IEB Working Paper 2020/07

WHAT DO DIVIDED CITIES HAVE IN COMMON? AN INTERNATIONAL COMPARISON OF INCOME SEGREGATION

Paolo Veneri, Andre Comandon, Miquel-Àngel Garcia-López, Michiel N. Daams

Version July 2020

Cities

WHAT DO DIVIDED CITIES HAVE IN COMMON? AN INTERNATIONAL COMPARISON OF INCOME SEGREGATION

Paolo Veneri, Andre Comandon, Miquel-Àngel Garcia-López, Michiel N. Daams

Postal Address:
Institut d'Economia de Barcelona
Facultat d'Economia i Empresa
Universitat de Barcelona
C/ John M. Keynes, 1-11
(08034) Barcelona, Spain
Tel.: + 34 93 403 46 46
ieb@ub.edu
http://www.ieb.ub.edu

IEB Working Paper 2020/07

WHAT DO DIVIDED CITIES HAVE IN COMMON? AN INTERNATIONAL COMPARISON OF INCOME SEGREGATION *

Paolo Veneri, Andre Comandon, Miquel-Àngel Garcia-López, Michiel N. Daams

ABSTRACT: This paper provides a comparative assessment of income segregation within cities in 12 countries. We use spatial entropy indexes based on small-scale gridded income data and consistent definition of city boundaries to ensure international comparability of our segregation measures. Results show considerable variation in the levels of income segregation across cities, even within countries, reflecting the diversity of cities within urban systems. Larger, more affluent, productive, and more unequal cities tend to be more segregated. Urban form, demographic, and economic factors explain additional variation in segregation levels through the influence of high-income households, who tend to be the most segregated. The positive association between productivity and segregation is mitigated in polycentric cities.

JEL Codes: D63, R12

Keywords: Segregation, income, functional urban areas, international urban comparison, spatial inequalities

Paolo Veneri
OECD Centre for Entrepreneurship, SMEs, Regions and Cities
2 rue André-Pascal, 75016 Paris, France
E-mail: paolo.veneri@oecd.org

Andre Comandon
University of California, Los Angeles (UCLA), United States
Email: acomandon@ucla.edu

Miquel-Àngel Garcia-López
Universitat Autònoma de Barcelona & IEB, Spain
E-mail: miquelangel.garcia@uab.cat

Michiel N. Daams
University of Groningen, Department of Economic Geography
Groningen, Netherlands
E-mail: m.n.daams@rug.nl

* Authors would like to acknowledge the support by the Gran Sasso Science Institute, Italy. Miquel-Àngel Garcia-López acknowledges financial support from the Ministerio de Ciencia e Innovación (research project RTI2018-097401- B-I00), Generalitat de Catalunya (research project 2017SGR1301), and the "Xarxa de Referència d'R+D+I en Economia Aplicada". Any errors remain the authors'. The views expressed herein are those of the authors and do not necessarily correspond to the views of the OECD or of its member countries.

1. INTRODUCTION

Cities unite people of different cultural, ethnic and socio-economic backgrounds. Within such a diversity, similar individuals often congregate and, simultaneously, separate from other groups. The spatial sorting of people in different neighborhoods according to their socio-economic and cultural characteristics is an inherent part of how cities grow and acquire their structure. A long line of research shows how this process occurs and why (Mossay & Picard, 2019; Schelling, 1971). This process is not in itself harmful in terms of social welfare, and, in some cases, even desirable to developing communities (Ellickson, 2006; Merry, 2016; Morrison, 2015). However, many cities are truly divided with potentially damaging consequences.

We call divided cities those cities where the combination of socio-economic inequality, in terms of income, correlates with the concentration of socio-economic classes in space. This combination exacerbates societal disparities and fosters a vicious cycle that breaks the mechanisms of upward mobility for low-income households, both in terms of income and neighborhood of residence (Nieuwenhuis, Tammaru, Van Ham, Hedman, & Manley, 2020).

A large amount of research on socio-economic or ethnic segregation in cities has tried to identify possible negative consequences of socio-economic and ethnic segregation on individual and social outcomes (e.g., Chetty & Hendren, 2018; Novara, Loury, & Khare, 2017; Oreopoulos, 2003). Most of the studies target specific cities or a set of cities in a single country. In these cases, empirical analyses, often with a longitudinal dimension, make it possible to identify the long terms consequences of living and growing up in isolated and disadvantaged neighborhoods, which is generally more likely to occur in highly segregated cities. Within this framework, the literature still lacks a comparative assessment of segregation patterns in different areas of the world. There are very few studies providing comparative evidence on what are the prevalent patterns of spatial inequalities in cities as well as on what are the factors that tend to explain the variation in levels of spatial inequalities. Notable efforts include the work by Tammaru, Marcińczak, & Van Ham (2016) documenting socio-spatial segregation trends in 13 European cities, collected studies in Maloutas & Fujita (2012), and previous work by Arbaci (2007). In such cases, the number of countries or cities, in addition to the geographical scope of these studies limits the ability to delve into possible differences in patterns (cf., Johnston, Poulsen, & Forrest, 2007; Musterd, 2005).

This paper aims to provide an assessment of income spatial inequalities – i.e., neighborhood segregation – in about 120 cities distributed across 12 countries. It does so by using state-of-the-art indicators of segregation based on entropy measures and considering a consistent definition of city boundaries – an issue hardly considered so far in comparative studies on spatial inequalities. In addition, we carefully address the harmonization of spatial data to further enhance international comparability. Our analysis sheds light on whether a city is more or less spatially segregated than another, what are the income groups that tend to be more spatially segregated and what are the city characteristics that are associated to higher levels of segregation, once controlling for factors that play a role at the national level. To our knowledge, no other works in the literature on the comprehensive assessment of income segregation offer such a wide geographical scope, while ensuring international comparability by applying a definition of cities, methods, and data harmonization, as this study does.

Our results show that segregation levels vary substantially across cities in the same country and they tend to be higher for the richest households, although there are exceptions in countries where segregation is lower. We investigate the factors explaining the variation in segregation levels through an econometric model where the main dependent variable is regressed against measures of city size, urban form, types of city government and other city-level socio-economic characteristics. Results confirm that city size and income inequalities in the city matters for explaining segregation levels. Larger, more affluent, and unequal cities are found to be more segregated, on average. A centralized urban form is also associated with higher segregation, while in polycentric urban structures productivity gains are less associated to increases in segregation.

The remainder of the paper is organized as follows. The next section introduces the data, methods and definitions we used to analyze segregation patterns. The third section provides an assessment of the patterns of segregation in all cities considered, documenting the overarching facts and trends. The fourth section identifies the different factors associated to the degree of segregation of different income groups. The fifth section offers concluding remarks and defines possible further research questions.

2. DATA AND DEFINITIONS

2.1. Defining cities

Measures of income segregation – and spatial inequality more generally – can be sensitive to both the size of the individual geographical sub-units (conventionally called neighborhoods) and to that of the overall urban area under investigation. Therefore, comparing levels of segregation across different countries first requires a consistent definition of cities or urban economic agglomerations. We apply the city definition based on the concept of functional urban area (FUA), see OECD (2012), where it is available and a close alternative in the other non-OECD countries.

The definition the OECD developed in collaboration with the European Union states that a FUA is a cluster of contiguous local units (i.e. municipalities, ward, census tracts, etc.) composed of a high-density urban center and a surrounding commuting zone. To achieve consistency, the method first identifies urban centers as clusters of contiguous 1 km^2 grid cells each with a density of at least 1,500 inhabitants per km^2 and a minimum combined population of 50,000 in the cluster. Next, the boundaries of urban centers are matched to the local units where most of the population lives within an urban center. We call the resulting units urban cores. Finally, all surrounding local units where at least 15% of the labor force commutes daily towards the core are considered part of the commuting catchment zone and they are added as a part of the FUA.

This paper only compares FUAs with at least 500,000 people. This ensures better data availability and that we discuss similar types of segregation for the entire sample. Segregation in large cities can be characterized by entirely isolated communities by virtue of their size or geographic locations. While the mechanism of segregation is the same in smaller cities (i.e. disparities in access to goods and services that correlate to groups' spatial concentration), the number of geographical units per city tend to be too small to allow for a reliable (and policy relevant) assessment of segregation.[1] In South Africa, Brazil and New Zealand where the exact definition of FUA is not available, we use the national definitions that come closest to the FUA standard, Metropolitan Regions in Brazil, Metropolitan Municipalities in South Africa, and the Metro Area of Auckland in New Zealand. Henceforth, unless a clear differentiation is necessary, we refer to FUAs as cities.

The building block of segregation measures is the neighborhood. Neighborhood is a catchall term with no consistent definition. This is not as much of an issue for single country studies where the neighborhood is whatever geographical subunit is available. However, we have access to large array of data, all using different units to report their data. No income data is directly available at grid level, but we create grids through apportionment (see next section). We use the term neighborhood to refer to the abstract concept of small geographical subunits for which data are available. We define a more precise unit of analysis below.

2.2. Indicators of income segregation

A large array of tools is available to measure segregation. Researchers have developed strategies for operationalizing segregation in a way so that its presence could be summarized with ease to compare cities since the 1950s (e.g., Duncan & Duncan, 1955). Since then, many advances have been made to capture different aspects of segregation. Rather than provide yet another overview of segregation measures and their respective merits (see for an extensive discussion Johnston & Jones, 2010; Johnston, Poulsen, & Forrest, 2014; Kramer, Cooper, Drews-Botsch, Waller, & Hogue, 2010; Massey & Denton, 1988; Musterd, 2005; Reardon, Firebaugh, O'Sullivan, & Matthews, 2006; Tarozzi & Deaton, 2009) we focus on how our methodological choices apply to international comparison.

