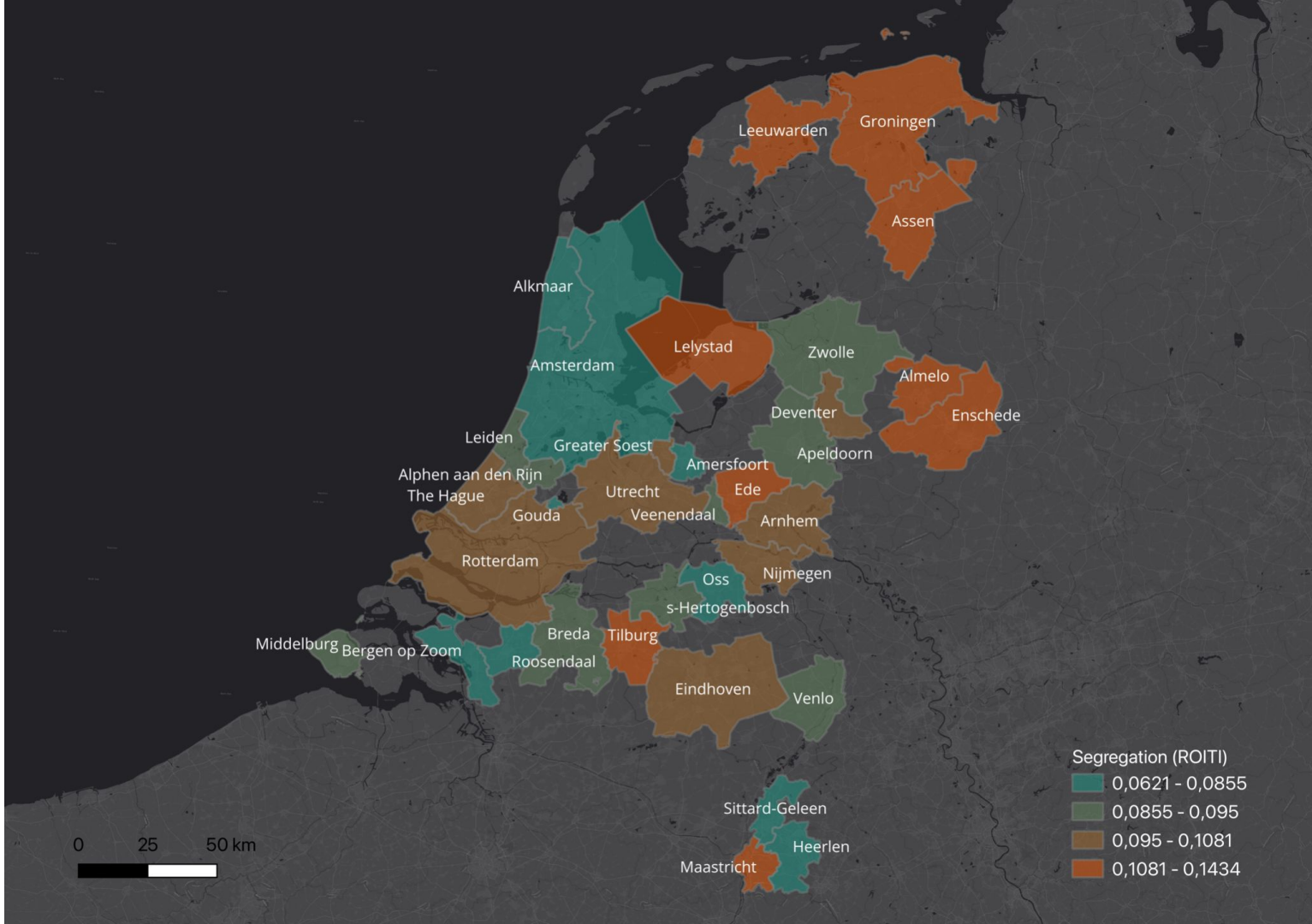
Results


Evolution of income inequality from 2011 to 2021
Gini Coefficient
0.5
0.4
0.3
0.2
0.1
0.0
Costa Rica
Slovak Republic
2011
2021
Heerlen
Assen
Sittard-Geleen
Venlo
Roosendaal
Almelo
Zwolle
Lelystad
Bergen op Zoom
Veenendaal
Oss
Deventer
Gouda
Leeuwarden
Alphen aan den Rijn
Middelburg
Apeldoorn
Alkmaar
Amersfoort
Arnhem
Enschede
Breda
Eindhoven
Rotterdam
Tilburg
's-Hertogenbosch
Soest
Nijmegen
Groningen
Utrecht
Ede
Leiden
's-Gravenhage
Amsterdam
Maastricht
Functional Urban Area