Our measure of income segregation is the ordinal entropy index applied to income data. The ordinal index weighs the concentration of individuals in more distant categories more heavily, which is important for income. Table 1 illustrates the difference with a conventional measure of segregation. A multigroup index would produce the same index value for scenario A and B (0.198). Yet, theoretically, the greater concentration of lowest- and highest-income individuals in scenario A is distinct from that in scenario B. People at the extremes of the income distribution are more likely to share a neighborhood in Scenario B and overall segregation is therefore lower in that scenario (0.162 vs 0.277).[2]

[TABLE 1 about here]

We choose the ordinal entropy index because despite the continuous nature of income, census authorities report it as ordered categories (e.g., 0 to $5,000). Unlike more widespread measures like the dissimilarity index which are limited to comparison between two groups, the entropy measure handles the ordinal nature of the data and any number of categories. Furthermore, we included data on the United Kingdom and Ireland, two countries that rather than income only report social class based on occupational categories.[3]

The ordinal entropy index is best fitted for categories with consistent definitions, school grade for example. In such cases, moving from grade 2 to grade 3 is equivalent to moving from grade 5 to grade 6. Income categories rarely have this kind of uniformity of definition. The right-skewed distribution of income require narrower categories towards the left hand of the distribution and wider ones in the right tail to achieve relatively even overall distribution within categories. For example, the difference between the bottom decile of the income distribution in Paris and the median is about €25,000; the difference between the median and the top decile is €70,000. As such, equally spaced income categories would have many more people at the lower end than at the top, which often contravenes privacy requirements. The rank-order approach estimates a continuous distribution using a polynomial function to transform the bins into a continuous variable. This method, however, performs reliably only when the number of categories spans most of the income distribution (Reardon, 2011). Our sample includes countries where information about the extremes of the income distribution is missing (e.g. the Netherlands and New Zealand) and where we use social class rather than income; therefore, we apply the ordinal entropy index for all estimates.

The outcome values of the Ordinal Entropy Index are between 0 and 1. The index is a measure of relative diversity with respect to the maximum diversity provided by the entire city. As most cities are income diverse, segregation arises when within-neighborhoods inequality is lower than we would expect given the composition of the city. The maximum value of 1 indicates that all neighborhoods have a single income group and a value of 0 means that every neighborhood has the same income distribution as the city. The actual values have no intuitive interpretation because they measure average uncertainty about the composition of neighborhoods (that is, a city of homogenous neighborhoods would have a value of one; there is no uncertainty about the distribution if all people belong to the same group). The interpretation is inherently comparative; the values gain meaning when used to compare values across cities. A detailed technical description of the spatial entropy procedure is provided in Appendix I.

Measures of segregation summarize spatial structure but are aspatial in that they treat all units as spatially unrelated. For example, if a city were divided into two neighborhoods, conventional measures of evenness would not capture the difference between a configuration where the two neighborhoods are adjacent and of the same area and one where they are spatially distant and of differing size. We consider these two spatial issues as follows.

First, ideally, the data would be available at small, uniformly sized spatial units to avoid the modifiable areal unit problem (Openshaw, 1984). However, the spatial units census authorities use to disseminate data reflect rules that protect the anonymity of residents. The census authority usually uses population threshold that must be met to constitute a unit. For example, census tracts in the United States are drawn to have an average population of 4,000. They, thus, vary in size depending on the residential density of the area. Dense areas in central cities are divided into small geographical units; low-density suburban units can cover vast areas of land.

The second spatial issue is the scale at which segregation is measured. In treating neighborhoods in dense cities (e.g. Paris) and sprawling ones (e.g Atlanta) the same, segregation indexes ignore that these neighborhoods exist in different local contextual settings (see Figure 1). A person living in the center of a low-density neighborhood may be surrounded by relatively homogenous socio-economic environment for several blocks and physical barriers (e.g. a park or highway) before any noticeable change in composition. This extreme scenario is represented in panel (b) of Figure 1, a large patch of homogeneity separated from the main urban area. In contrast, someone living in a neighborhood with same composition but that is part of a dense network of neighborhoods may experience a rapid succession of socio-economic environments in the surrounding blocks, like panel (e). There is little that can be done to address the data issues tied to the spatial structure of cities (i.e. an index that controls for intra-city density variations). The spatial entropy can, however, mitigate the effect of the different scale at which the data are reported. That is, as long as the size of the squares are the same in Figure 1, the index differentiates between structures.

[FIGURE 1 about here]

The spatial entropy index mitigates the influence of both modifiable areal units and the contextual settings by creating new spatial units. The index is obtained from a set of uniformly sized local grid cells that are, in turn, based on the underlying census data (c.f., the technical description in Appendix I). This means that we can apply spatial units at the same scale (of choice) in all cities and countries. In contrast to neighborhoods which are based on administratively defined boundaries, the spatial units are the local environment surrounding any given point (p in the mathematical definition in the Appendix) at a specified radius. In other words, if we had information on the location of every household, each household would be a point and the local environment would capture the composition of all households within a set distance from the household.

Regarding the choice of scale for the local environment, however, there is no theoretical guidance. Therefore, single-city or -country studies have defined multiple scales to track how segregation changes across definitions (Clark et al., 2015; Fowler, 2016; Jones et al., 2018; Reardon et al., 2008). We are, however, more interested in maximizing the comparability of cities across countries. Comparison requires the additional consideration that, although our entropy measure of income segregation can be flexibly obtained across any distance interval, the outcome values may be driven by the scale of the underlying data. Indeed, countries report their data at different scales. Australia, for example, releases data at a finer resolution (median geographical unit is 0.17 km^2) than most other countries. The comparison of Australian data with France's larger units (median geographical unit is 0.76km^2) would risk the equivalent of comparing panels (e) and (f) in Figure 1. Higher segregation in Australia could simply be due to the higher resolution of the data. The magnitude of that bias is non-negligible (Wong, 2004).

The spatial index mitigates the bias by creating a new spatial unit that is the same scale in France and Australia. With reference to Figure 1, we are using the scale of squares in panel (f) to calculate segregation in (e). We choose 1 km^2 as an intermediate scale that balances the restrictions the data imposed. That is, the new local environment is large enough to encompass most neighborhoods, even in countries that report census results in larger units, but small enough to capture variation. We verify the scale dependence of our results by calculating indexes at four scales (500m, 1km, 2km, and 4km). While there is variation in outcomes based on scale, the substantive interpretation changes little in terms of comparison.

2.3. Data

Income is an imperfect measure of wellbeing and security, but it captures broad structural trends in spatial inequality. A challenge to measuring segregation across different countries is the different ways in which statistical agencies define income, which can refer to household disposable income, personal net income, personal gross or taxable income, among others. Table 2 summarizes the source of the data and important summary measures about data quality and coverage. Entropy indexes are obtained using data for cities across 12 countries; Canada, Ireland, New Zealand, Australia, the Netherlands, France, Denmark, Mexico, Great Britain, South Africa, Brazil, and the United States.

[TABLE 2 about here]

The average areas we report in Table 2 show that in many countries the average geographical unit is quite a bit larger than 1 km^2, but this is largely due to areas at the periphery of cities that are outliers in terms of surface. These areas are outliers, but we include them because they tend to have a small cumulative population and therefore have little influence on entropy, which weighs unit population. The method we use simply divides large units into cells and allocate the population of that unit proportionally (that is, each cell replicates the distribution of the larger unit divided by the portion of that unit's area in that cell).

The second issue of note is the method of collection. The complicated nature of income data means that they are often collected through surveys administered to a sub-sample of the entire census population (e.g., Canada) or on a yearly-basis to create representative samples over an extended period of time (five years in the cases of France and the United States). Most countries in our sample, however, collect income or class information as part of the census. This gives us greater confidence in our results as the sampling procedure can introduce significant bias in the index. Reardon, Bischoff, Owens, and Townsend (2018) found that the sampling rate could be especially significant where unit populations are small and when the sample from which estimates are drawn have a low sampling rate. This might be a source of concern primarily in the United States and Brazil where sampling rates can be quite low. However, because we use segregation indexes that aggregate individual units, the sampling errors are reduced.

The definition of income is another source of potential bias. All the countries report household income, but some only provide total income (e.g., South Africa and the United States) which may not reflect disposable income. Some countries report both total and after-tax and redistribution income (e.g., Canada and France). Based on these data we were able to compare the segregation measure using one or the other and found a near perfect correlation of 0.99. Finally, the United Kingdom and Ireland only report occupational classes, which we treat as a proxy (Tammaru et al., 2019).

A last potential issue that we face is related to discrepancies in geographic coverage. Countries like the United States and the United Kingdom make data available at a small scale for the entire country, making it possible to include all units within the boundaries of FUAs. However, countries like Canada, France, the Netherlands, and Mexico have more limited coverage. There is a trade-off between coverage and accuracy. In the cases of Canada, France, and the Netherlands we combined the main geographical units with the next smallest one to create a more complete dataset. The inclusion in the case of Canada is unlikely to make much of a difference as most of the population is included in the census tracts counts, but is more important in the cases of France and the Netherlands that have a strict cut-off for the number of people that need to be in a unit before income data are released. In both cases, the inclusion of municipal data (which tend to be similarly sized to many denser tracts within urban cores) compensates for the more limited data coverage.

In the cases of Mexican cities, data necessary to fill in those gaps are missing. Much like Canada, however, the limited coverage is not too much of a concern because data cover most of the urbanized area and close to 75% of the total FUA population on average. Nonetheless, especially in the context of a country with large disparities and difficult to measure income levels, this missing data could possibly lead to bias and should be taken into consideration when interpreting the results.

3. COMPARING SEGREGATION ACROSS COUNTRIES AND CITIES

3.1. Where is segregation prevalent?

There is a substantial variation in the levels of income segregation across cities both within and between countries. As depicted in Figure 2, countries sort into two groups. One includes Brazil, the United States, and South Africa. Those countries have considerably higher levels of segregation than the other countries in the sample. Their national level is between 0.1 and 0.15 compared to the rest of the sample which is closer to 0.05. In addition, while there is a correlation between the number of cities in each country and the variation across cities, countries like France and Mexico have a tighter distribution than Canada and the United Kingdom.

In the case of the United Kingdom and the United States, the data allow us to assess segregation levels in two points in time.[4] From such a time perspective, there is a noticeable increase levels of segregation in UK cities between 2001 and 2011 (a little over 60%). However, such an increase originates from a small number of cities, such as Leeds, Liverpool, Manchester, and Sheffield, where segregation levels increased substantially. During the same period, segregation decreased in London and Newcastle. Interestingly, the UK-wide standard deviation for the level of segregation across cities nearly doubled from 0.009 to 0.017. This indicates that cities have been following different patterns in spatial inequalities, resulting in a higher heterogeneity in income segregation levels. Noteworthy is that in nearly half of the UK cities segregation increased less than the national average and remained fairly constant.

While changes over time in the United States must be interpreted with caution due to different sampling strategies in 2000 and 2014, the data show a relatively uniform and modest increase in the variation of segregation across cities between 2000 and 2014 (around 3%). The variation in segregation across cities remained almost stable, as indicated by the standard deviation of segregation levels changing from 0.017 to 0.018. Over the same period, twenty-eight out of sixty-two US cities saw an increase in segregation, while 20 saw a decrease. The remaining 22 cities were stable, as changes in their absolute levels of segregation are negligible.

[FIGURE 2 about here]

The cities in South Africa, United States, and Brazil point to the importance of history and inequality in explaining levels of segregation. Many of these cities have among the highest inequality levels in the world as well as histories of segregation (Christopher, 2005; Massey & Denton, 1993; Reardon & Bischoff, 2011; Telles, 2004). In addition, in all three cases, income segregation intersects with racial and ethnic segregation in such a way that segregation levels are compounded (Marteleto, 2012; Nightingale, 2012).

Despite high levels of inequality, segregation in Mexico appears relatively low. Some features of income data in Mexico may partially explain this pattern. While the small area data covers most households in urban areas, there might be some gaps in data collection as surveys likely leave out the most disadvantaged. This possibility is supported by the New Zealand data, which have measured how many households did not answer to the underlying survey. These non-response issues suggest that reporting may be lower in low income areas. If a similar systematic underreporting occurred in the case of Mexico, this would result in a possible downward bias. It might also be the case that Mexico has a different pattern of segregation with respect to other countries. Research on specific cities suggest that segregation, especially at the scale of the current analysis, is generally low among low and middle-income households and higher among high income groups (Aguilar & Mateos, 2011; Monkkonen, Comandon, Montejano Escamilla, & Guerra, 2018). This is consistent with the results discussed in the following paragraphs and suggests that Mexico has a different pattern of spatial clustering of households in space with respect to most countries analyzed here.

Cities in Australia and New Zealand as well as in Denmark, France, and the Netherlands have relatively low levels of segregation compared to cities in Canada and in the United States (Figure 2). In contrast to the Anglo-Saxon countries, these countries have low inequality levels, especially the Netherlands and Denmark. The similar levels of segregation across cities in these three countries suggest that the link between inequality and segregation is not straightforward, as documented by Scarpa (2015) in the case of Malmö, Sweden. While there is evidence that segregation levels tend to be higher in contexts characterized by high income inequalities, many other factors could be at play – both at national and local level – to explain differences in segregation levels across cities. The next sub-section provides details on how segregation patterns change with income groups, while the following section analyses more deeply what observable city characteristics tend to be associated with higher segregation levels.

3.2. Who tends to be more segregated?

The average levels of segregation of cities described in the previous section account for the way all income groups are distributed in space. Consequently, those indicators should be interpreted as an overall measure of segregation patterns across all different income quantiles or, as Roberto (2015) argues, a measure of diversity. However, such indicators may overlook high heterogeneity of spatial inequalities along the income distribution. The curves depicted in Figure 3 show how segregation levels changed for different income groups in three randomly selected cities per country.[5] More specifically, the curves reflect how segregated a specific income group is from the rest of the population in the city. The leftmost data point of the curves reflects the segregation of households living below the first income threshold relative to the rest of their city's population. The rightmost data point on the same segregation curve shows segregation of households in the highest income bin from other households.[6] In the case of the United States (Figure 3 j) and New Zealand, we show cities in different points in time, to exploit visually the time variation, when available. The figure reveals increased levels of segregation of cities in both countries.

Overall, rich households tend to be more segregated than poorer ones, a pattern consistent with existing literature on specific countries. Reardon et al. (2006), for example, show that for several metropolitan areas in the United States, segregation levels at the lowest income levels are higher than for most of the population, but remain much lower than for the richest households. Similar patterns have also been observed outside the United States. Floch (2017) finds higher levels of segregation for the highest income group across twelve French cities. One possible explanation for such patterns is that poorer residents might end up in more diverse neighborhoods than the richest residents. People in the most affluent group have the widest range of location choices and, by extension, have a greater ability to live close to other people of similar income levels. The process of residential sorting that yields more homogeneous neighborhoods at the top of the income ladder contributes to the lower levels of segregation at all other income levels. As high-income households bid up the price of the most desirable locations, all other household have to compete for the remaining spaces.

More detailed information on the income distribution, especially for the bottom income groups, makes it possible to analyze segregation patterns more precisely. Sharp peaks at the extremes should be expected when the category is very small. Among the countries considered in this work, Australia and South Africa include an income category for households with no reported income. Households falling in this category show extremely high segregation levels in the case of Australia (Figure 3, panel a), where very few people have no reported income. When a very small group concentrates in the same area, this will result in high observed segregation. South African cities (Figure 3, panel i), on the other hand, have lower levels of segregation at the lower extreme, consistent with the patterns recorded in some cities in the United States. However, South African cities have a different set of confounding factors. The category for no income is either the largest or one of the largest for each city, hiding much variation, especially when considering the significance of the informal economy.

[FIGURE 3 about here]

While cities in nearly all examined countries have higher segregation levels at higher income levels, this does not happen everywhere. Figure 4 shows how segregation levels compare between the highest and lowest 20% of the income distribution in each country.[7] Most countries show a consistent pattern of upper quantile being considerably more segregated than the lower quintile. Ratios between these deciles are for most cities between 0.6 and 0.8. In some countries, such as in France, segregation levels are similar at both ends of the income distribution. The Netherlands and Denmark – two countries with overall low levels of income inequalities and low levels of segregation – are the only countries where segregation tends to be higher for the poor than for the rich.[8] In these countries, the higher segregation of poor households might reflect the spatial organization of social housing. In the Netherlands, for example, social housing is sizeable (Elsinga, Haffner, and Van Der Heijden, 2008) and, if concentrated in specific neighborhoods, can yield higher spatial concentration of the low-income groups.

The remaining countries show a diversity of patterns, but also some consistencies. For example, cities in Canada and New Zealand show the lowest segregation at the bottom of the income distribution and then a steadily increase for richer households. French and Australian cities, on the other hand, tend to have much higher segregation at the top of the income distribution combined with almost constant segregation levels for the other income groups. The United States falls somewhat in-between with a steady, but steep increase from the sixth decile of the income distribution. Over time, segregation patterns in US cities seem to have increased for almost all income groups, meaning that neighborhoods have evolved towards a relatively higher degree of socio-economic homogeneity rather than towards more mixed patterns. There is more variation at the bottom of the income distribution from Canada's steep decrease to the United States and Australia's upward curvature.

[FIGURE 4 about here]

4. WHAT DRIVES INCOME SEGREGATION?

We now turn our attention to analyzing the relationship between income segregation and a set of city characteristics. Given some of the limitations of our data in terms of number of observations, time coverage and sufficiently large set of controls, results should not be interpreted as evidence for causal relationships. The tests still shed light on the factors associated with higher levels of segregation. To do so, we rely on an econometric approach and estimate the following Equation:

$$ln(SE_t) = \delta_0 + \delta_1 \times City\,Size_t + \delta_2 \times Government_t + \delta_3 \times Economy_t + \delta_4 \times Demography_t + \sum_j (\mu_j \times Country\ dummy_j) + \sum_t (e_t \times Year_t) \quad (3)$$

where the main dependent variable is the log of the measure of income segregation (spatial entropy at 1 km^2 scale).

We estimate Equation (3) using gradual specifications. Initially, we consider a set of explanatory variables that is grouped into four categories (city size, government, economy, and demography) based on previous empirical and theoretical studies. Thereafter, we include additional measures of urban form and segregation and investigate possible differences in estimates for the poor and the rich. The definition of each variable as well as its source, often the OECD Metropolitan Database, is provided in Appendix I (Table A1).

We estimate Equation (3) using Ordinary Least Squares (OLS) while pooling data for 107 cities in the years circa 2001 and 119 cities in the years circa 2011. The whole set of cities included in the analysis is reported in Table 3. Those cities represent a smaller sample with respect to the entire set of cities for which we assessed segregation levels. The reason why not all cities were included in the regression analysis lies on the lack of the independent variables, notably for South Africa and Brazil, which are not covered by the OECD Metropolitan Database. Based on each specification, we estimate two models; these models either do, or do not, account for possible omitted spatial processes. Possible omitted spatial processes are accounted for at the national scale through the inclusion of country dummies. At the consistent scale of countries, a one-unit shift in each of the observed city-level variables is plausible. We acknowledge that the dummies may absorb information about segregation effects in the case of the few observed cities that are the only cities in their respective countries, such as Dublin and Copenhagen.