Results

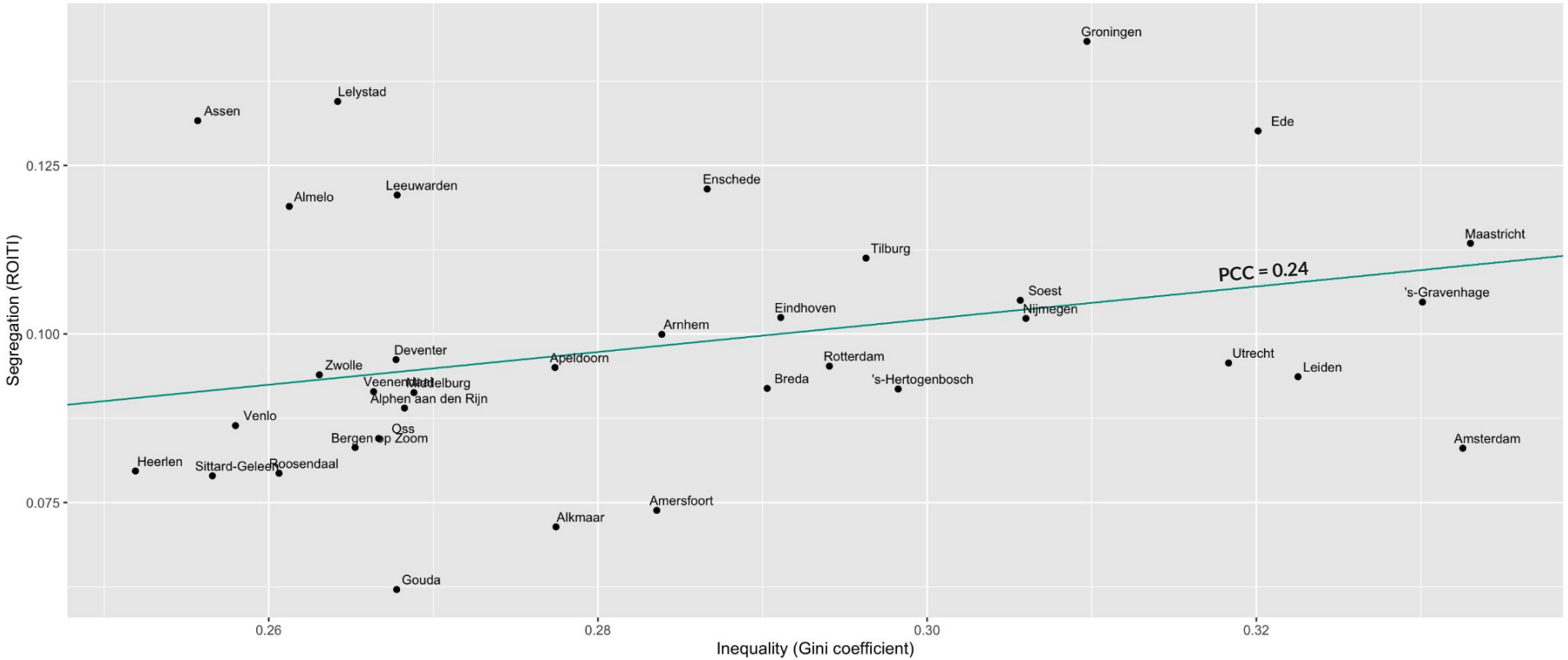

Evolution of urban economic segregation from 2011 to 2021
ROITI
0.20
0.15
0.10
0.05
0.00
United States metropolitan areas
2011
2021
Gouda
Alkmaar
Amersfoort
Siltard-Geleen
Roosendaal
Heerlen
Amsterdam
Bergen op Zoom
Oss
Venlo
Alphen aan den Rijn
Middelburg
Veenendaal
's-Hertogenbosch
Breda
Leiden
Zwolle
Apeldoorn
Rotterdam
Utrecht
Deventer
Arnhem
Nijmegen
Eindhoven
's-Gravenhage
Soest
Tilburg
Maastricht
Almelo
Leeuwarden
Enschede
Ede
Assen
Lelystad
Groningen
Functional Urban Area

Results


Leeuwarden
Groningen
Assen
Alkmaar
Lelystad
Zwolle
Amsterdam
Almelo
Deventer
Enschede
Leiden
Greater Soest
Apeldoorn
Amersfoort
Alphen aan den Rijn
The Hague
Utrecht
Ede
Gouda
Veenendaal
Arnhem
Rotterdam
Nijmegen
Oss
s-Hertogenbosch
Middelburg
Bergen op Zoom
Breda
Tilburg
Roosendaal
Eindhoven
Venlo
Sittard-Geleen
Heerlen
Maastricht
0
25
50 km
Segregation (ROITI)
0,0621 - 0,0855
0,0855 - 0,095
0,095 - 0,1081
0,1081 - 0,1434