4.1 The role of city size, government, economy and demography

The first out of the four literature-led categories of variables includes population size in order to capture the relationship between inequality and city size. We expect this relationship to be positive as, in recent decades, relatively large wage dispersion has been documented for larger cities (Baum-Snow and Pavan, 2013).

As is usual, we add controls for the economy of the city. The employment rate controls for the state of the labor market (and the city capacity to integrate low-skill workers) and, in general, for the level of development of the city (Pendall and Carruthers, 2003). We also add labor productivity, which is related to the literature on inequality and city size (Baum-Snow and Pavan, 2013). We hypothesize that higher productivity translates to higher wages, especially for skilled workers, leading to greater relative wage differences between skilled and low-skilled workers. This increasing difference in terms of wages may result in higher preferences heterogeneity (e.g., commuting, public goods, and amenities) and in higher willingness for the rich to outbid the poor, therefore exacerbating income segregation. This may in part depend on labor market conditions, such as complementarity between skilled and unskilled workers, as described in Benabou (1993). On the other hand, the effect of higher productivity might work also in the opposite direction, to the extent that an increase quality and quantity of amenities and public services becomes available for the poor following an increase in productivity and wages. Noteworthy is that, when controlling for productivity, we interpret the density effect for a given level of productivity, and vice-versa.

Based on Pendall and Carruthers (2003) and Tiebout's (1956) theory of spatial sorting, we also study the role of local governments. People sort according to their preferences for different types and levels of public goods provision. However, Tiebout's theory has limited implications for the ability of people to exclude others. Income inequality means that high income households can not only outbid lower income households in municipalities with efficient provision of public goods (e.g., lower tax burden for high quality public schools), but also devise strategies to exclude them (Fennell, 2006; Trounstine, 2018). Thus, the higher the number of local governments (e.g., municipalities) within the region relative to the total FUA population, the higher the potential heterogeneity in local taxation schemes, which could foster the sorting of households in space, further pushing segregation.

Finally, following Pendall and Carruthers (2003), Galster and Cutsinger (2007) and Garcia-López and Moreno-Monroy (2018), we controlled for city demography (youth and old age ratios). The idea is that families in different stages of their life cycle have different preferences that might affect their location patterns. For example, young people with children compete for better school districts, leading to more income segregation. Also, these households may necessitate more space and thus seek after cheaper land, thereby separating themselves from older (richer) households without children. While other socio-economic factors (such as race and ethnicity, among others) play a critical role, inconsistent data availability and the scope of the study made it challenging to include these variables.

[TABLE 3 about here]

We now turn to our gradual specifications, based on Equation (3) and as shown in Table 4, which we use to examine the association of relevant variables with income segregation. That is, in Columns 1a-1b, we first include the log of the city population as an explanatory variable. Then, in Columns 2a-2b, we add the log of the administrative fragmentation index (ratio of number of local governments and population). Columns 3a-3b include two variables related to the economy of the city in log form: labor productivity, measured as the ratio between city GDP and total employment, and employment rate, measured as the percentage of city employment over total labor force. Finally, the log of two demographic variables is also added (Columns 4a-4b): the youth dependency ratio, measured as the ratio between the youth population (0–14 years old) over the working age population (15–64 years old), and the old age dependency ratio, measured as the ratio between the population 65 years old and over and the working age population. All these explanatory variables are obtained from the OECD

Metropolitan Database.[9] Recall that each gradual specification is estimated without and with country fixed-effects, as reported in Columns (a) and (b), respectively.

Results in Table 4 show a positive and significant relationship between income segregation and some of our explanatory variables. Across the gradual specifications (1-4), we consistently find that the higher the population, labor productivity, or youth dependency ratio of a city, the higher the degree of (income) spatial segregation. The coefficients for these variables, in contrast to administrative fragmentation, the employment rage and old age dependency, are both statistically and economically significant. Noteworthy is that the estimated effect of population size on segregation levels initially fluctuates depending on whether spatial controls are included, in restricted specifications (1-2), but becomes stable after the inclusion variables that capture the state of the observed cities' economies. Overall, the effect-sizes obtained for the variables in the full specification (4a-4b) are qualitatively similar to those in models (3a-4a) and stable in terms of sign and significance.

[TABLE 4 about here]

4.2 The role of urban form

Following Pendall and Carruthers (2003), Galster and Cutsinger (2007), Combes and Gobillon (2015), and Ahlfeldt and Pietrostefani (2017; 2019), we also consider city size in terms of city density by adding the city land area as a control variable. Existing research shows that density is positively linked to productivity (higher wages), housing prices, rents, access to services, and efficiency of public services. Simultaneously, some of these traits may also lead to higher income segregation. For example, according to Ahlfeldt and Pietrostefani (2019), an increase in density leads to a decrease in net wages (higher wages are more than offset by even higher values of space) which is compensated by higher amenities. Thus, the impact of an increase of density on income segregation is ambiguous and it depends on how amenities are evaluated by the people and how public goods can be replaced by private ones. Then, we may assume that wealthier households have a higher willingness to pay for amenities and public or private services or, at least, a greater capacity to pay for them. This difference in ability to access space can increase income segregation.

We also consider the role of different types of urban forms: monocentric versus polycentric cities, compact versus dispersed cities, and centralized versus decentralized cities. These urban forms are related to different spatial distributions of jobs within cities and, as a result, they might be related to different residential location patterns. For example, McMillen (2001) highlighted that subcenters in a polycentric city enjoy some of the same agglomeration economies (i.e., higher wages) as the central business district (CBD), but offer lower commuting costs and, in particular, lower housing prices for suburban workers. One possible mechanism at play is that lower housing prices allow households of different income groups to compete for housing in the same area and, as a result, we may observe more mixed locations patterns. Thus, we may observe less segregation in polycentric cities (Garcia-López and Moreno-Monroy, 2018). However, if there is also job segregation, with qualified jobs in the CBD, less qualified jobs in the subcenters and non-qualified jobs elsewhere, polycentric cities might lead to higher income segregation.

In Table 5, we focus on alternative measures of urban form. The purpose is to test the effect of different types of urban spatial structures (e.g., monocentric versus polycentric cities, compact versus dispersed cities, or centralized versus decentralized cities). In Column 1, we include the log of the city land area allowing us to interpret the coefficient of our city size variable, the city's total population, in terms of population density. In Column 2, we expand on the relationship to density by substituting the population variable with an explicit measure of population density. In both cases, results for the log of population (Column 1) and the log of population density (Column 2) show that a higher population density is also directly related to a higher degree of income segregation.[10] The positive association between city size and socio-economic segregation is consistent with several other works in the literature (Farley 1991 and Jargowski 1996, among others). However, this association may be spurious due to the lack of small homogenous tracts in small cities where high density areas that meet the statistical standard to form a tract are limited (Krupka, 2007). In this work, we take this possible bias into account by considering only cities over 500,000 inhabitants and by applying dasymmetric mapping (see Appendix I) to smooth income data at a regular and consistent size across cities. In Table 5, Columns 3 and 4, we test whether there are significant differences in segregation levels between monocentric and polycentric cities. We first interact the log of population with a dummy for polycentric cities (and add also this dummy as an explanatory variable) in Column 3. The polycentricity dummy takes on a value of 1 when the number of urban cores within a city is greater than 1 according to the OECD Metropolitan Database. The results imply that the positive and significant relationship between population density and income segregation is related to monocentric cities (0.066). Despite the magnitude of the estimate is much smaller for polycentric cities (0.001 = 0.066 - 0.065), the difference between the two types of spatial structures is not statistically significant. As a result, the above mentioned positive relationship between segregation and density applies to both types of urban form.

In Column 4, we split the overall city population between central city population (people living in the urban cores) and suburban population (people living in the commuting zones of the urban cores). Results for these variables show that the effect of a larger central population is not significantly different between monocentric and polycentric cities and is positively related to higher income segregation levels (0.064). A higher suburban population is found to be positively related to income segregation only in monocentric cities (0.011) as the relationship between suburban population and income segregation is negative in polycentric cities (-0.065 = 0.011 - 0.076).

Finally, since the two types of urban spatial structure may also be related to differences in productivity levels, we an additional interaction between the polycentricity dummy and labor productivity (Table 5, Column 5). This interaction is significant and negative and, in absolute values, higher than the coefficient for monocentric cities. Overall, results suggest that cities with higher labor productivity levels tend to have higher income segregation levels in monocentric cities (0.255), and lower income segregation in polycentric cities (-0.044 = 0.255 - 0.299). One potential mechanism underlying the negative association between polycentricity and segregation levels might be lower commuting (or congestion) and housing costs in polycentric cities.

In Columns 6 through 9 of Table 5, we analyze the effect of other measures of urban spatial form. Departing from the specification in Column 1, we add a measure for the degree of spatial concentration of the population, the Theil's entropy index, in Column 6. The index is consistent with that used by Tsai (2005) and it is a Theil entropy index computed using population values of all local units within a FUA. Differently from other indicators of concentration, entropy measures are more comparable across cities as they are not influenced by the number of local units within each urban area.