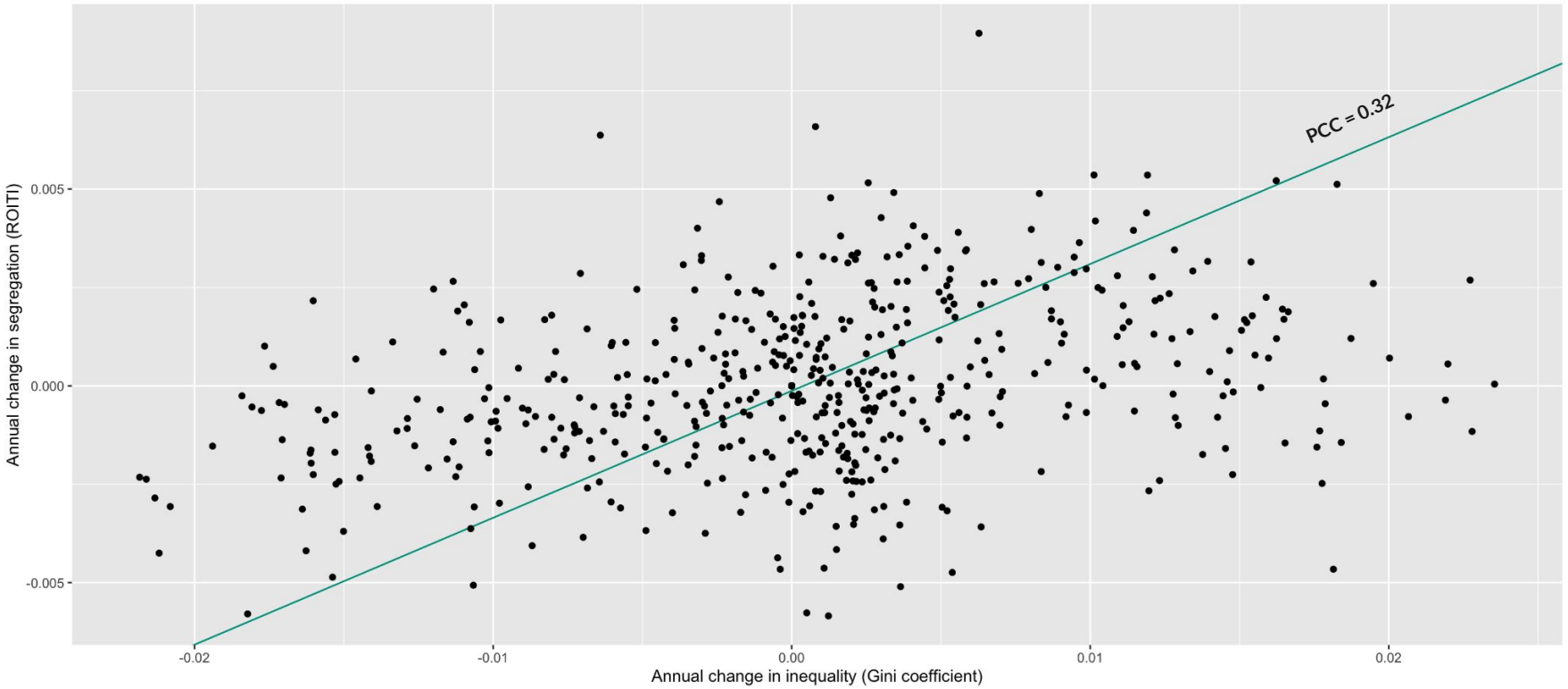
Results

Relationship between inequality and segregation in 2021



Results

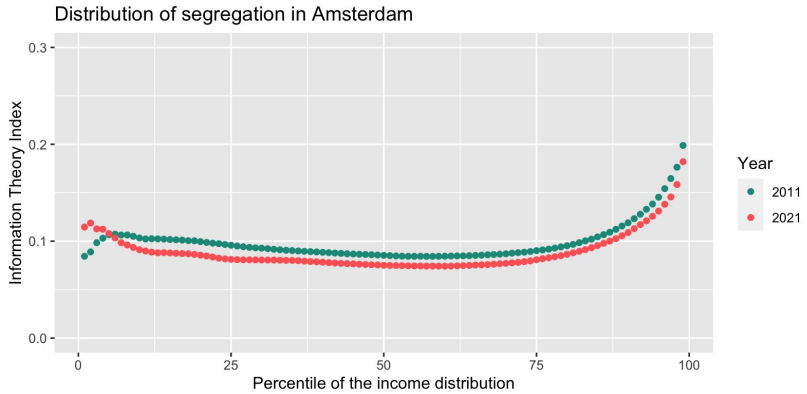
Relationship between changes in inequality and changes in segregation