In the column 7 we added the average weighted distance to CBD to measure the degree of spatial centralization of the population, meaning the degree of concentration of the population in the single main center; Both spatial concentration and decentralization are added in Column 8; and, departing from the specification in Column 8, the number of population centers (based on the OECD variable 'polycentricity') are added in Column 9. Since the Theil's concentration index ranges between 0 and 1, with 0 indicating perfect concentration (see Veneri 2018 for further explanations), results for these regressions show that cities with lower (higher) spatial concentration indexes are related to lower (higher) levels of income segregation. Results for the average distance to CBD show that less (more) centralized cities are related to lower (higher) segregation levels. Finally, in line with the results in Columns 3 to 5, a higher (lower) number of city centers (with a minimum of 1 for monocentric cities, and more than 1 for polycentric cities) are related to lower (higher) levels of income segregation. All these results clearly show that urban spatial structure – measured in terms of the patterns of population distribution across the urban space – and income segregation are deeply interrelated.

Finally, in all regressions in Table 5 the variables labor productivity and youth dependency ratio keep showing a positive and significant relationship with income segregation. The only exception is the abovementioned interaction of labor productivity and the dummy for polycentric cities (Column 5).

[TABLE 5 about here]

4.3 Different spatial scales

As a robustness check, we estimate Equation (1) using our income segregation measure computed at different spatial scales. This new set of regressions departs from the baseline specification in Table 5 Column 9.

Table 6 reports results when using a segregation index computed at the 500 meters spatial scale in Column 1, 2,000 meters in Column 2, and of 4,000 meters in Column 3. We use an a-spatial segregation index computed using the smallest available intra-city unit (i.e., municipalities, wards, or census tracts) in Column 4. In all regressions, results are not significantly different from the preferred specification in Table 4 (Column 9) and further confirm that cities with a higher (lower) population density, degree of spatial concentration and centralization, labor productivity, and youth dependency ratio are associated with higher (lower) levels of income segregation.

[TABLE 6 about here]

4.4 The poor vs. the rich

Finally, we estimate Equation (1) for different groups of population according to their income level. The idea is to test whether the above studied 'average' relationships remain as such across the income distribution and, in particular, for the lowest and highest income levels (the poor and the rich).

Table 7 reports results for the lowest income groups ('the poor') and for the highest income groups ('the rich') using our income segregation index computed only for the 10th and 20th income percentiles (Columns 1 and 2) and for the 80th and 90th income percentiles (Columns 3 and 4), respectively.

This new set of results clearly shows that the level of segregation of the poor is only (positively) related to the labor productivity of the city and to the degree of spatial centralization. On the other hand, the results for the highest income levels are quite similar to the 'average' results and show that the segregation of the rich are (positively) related to the city size (population density), the degree of spatial centralization of the city, the labor productivity and the youth dependency ratio.

[TABLE 7 about here]

The econometric results in Table 7 further show that specific characteristics of cities in terms of size, economic development, demographic composition and spatial configuration can explain the variance observed in segregation levels across different countries. Larger and denser cities tend to be more segregated as well as cities with higher labor productivity. At the same time, cities with higher proportions of elderly population tend to have higher income segregation levels.

The econometric results here also show that the spatial configuration of cities can explain the observed patterns of income segregation. First, a polycentric spatial configuration seems to mitigate the association between segregation and density, as well as between segregation and urban size. Second, cities where larger share of the population is concentrated around the main center show on average higher income segregation levels, both in a monocentric and polycentric configuration. Third, while a higher labor productivity is associated to higher income segregation levels in monocentric cities, the opposite relationship is found for polycentric cities. In our sample, after controlling for other characteristics, cities with more suburbanized and decentralized population have lower degree of income segregation, on average. Finally, the econometric results show that income levels matter: labor productivity and the degree of spatial decentralization of the city are related to income segregation levels of the poor and the rich. Other abovementioned determinants are only related to the segregation level of the rich.

5. CONCLUSIONS

In well-functioning and inclusive cities, people of all backgrounds, while physically separated, can access opportunities and high-quality services that ensure socio-economic mobility along the income ladder and, as a consequence, across neighborhoods of residence. While the literature on the neighborhood effect is still not conclusive, a comprehensive assessment of patterns of income segregation in cities can help framing future research in a more informed context.

This paper has provided an assessment of income segregation within cities in twelve countries, trying to maximize international comparability. In the twelve observed countries, within-city segregation of households varies considerably across cities – even within the same country. This finding opened the way to identify regularities between the characteristics of cities and their patterns of segregation. Results shows that the location patterns of households at the extremes of the income distribution, the most and least affluent households, are often reflected in the observed segregation levels. In most cities, segregation is driven by location patterns of the richest households, although there are exceptions in the least unequal countries.

Our study analyses the city-characteristics explaining the observed variation in segregation levels. Given the empirical framework and the data limitation in such an international comparative analysis, results should not be interpreted as causal evidence. However, the results of this paper allow to broaden the debate on how income segregation can be linked to different institutional, geographical and economic settings, which could be taken into account by policy makers. Our study highlights that the highest levels of segregation occur in large, young, and highly productive cities. In those cities, residents at the top of the income ladder might outbid the poor in a way that is more concentrated in space, triggering higher segregation levels. In more affluent cities, under the hypothesis that one type of social group is preferred as neighbors by all, see Becker and Murphy (2003), higher competition for better locations might increase further the willingness of the most affluent to separate from lower income groups, further exacerbating segregation levels.

Our results also suggest that those cities tend to have relatively large populations in their urban core(s). A relatively suburbanized polycentric urban structure, on the other hand, seems to show lower productivity levels. This result might suggest that polycentric urban structures reflect an adjustment of the city to high congestion and land prices in the main center, combining the productivity advantages of density with lower segregation levels.

Some of the city characteristics taken into account in this study are less powerful to explain differences across cities in the segregation of the poor. For the latter, there are several factors that might be at play and for which further research is needed. More specifically, land-use regulation, transport, housing, and education are important policy domains that might be able to affect segregation and the capacity for the least affluent households to access the opportunities offered by the city.

Our study also highlights the importance of advancing the measurement and understanding of patterns of inequalities within cities. Methods for extrapolating entire income distributions from available data have made great stride, but hit a wall where data are unavailable at the extremes or spatially. International comparative research is an opportunity to harness the benefits of several data sources to refine methods for measuring segregation. Australia, for example, offers detailed data over much of the income distribution at a high spatial resolution, which could support efforts to understand the correlation of income levels with the built environment. In a growing number of countries, high-quality longitudinal data on household income become available and represent an important source of information to understand the patterns of spatial inequalities within cities. This can also help to test some of the assumptions made in comparing countries that use different methods of collection and lead to better tools to mitigate the influence of such differences.

[1] For example, McAllen, USA, one of the smaller regions in the sample, includes 119 geographical units, which near the minimum number of units to derive plausible estimates of the spatial structure (Reardon, 2008). The amount of information available about the spatial structure drops quickly for smaller regions. For example, Corpus Christi (Texas, United States) with a population just under 500,000 has only 92 units, several of which are very large tracts that offer limited information.

[2] The index does not differentiate between categories. That is, if all individuals in the neighborhoods were in a single income category, it would not matter which income category they were in because the neighborhoods would be homogenous regardless (Reardon, 2009). As a measure of evenness, the entropy index depends only on the relative distribution among the categories.

[3] The use of social class as a proxy for income has a precedent in comparative work on segregation (Tammaru, Marcinzak, Aunap, Van Ham, & Janssen, 2019). The correlation between social class (based on the kind of work one does, in the case of the UK and Ireland) and income tends to be very high and the implication for segregation similar. However, some caveats are required. Some occupational categories – e.g., self-employed individuals who may have a fledging small business or own a successful law practice – can include many different levels of income. Categories will also have much variation within them because of differences in experience level, for example. In the analysis we shy away from putting these countries on the same level as others but include them for reference and to add to the coverage.

[4] The geographical unit for both years is not consistent, which introduces some bias. In the case of the United States, the increase is too small to be significant, but in the United Kingdom the large increase reflects more than the change in the underlying data.

[5] While the cities were randomly selected, we checked against a larger sample of cities to make sure the pattern were representative of the national trends. While there are deviations, the curves are illustrative of the most common shape. In countries with only a few cities, deviations can suggest a lack of general pattern. These deviations are interesting in themselves and one of the goals of the paper is to push for more research into the variation within country.

[6] In most countries the bin for 'no income' does not exist and nor does the bin that would include all income above the highest threshold and is as such not comparable across countries. Due to this lack of data the segregation of households at the lower and upper tails of the income distribution remains unobserved for most countries (Australia is an exception as it has a 'no-income' bin.

[7] This approach is chosen to minimize the influence of data at the extremes and because all countries have data available at the least between the 20th and 80th percentiles, making the data more comparable. Cut-offs are chosen by calculating the distance between the chosen deciles and the percentiles that correspond to each income bin. So, for example, if there is an income bin that represents 8% of the population and the next two bring the cumulative population to 15% and 26% respectively, the first is chosen as the first decile and the second as the second decile.

[8] However, consider that due to the lack of data for these countries on the upper and lower ends of the income distribution segregation at the 'true' extremes of income-classes remains unmeasured.

[9] https://stats.oecd.org/Index.aspx?DataSetCode=CITIES

[10] One may note that in Columns 1 and 2, the estimates are positive and significant for both population density and productivity. When productivity is held constant, population density levels may vary with the underlying nature of cities. For example, at a given level of productivity, the source of this productivity may be in sectors that are associated with either centripetal or centrifugal patterns of agglomeration or density (McCann, 2008). This coherence between particular sectors and urban density can be associated with segregation, for instance through a channel such as the complementarity of skilled and unskilled workers (c.f., Benabou, 1993). We further examine the association of segregation with specific forms of urban density, monocentric and polycentric structure, below.

REFERENCES

Aguilar, A. G. and P. Mateos. 2011. "Diferenciación sociodemográfica del espacio urbano de la Ciudad de México," *Eure (Santiago)*, *37*(110), 5–30.