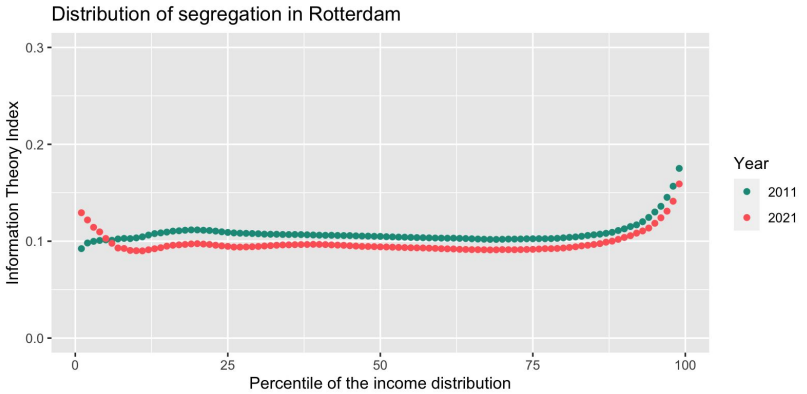


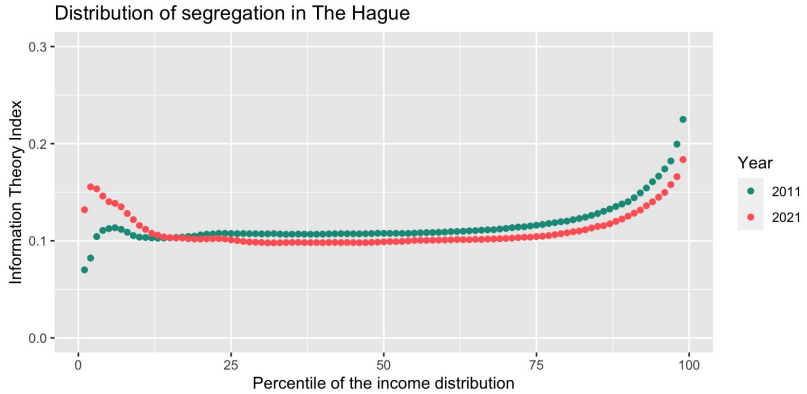
Results

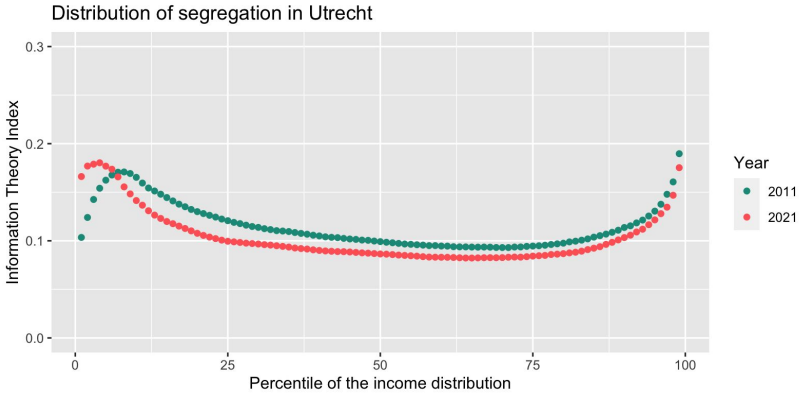
	(1A) Static values	(1B)	(2A) Annual changes	(2B)	(3A) 5-year intervals	(3B)	(4A) 10-year intervals	(4B)
Gini	**0.2439*** **	**0.1518*** **	**0.0947*** **	**0.0938*** **	**0.6132*** **	**0.6255*** **	**0.0743***	**0.1080*** **
	(0.0367)	(0.0370)	(0.0115)	(0.0113)	(0.0521)	(0.0521)	(0.0296)	(0.0256)
Change in data collection methods	**0.0194*** **	**0.0139*** **	**0.0221*** **	**0.0220*** **	**-0.0143*** **	**-0.0125*** **	**-0.0277*** **	**-0.0221*** **
	(0.0007)	(0.0001)	(0.0004)	(0.0004)	(0.0008)	(0.0014)	(0.0008)	(0.0009)
Number of households (log)		0.0136		-0.0046		0.0313929		0.0034
		(0.0118)		(0.0043)		(0.0199)		(0.01059)
Average income (log)		**0.0222*** **		-0.0019		**-0.0202***		**-0.0421*** **
		(0.0044)		(0.0014)		(0.0083)		(0.0048)
N	630	650	595	595	455	455	280	280
Adjusted R^2	**0.7279**	**0.7517**	**0.8452**	**0.8514**	**0.4612**	**0.4663**	**0.8175**	**0.87258**

Results


Distribution of segregation in Amsterdam
Information Theory Index
0.0
0.1
0.2
0.3
0
25
50
75
100
Percentile of the income distribution
Year
2011
2021


Distribution of segregation in Rotterdam
Information Theory Index
0.0
0.1
0.2
0.3
0
25
50
75
100
Percentile of the income distribution
Year
2011
2021


Distribution of segregation in The Hague
Information Theory Index
0.0
0.1
0.2
0.3
0
25
50
75
100
Percentile of the income distribution
Year
2011
2021


Distribution of segregation in Utrecht
Information Theory Index
0.0
0.1
0.2
0.3
0
25
50
75
100
Percentile of the income distribution
Year
2011
2021

Next steps

- Calculate the spatial version of the Rank-Ordered Information Theory Index
- Testing the time lag through Granger Causality tests and/or VAR models
- Refine the regression models (e.g. adding FUA time-varying characteristics)

Obrigado!

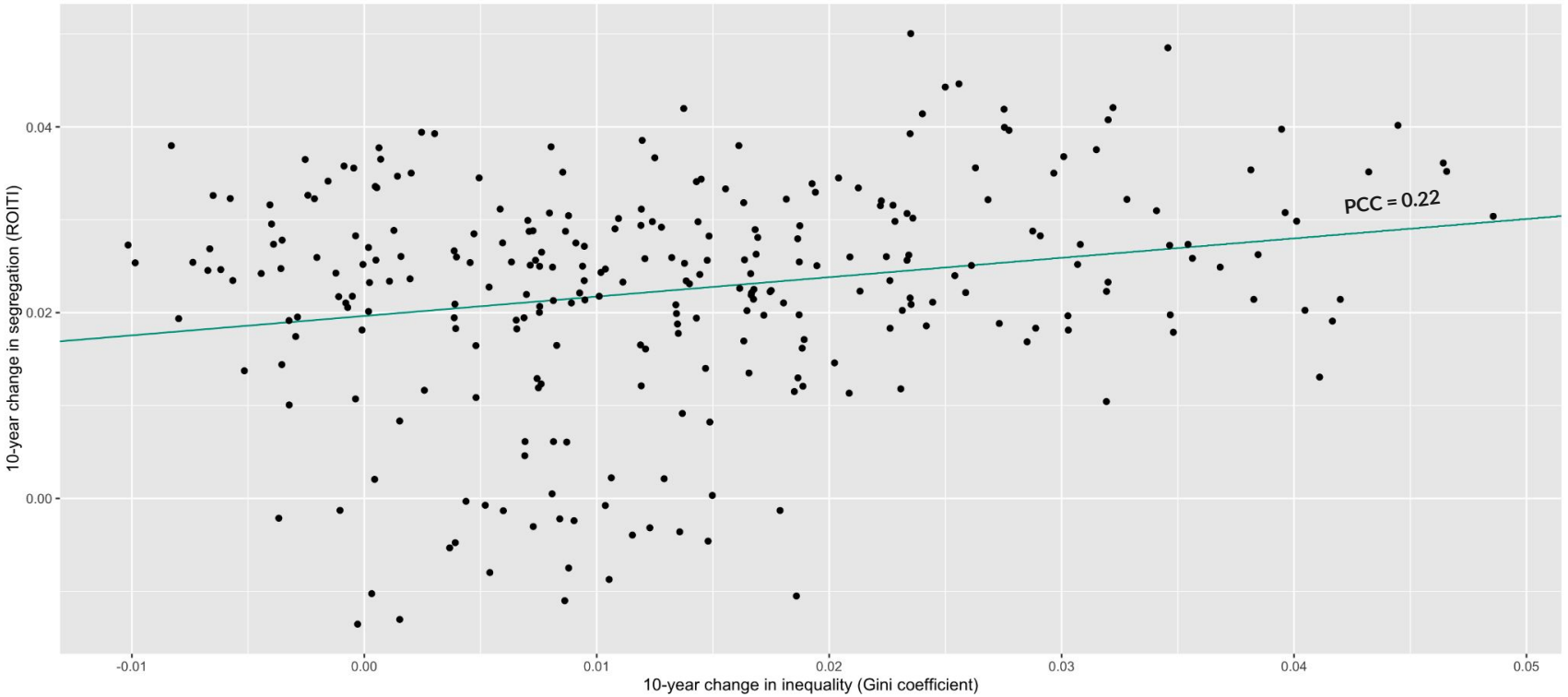
j.sanmillantejedor@tudelft.nl

@javisanmillan

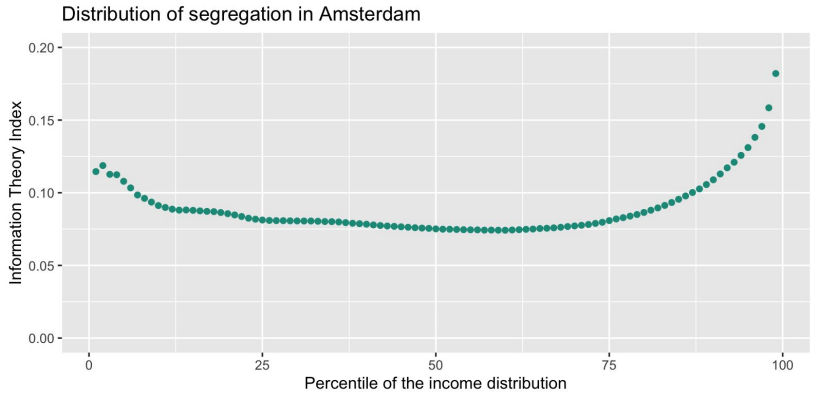
@ERC_SEGUE

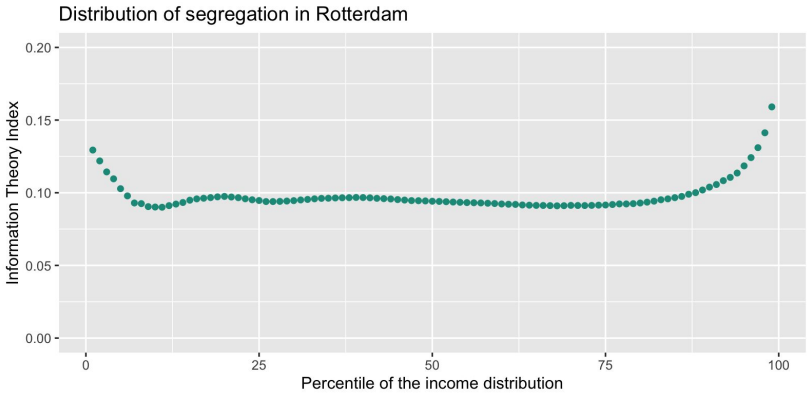
Results

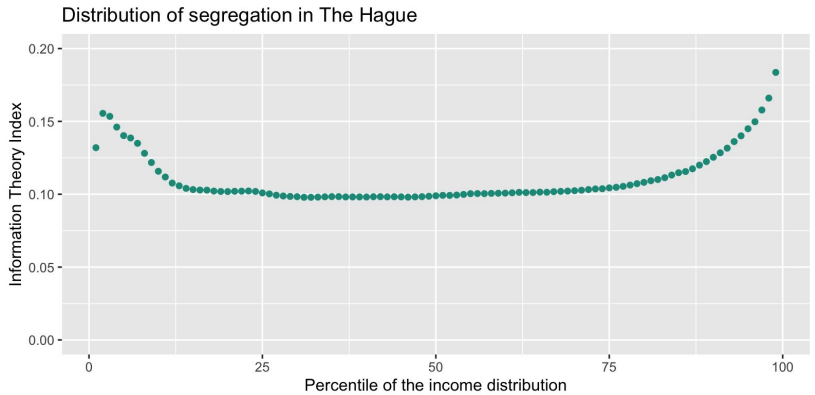
Relationship between decennial changes in inequality and changes in segregation