Ahlfeldt, G. and E. Pietrostefani 2019. "The economic effects of density: A synthesis," *Journal of Urban Economics* 111, 93–107.

Ahlfeldt, G. and E. Pietrostefani 2017. "The compact city in empirical research: A quantitative literature review," SERC discussion papers, Spatial Economics Research Centre, LSE.

Baum-Snow, N. and R. Pavan 2013. "Inequality and city size. *The Review of Economics and Statistics* 95(5), 1535–1548.

Benabou, R. (1993). "Workings of a city: location, education, and production", *The Quarterly Journal of Economics*, 108(3), 619–652.

Bischoff, K. 2008. "School district fragmentation and racial residential segregation: How do boundaries matter?" *Urban Affairs Review*, *44*(2), 182–217.

McCann, P. (2008). "Globalization and economic geography: The world is curved, not flat." *Cambridge Journal of Regions, Economy and Society*, 1(3), 351–370.

Chetty, R. and N. Hendren 2018. "The impact of neighbourhoods on intergenerational mobility I: Childhood exposure effects," *The Quarterly Journal of Economics* 133(3), 1107–1162.

Christopher, A.J. 2001. "Urban segregation in post-apartheid South Africa," *Urban Studies*, 38(3), 449–466.

Christopher, A. J. 2005. "The slow pace of desegregation in South African cities, 1996-2001," *Urban Studies*, *42*(12), 2305–2320.

Combes, P.P. and L. Gobillon 2015. "The empirics of agglomeration economies," In Gilles Duranton, J. Vernon Henderson, and William Strange (eds.) *Handbook of Regional and Urban Economics Volume 5*. Elsevier, 247–348.

Ellickson, R. C. 2006. "The puzzle of the optimal social composition of neighborhoods," In *The Tiebout model at fifty: Essays in public economics in honor of Wallace Oates*.

Elsinga, M., M Haffner, and H. Van Der Heijden. 2008. "Threats to the Dutch unitary rental market," *International Journal of Housing Policy*, *8*(1), 21–37.

Enos, R. D. 2017. "*The spaces between us: social geography and politics*," Cambridge: Cambridge University Press.

Farley, R. 1991. "Residential segregation of social and economic groups among blacks, 1970–1980," in: C. Jencks and P.E. Peterson (Eds) *The Urban Underclass*, 274–298. Washington, DC: Brookings Institution.

Fennell, L. A. 2006. "Exclusion's attraction: Land use controls in Tieboutian perspective. In W. Fischel (Ed.) ," *The Tiebout model at fifty: Essays in public economics in honor of Wallace Oates*.

Galster, G., and J. Cutsinger. 2007. "Racial settlement and metropolitan land-use patterns: Does sprawl abet-black-white segregation?" *Urban Geography* 28(6), 516–553.

Garcia-López, M.A. and A.I. Moreno-Monroy. 2018. "Income segregation in monocentric and polycentric cities: Does urban form really matter?" *Regional Science and Urban Economics*, 71, 62–79.

Jargowski P. A. 1996. "Take the money and run: economic segregation in US metropolitan areas," *American Sociological Review*, 61(6), 984–998.

Johnston, R. and K. Jones. 2010. "Measuring segregation—A cautionary tale," *Environment and Planning A*, *42*(6), 1264–1270.

Johnston, R., Poulsen, M., and J. Forrest. 2010. "Moving on from indices, refocusing on mix: On measuring and understanding ethnic patterns of residential segregation," *Journal of Ethnic and Migration Studies*, *36*(4), 697–706.

Klinthäll, M. and S. Urban. 2016. "The strength of ethnic ties: Routes into the labour market in spaces of segregation," *Urban Studies*, 53(1), 3–16.

Kramer, M. R., Cooper, H. L., Drews-Botsch, C. D., Waller, L. A., and C. R. Hogue. 2010. "Do measures matter? Comparing surface-density-derived and census-tract-derived measures of racial residential segregation," *International Journal of Health Geographics*, *9*(1), 1.

Krupka, D. J. 2007. "Are big cities more segregated? Neighbourhood scale and the measurement of segregation, *Urban Studies*, 44(1), 187–197.

Li, H., Campbell, H. and S. Fernandez. 2013. "Residential segregation, spatial mismatch and economic growth across US metropolitan areas," *Urban Studies*, 50(13), 2642–2660.

Maloutas, T. and K. Fujita. 2012. "*Residential segregation in comparative perspective: making sense of contextual diversity*," Burlington, VT: Ashgate.

Marteleto, L. J. 2012. "Educational inequality by race in Brazil, 1982–2007: Structural changes and shifts in racial classification," *Demography*, *49*(1), 337–358.

Massey, D. S. and N. A. Denton. 1988. "The dimensions of residential segregation," *Social Forces*, *67*(2), 281–315.

Massey, D. S. and N. A. Denton. 1993. "*American apartheid: segregation and the making of the underclass*," Cambridge, Mass: Harvard University Press.

Merry, M. 2016. "*Equality, citizenship, and segregation: A defense of separation*," Dordrecht: Springer.

McMillen, D.P. 2001. "Nonparametric employment subcenter identification," *Journal of Urban Economics* 50(3), 448–473.

Monkkonen, P., Comandon, A., Montejano Escamilla, J. A., and E. Guerra. 2018. "Urban sprawl and the growing geographic scale of segregation in Mexico, 1990–2010," *Habitat International*, *73*, 89–95.

Morrison, P.S. 2015. "The inequality debate. The neglected role of residential sorting," *Policy Quarterly*, 11/2, 72–79, Retrieved from http://igps.victoria.ac.nz/publications/files/05ff5e0d1bb.pdf.

Mossay P., and Picard, P.2019. "Spatial segregation and urban structure," *Journal of Regional Science*, 59(3), 480-507.

Nieuwenhuis, J., Tammaru, T., van Ham, M., Hedman, L., and D. Manley. 2020. "Does segregation reduce socio-spatial mobility? Evidence from four European countries with different inequality and segregation contexts," *Urban Studies* 57(1), 176–197.

Nightingale, C. H. 2012. "*Segregation: A global history of divided cities*," Chicago; London: The University of Chicago Press.

Novara, M., Loury, A., and A. Khare. 2017. "*The cost of segregation: Lost income, lost lives, lost potential*," Chicago, IL: Metropolitan Planning Council.

Openshaw, S. 1984. "*The modifiable areal unit problem*," Norwich: Geobook.

Oreopoulos, P. 2003. "The long-run consequences of living in a poor neighborhood," *The Quarterly Journal of Economics* 118, 1533–1175.

Pendall, R. and J. Carruthers, 2003. "Does density exacerbate income segregation? Evidence from

US metropolitan areas, 1980 to 2000," *Housing Policy Debate* 14(4), 541–589.
Reardon, S. F. 2009. "Measures of Ordinal Segregation," in Y. Flückiger, S. F. Reardon and J. Silber (Eds.), *Occupational and residential segregation*, Vol. 17. 129-155), Bingley: Emerald.
Reardon, S. 2011. "*Measures of income segregation*. Retrieved from Stanford Center on Poverty and Inequality website: http://inequality.stanford.edu/sites/default/files/reardon_measures-income-seg.pdf
Reardon, S. F., Firebaugh, G., O'Sullivan, D., and S. Matthews2006. "A new approach to measuring socio-spatial economic segregation. *Measurement of segregation: New directions and results*," Presented at the General Conference of The International Association for Research in Income and Wealth, Joensuu, Finland. Retrieved from http://www.iariw.org/papers/2006/reardon.pdf
Reardon, S. F., Bischoff, K., Owens, A., and J. B. Townsend. 2018. "Has ncome segregation really increased? Bias and bias correction in sample-based segregation estimates," *Demography*, 32.
Reardon, S. F. and K. Bischoff2011. "Income inequality and income segregation," *American Journal of Sociology*, 116(4), 1092–1153.
Roberto, E. 2015. "The divergence index: A decomposable measure of segregation and inequality," *ArXiv:1508.01167 [Physics, Stat]*. Retrieved from http://arxiv.org/abs/1508.01167
Scarpa, S. 2015. "The impact of income inequality on economic residential segregation: The case of Malmö, 1991-2010," *Urban Studies*, 52(5), 906–922.
Schelling, T. C. 1971. "Dynamic models of segregation," *The Journal of Mathematical Sociology*, *1*(2), 143–186.
Tammaru, T., Marcińczak, S., Aunap, R., van Ham, M., and H. Janssen. 2019. "Relationship between income inequality and residential segregation of socioeconomic groups," *Regional Studies*, 1–12. https://doi.org/10.1080/00343404.2018.1540035
Tarozzi, A. and A. Deaton. 2009. "Using census and survey data to estimate poverty and inequality for small areas," *The Review of Economics and Statistics*, 91(4), 773–792.
Telles, E. 2004. "*Race in another America: The significance of skin color in Brazil*," Princeton, NJ: Princeton University Press.
Tiebout, C. M. 1956. "A pure theory of local expenditures," *Journal of Political Economy*, 64, 416–424.
Tammaru T, Marcińczak S, Van Ham M., Musterd, S. (eds) 2016. "*Socio-Economic Segregation in European Capital Cities*," East Meets West. London: Routledge.
Trounstine, J. 2018. "*Segregation by Design: Local Politics and Inequality in American Cities*,"
Tsai Y.-H. (2005), Quantifying Urban Form: Compactness versus 'Sprawl', *Urban Studies* 42: 141-61.
Vang, Z.M. 2010. "Housing Supply and Residential Segregation in Ireland," *Urban Studies*, 47(14), 2983–3012.
Veneri, P. 2018. "Urban spatial structure in OECD cities: Is urban population decentralising or clustering?" *Papers in Regional Science* 97(4), 1355–1374.