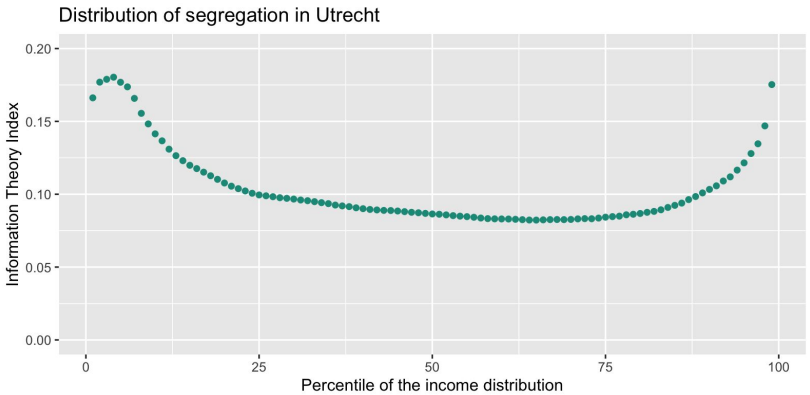


Results

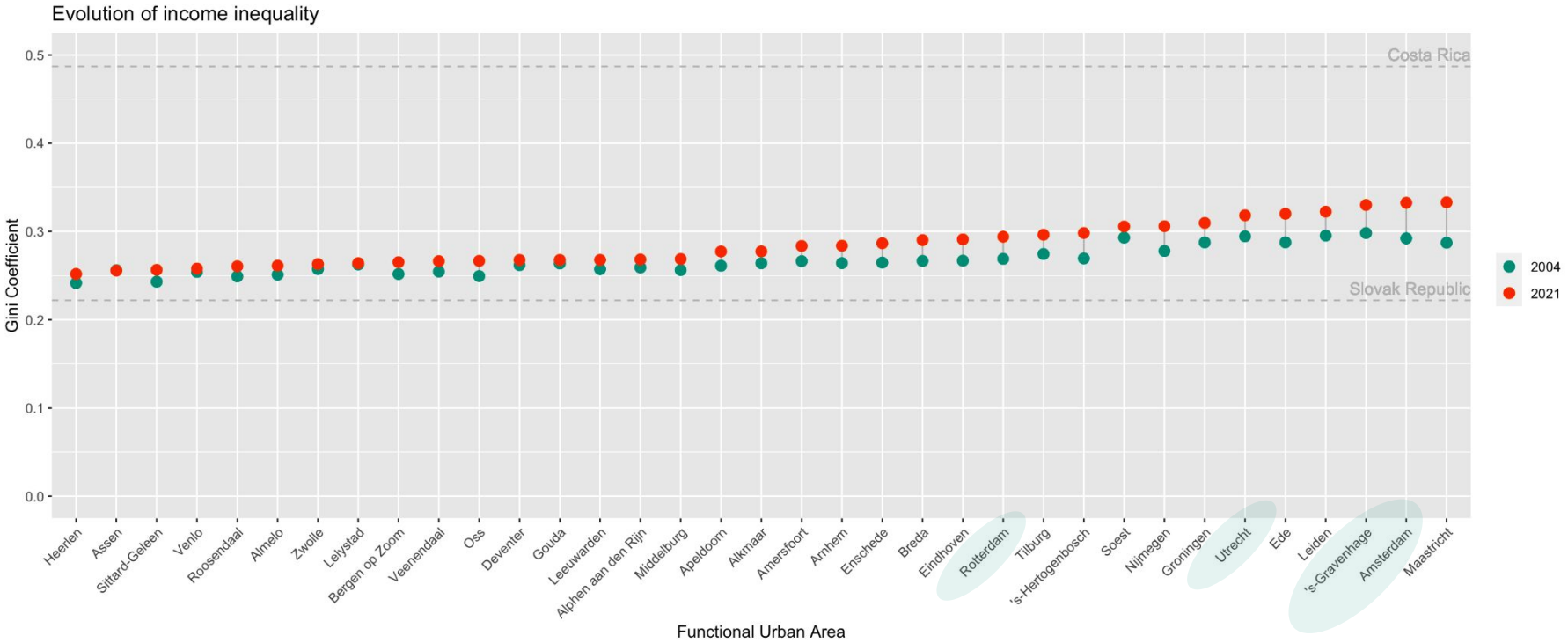

Distribution of segregation in Amsterdam
Information Theory Index
0.20
0.15
0.10
0.05
0.00
0
25
50
75
100
Percentile of the income distribution


Distribution of segregation in Rotterdam
Information Theory Index
0.20
0.15
0.10
0.05
0.00
0
25
50
75
100
Percentile of the income distribution