TABLES

Table 1: Illustration of ordinal segregation index (adapted from Reardon, 2009).

Income category	Neighborhood				
	1	2	3	4	Total
	Scenario A				
0-50	50	40	10	0	100
51-100	40	30	20	10	100
101-150	10	20	30	40	100
151-200	0	10	40	50	100
Total	100	100	100	100	400
	Scenario B				
0-50	40	30	20	10	100
51-100	50	40	10	0	100
101-150	0	10	40	50	100
151-200	10	20	30	40	100
Total	100	100	100	100	400

Note: The table represents a stylized city of 400 people divided into four neighborhoods (columns) and four income categories (rows) of 100 people each. While the multi-group index is 0.198 in both scenario, the ordinal index is 0.277 in Scenario A and 0.162 in Scenario B.

Table 2: Summary of source and key measures

Country		Census authority	Avg. household per unit	avg. area km2 (median)	Income bins	Sampling rate*
Australia	2010	Australian Bureau of Statistics	134	1.57 (0.17)	15	100%
Brazil	2010	Instituto Brasileiro de Geografia e Estatística	206	5.18 (0.07)	10	50% to 5% depending on city size
Canada	2011	Statistics Canada - National Household Survey	2007	4.25 (1.6)	13	25%
Denmark	2013	Dansk Demografisk Database	1674	9.2 (6.9)	5	100% from Administrative record
France	2011	Institut National de la Statistique et des études économique	1318	5.62 (0.76)	11	40% over 5-year period
Ireland[a]	2006	Central Statistics Office	321	0.27 (0.05)	8	100%
Ireland[a]	2011	Central Statistics Office	98	0.77 (0.05)	8	100%
Mexico	2000	Instituto Nacional de Estadística y Geografia (INEGI)	654	0.5 (0.33)	12	100%
Netherlands[b]	2008	Statistics Netherlands	1637	2.82 (0.75)	5	> 32%
New Zealand	2001-13	Statistics New Zealand	906	3.07 (1.5)	6	100%
South Africa	2011	Statistics South Africa	189	0.79 (0.11)	12	100%
United Kingdom[a]	2001	Office for National Statistics	109	0.37 (0.05)	8	100%
United Kingdom[a]	2011	Office for National Statistics	228	0.35 (0.05)	8	100%
United States	2000	U.S. Census Bureau	1693	19.1 (2.4)	16	15%
United States[c]	2014	U.S. Census Bureau - American Community Survey	1681	17 (2.4)	16	8% over 5-year period

Notes: *All countries with 100% sampling rate refer to collection of the data as part of the census unless otherwise indicated. In all cases, the sampling strategy is prone to bias due to systematic differences in response rate across the income distribution.

a. For Ireland and the United Kingdom, we use social class rather than income. The categories for Ireland are: A = Employers and managers, B = Higher professional, C = Lower professional, D = Non-manual, E = Manual skilled, F = Semi-skilled, G = Unskilled, H = Own account workers, In the United Kingdom, the categories are: 1 Higher managerial, administrative and professional occupations, 2Lower managerial, administrative and professional occupations 3Intermediate occupations, 4 Small employers and own account workers, 5 Lower supervisory and technical occupations, 6 Semi-routine occupations, 7 Routine occupations, 8 Never worked and long-term unemployed

b. The initial sampling rate for the observed cities is 32% of all inhabitants of the age of 15 or older. Subsequently, household members are also added into the sample.

c. For more information on the American Community Survey (ACS), please consult https://www.census.gov/programs-surveys/acs/methodology/design-and-methodology.html and https://www2.census.gov/programs-surveys/acs/tech_docs/accuracy/MultiyearACSAccuracyofData2016.pdf

Table 3: Countries, cities and years

Country	City (FUA)	Year
Australia	Adelaide, Brisbane, Gold Coast-Tweed Heads, Melbourne, Perth, Sydney	2010
Brazil*	Agreste, Aracaju, Baixada Santista, Belém, Belo Horizonte, Brasilia, Campina Grande, Campinas, Cariri, Curitiba, Florianópolis, Fortaleza, Foz do Rio Itajaí, Goiânia, Grande São Luís, Grande Teresina, João Pessoa, Londrina, Maceió, Maringá, Natal, Norte/Nordeste Catarinense, Petrolina/Juazeiro, Porto Alegre, Recife, Rio de Janeiro, São Paulo, Salvador, Vale do Aço, Vale do Itajaí	2006
Canada	Calgary, Edmonton, Hamilton, Montreal, Ottawa-Gatineau, Quebec, Toronto, Vancouver, Winnipeg	2011
Denmark	Copenhagen	2013
France	Bordeaux, Grenoble, Lille, Lyon, Marseille, Montpellier, Nantes, Nice, Paris, Rennes, Rouen, Saint-Étienne, Strasbourg, Toulon, Toulouse	2011
Ireland	Dublin	2006, 2011
Mexico	Acapulco, Aguascalientes, Centro, Chihuahua, Cuernavaca, Guadalajara, Juárez, León, Mexicali, Mexico City, Monterrey, Morelia, Mérida, Puebla, Querétaro, Reynosa, Saltillo, San Luis Potosí, Tampico, Tijuana, Toluca, Torreón, Veracruz	2000
Netherlands	Amsterdam, Eindhoven, Rotterdam, The Hague, Utrecht	2008
New Zealand*	Auckland	2001-2006-2013
South Africa*	Buffalo City, Cape Town, Ekurhuleni, Johannesburg, Emufuleni, eThekwini, Mangaung, Nelson Mandela Bay, Tshwane	2011
United Kingdom	Birmingham, Bradford, Bristol, Cardiff, Leeds, Leicester, Liverpool, London, Manchester, Newcastle, Nottingham, Portsmouth, Sheffield	2001, 2011
United States	Akron, Albany, Albuquerque, Atlanta, Austin, Baltimore, Baton Rouge, Birmingham, Boston, Buffalo, Charleston, Charlotte, Chicago, Cincinnati, Clearwater/St Petersburg, Cleveland, Colorado Springs, Columbia, Columbus, Dallas, Dayton, Denver, Des Moines, Detroit, El Paso, Fort Worth, Fresno, Grand Rapids, Harrisburg, Houston, Indianapolis, Jacksonville, Kansas City, Las Vegas, Little Rock, Los Angeles, Louisville, Madison, McAllen, Memphis, Miami, Milwaukee, Minneapolis, Nashville, New Orleans, New York, Norfolk-Portsmouth-Chesapeake-Virginia Beach, Oklahoma City, Omaha, Orlando, Philadelphia, Phoenix, Pittsburgh, Portland, Providence, Raleigh, Richmond, Sacramento/Roseville, Saint Louis, Salt Lake City, San Antonio, San Diego, San Francisco, Seattle, Tampa, Toledo (only 2000 data), Tucson, Tulsa, Washington, Wichita	2000, 2014

Notes: * Denotes countries that are excluded from the regression analysis because they are missing from the Metropolitan data base

Table 4: Determinants of income segregation (I): city size, government, economy, demography

Dependent variable:	ln(SE 1-km index)							
	[1a]	[1b]	[2a]	[2b]	[3a]	[3b]	[4a]	[4b]
ln(Population)	0.122[a]	0.062[a]	0.137[a]	0.065[a]	0.061[a]	0.052[a]	0.056[b]	0.050[b]
	(0.030)	(0.014)	(0.030)	(0.017)	(0.005)	(0.014)	(0.009)	(0.016)
ln(Fragmentation)			0.093	0.012	0.012[c]	0.016	0.010	0.014
			(0.069)	(0.026)	(0.004)	(0.026)	(0.023)	(0.028)
ln(Labor productivity)					0.144[b]	0.154[b]	0.127[c]	0.203[a]
					(0.019)	(0.055)	(0.062)	(0.054)
ln(Employment rate)					-0.235	-0.248	-0.256	-0.100
					(0.291)	(0.200)	(0.111)	(0.240)
ln(Youth dependency ratio)							0.400[b]	0.391[b]
							(0.103)	(0.160)
ln(Old age dependency ratio)							0.044	0.066
							(0.043)	(0.080)
Country FE	No	Yes	No	Yes	No	Yes	No	Yes
Adjusted R^2	0.063	0.725	0.155	0.725	0.585	0.729	0.672	0.735

Note: 226 observations (107 in 2001, 119 in 2011). All regressions include year fixed effects. Robust standard errors in parenthesis clustered by country. [a], [b] and [c] indicate significant at 1, 5, and 10 percent level, respectively.