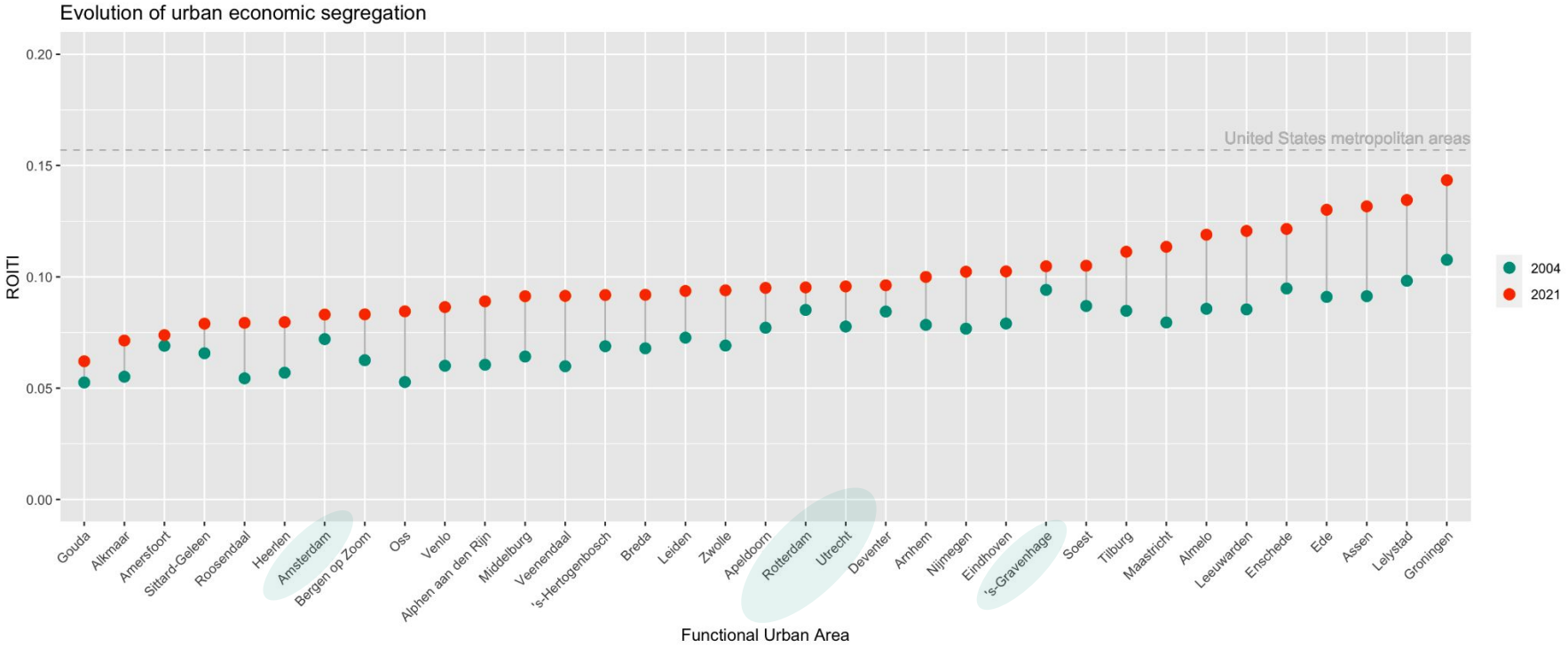

Distribution of segregation in The Hague
Information Theory Index
0.20
0.15
0.10
0.05
0.00
0
25
50
75
100
Percentile of the income distribution


Distribution of segregation in Utrecht
Information Theory Index
0.20
0.15
0.10
0.05
0.00
0
25
50
75
100
Percentile of the income distribution

Results


Evolution of income inequality
Gini Coefficient
0.5
0.4
0.3
0.2
0.1
0.0
Costa Rica
Slovak Republic
2004
2021
Heerlen
Assen
Sittard-Geleen
Venlo
Roosendaal
Almelo
Zwolle
Lelystad
Bergen op Zoom
Veenendaal
Oss
Deventer
Gouda
Leeuwarden
Alphen aan den Rijn
Middelburg
Apeldoorn
Alkmaar
Amersfoort
Arnhem
Enschede
Breda
Eindhoven
Rotterdam
Tilburg
's-Hertogenbosch
Soest
Nijmegen
Groningen
Utrecht
Ede
Leiden
's-Gravenhage
Amsterdam
Maastricht
Functional Urban Area

Results


Evolution of urban economic segregation
ROITI
0.20
0.15
0.10
0.05
0.00
United States metropolitan areas
2004
2021
Gouda
Alkmaar
Amersfoort
Siltard-Geleen
Roosendaal
Heerlen
Amsterdam
Bergen op Zoom
Oss
Venlo
Alphen aan den Rijn
Middelburg
Veenendaal
's-Hertogenbosch
Breda
Leiden
Zwolle
Apeldoorn
Rotterdam
Utrecht
Deventer
Arnhem
Nijmegen
Eindhoven
's-Gravenhage
Soest
Tilburg
Maastricht
Almelo
Leeuwarden
Enschede
Ede
Assen
Lelystad
Groningen
Functional Urban Area