Table 5. Determinants of income segregation (II): urban form

Dependent variable:	ln(SE 1-km index)								
	Land area and population density		Monocentric and polycentric cities			Number of centers, Centralization and concentration			
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]
ln(Population)	0.044[a]		0.066[a]			0.036[b]	0.081[a]	0.075[a]	0.086[a]
	(0.010)		(0.014)			(0.016)	(0.017)	(0.018)	(0.013)
ln(Pop) x D Poly			-0.065						
			(0.034)						
ln(Central population)				0.064[a]	0.062[a]				
				(0.009)	(0.10)				
ln(Central pop) x D Poly				-0.020	0.025				
				(0.022)	(0.021)				
ln(Suburban pop)				0.011[a]	0.011[a]				
				(0.002)	(0.002)				
ln(Sub pop) x D Poly				-0.076[b]	-0.067[a]				
				(0.024)	(0.017)				
ln(Land area)	0.011	0.055[c]	0.010	0.008	0.006	0.010	0.021	0.027	0.029
	(0.024)	(0.026)	(0.026)	(0.031)	(0.030)	(0.023)	(0.022)	(0.025)	(0.026)
ln(Population density)		0.044[a]							
		(0.010)							
Theil concentration index						0.118[c]		0.147[b]	0.168[b]
						(0.058)		(0.057)	(0.057)
Average distance to CBD							-0.008[b]	-0.009[b]	-0.009[b]
							(0.003)	(0.003)	(0.003)
Number of centers									-0.032[b]
									(0.014)
ln(Fragmentation)	0.013	0.013	0.016	0.001	0.000	0.002	0.009	0.010	0.013
	(0.027)	(0.027)	(0.028)	(0.021)	(0.021)	(0.013)	(0.012)	(0.011)	(0.012)
ln(Labor productivity)	0.208[b]	0.208[a]	0.230[a]	0.234[b]	0.255[a]	0.195[a]	0.182[b]	0.214[a]	0.226[b]
	(0.053)	(0.053)	(0.061)	(0.064)	(0.057)	(0.051)	(0.062)	(0.044)	(0.044)
ln(LProd) X D Poly					-0.299[b]				
					(0.128)				
ln(Employment rate)	-0.106	-0.106	-0.189	-0.167	-0.144	-0.023	-0.069	-0.075	-0.096
	(0.243)	(0.243)	(0.262)	(0.259)	(0.238)	(0.256)	(0.237)	(0.238)	(0.231)
ln(Youth dependency ratio)	0.377[c]	0.377[c]	0.367[c]	0.386[c]	0.396[c]	0.373[c]	0.332[c]	0.329[c]	0.332[c]
	(0.174)	(0.174)	(0.165)	(0.187)	(0.173)	(0.192)	(0.173)	(0.175)	(0.166)
ln(Old age dependency ratio)	0.076	0.076	0.078	0.141	0.148	0.038	0.045	0.043	0.043
	(0.097)	(0.097)	(0.098)	(0.097)	(0.094)	(0.067)	(0.075)	(0.975)	(0.077)
Dummy Polycentricity			0.885	1.189[a]	3.844[b]				
			(0.494)	(0.233)	(1.268)				
Adjusted R^2	0.736	0.736	0.740	0.750	0.752	0.719	0.723	0.725	0.727

Note: 226 observations (107 in 2001, 119 in 2011). 221 observations for columns 8-9. All regressions include country and year fixed-effects. Robust standard errors in parenthesis clustered by country. [a], [b] and [c] indicate significant at 1, 5, and 10 percent level, respectively.

Table 6. Determinants of income segregation (III): spatial vs. a-spatial segregation indices

Dependent variable:	ln(SE 1-km index)			
	500 m	2 km	4 km	A-Spatial
	[1]	[2]	[3]	[4]
ln(Population)	0.080[a]	0.107[a]	0.164[b]	0.080[b]
	(0.020)	(0.013)	(0.060)	(0.026)
ln(Land area)	0.028	0.028	0.025	0.025
	(0.028)	(0.026)	(0.023)	(0.028)
Theil concentration index	0.166[b]	0.175[b]	0.202[b]	0.159[c]
	(0.061)	(0.056)	(0.065)	(0.074)
Average distance to CBD	-0.008[b]	-0.010[a]	-0.016[a]	-0.008[b]
	(0.003)	(0.003)	(0.005)	(0.003)
Number of centers	-0.027[c]	-0.044[b]	-0.072[b]	-0.023[c]
	(0.013)	(0.016)	(0.026)	(0.011)
ln(Fragmentation)	0.009	0.020	0.035	0.005
	(0.012)	(0.015)	(0.027)	(0.011)
ln(Labor productivity)	0.248[a]	0.197[a]	0.152	0.260[a]
	(0.040)	(0.055)	(0.117)	(0.048)
ln(Employment rate)	-0.118	-0.021	-0.033	-0.106
	(0.206)	(0.256)	(0.279)	(0.167)
ln(Youth dependency ratio)	0.309[c]	0.393[c]	0.365	0.289[c]
	(0.158)	(0.180)	(0.354)	(0.127)
ln(Old age depen. ratio)	0.047	0.039	-0.003	0.058
	(0.075)	(0.082)	(0.098)	(0.073)
Dummy Polycentricity	0.702	0.752	0.765	0.671

Note: 226 observations (107 in 2001, 119 in 2011). All regressions include country and year fixed effects. Robust standard errors in parenthesis clustered by country. [a], [b] and [c] indicate significant at 1, 5, and 10 percent level, respectively.

Table 7. Determinants of income segregation (IV): poor vs. rich

Dependent variable:	ln(SE 1-km index)			
	Poor		Rich	
Percentiles:	10th	20th	80th	90th
	[1]	[2]	[3]	[4]
ln(Population)	0.079	0.063	0.115[b]	0.130[b]
	(0.052)	(0.059)	(0.030)	(0.033)
ln(Land area)	-0.016	0.038	0.011	0.002
	(0.024)	(0.045)	(0.033)	(0.038)
Theil concentration index	0.235[a]	0.173	0.064	0.065
	(0.102)	(0.157)	(0.112)	(0.130)
Average distance to CBD	-0.010	-0.013[a]	-0.006[c]	-0.008[b]
	(0.006)	(0.003)	(0.003)	(0.003)
Number of centers	-0.083	-0.030	-0.025	0.010
	(0.064)	(0.033)	(0.025)	(0.018)
ln(Fragmentation)	0.043[b]	0.010	-0.013	-0.008
	(0.009)	(0.013)	(0.015)	(0.014)
ln(Labor productivity)	0.361[a]	0.403[a]	0.261[a]	0.319[a]
	(0.055)	(0.077)	(0.048)	(0.050)
ln(Employment rate)	-0.011	-0.080	-0.040	0.029
	(0.296)	(0.180)	(0.160)	(0.163)
ln(Youth dependency ratio)	-0.165	0.002	0.479[c]	0.527[b]
	(0.155)	(0.159)	(0.186)	(0.186)
ln(Old age depen. ratio)	0.011	-0.010	0.038	0.157
	(0.130)	(0.143)	(0.067)	(0.085)
Adjusted R^2	0.762	0.787	0.659	0.749

Note: 226 observations (107 in 2001, 119 in 2011). All regressions include country and year fixed effects. Robust standard errors in parenthesis clustered by country. [a], [b] and [c] indicate significant at 1, 5, and 10 percent level, respectively.

FIGURES

Figure 1: Stylized illustration of different spatial configurations of segregation. Each square represents a neighborhood in a city with two groups of residents. Configurations (a), (b), (c), and (e) have the same segregation level when measured aspatially. That is, each neighborhood has only one group. The spatial index, instead, captures the difference in context so that (b) is the most segregated because (b) has the greatest physical distance between blue square and black circles. In contrast, segregation in (e) is very low at the scale of the city. Configurations (d) and (f) illustrate the possible limitations of underlying data. The segregation index is 0 in both cases because this is 'true' in the case of (d) but in (f) the definition of neighborhoods is not fine-grained enough to capture the internal segregation. Adapted from Ejdemyr, Simon, Segregation Measures in R, https://sejdemyr.github.io/r-tutorials/statistics/measuring-segregation.html. Last accessed: March 19, 2020.

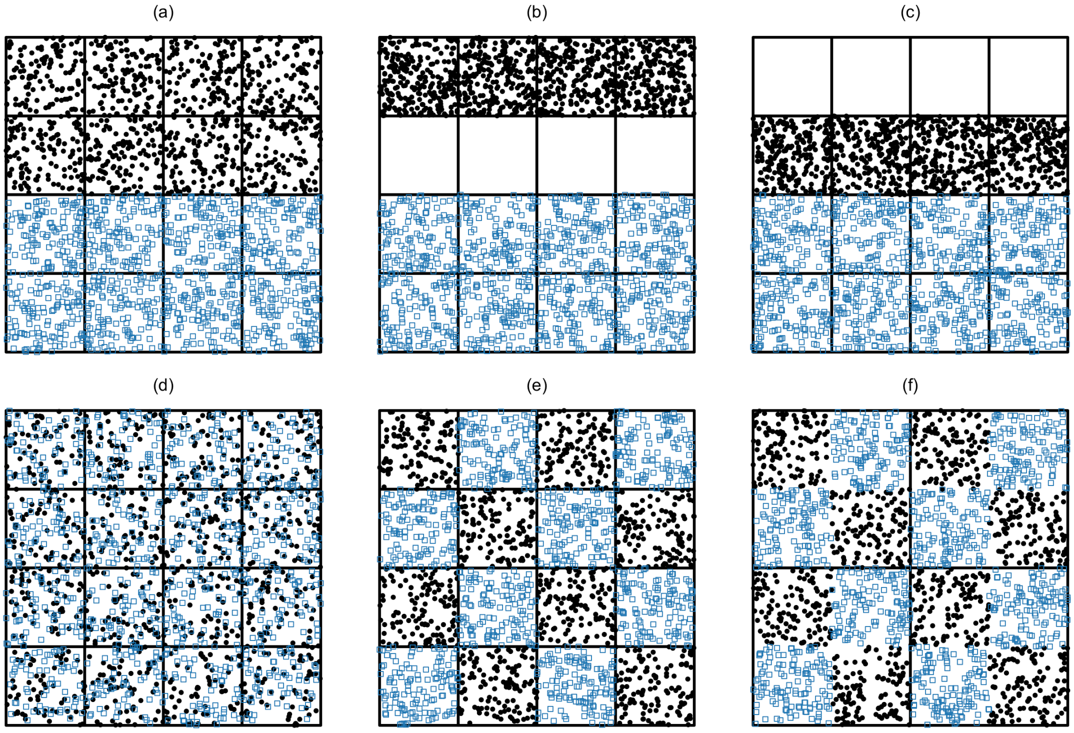
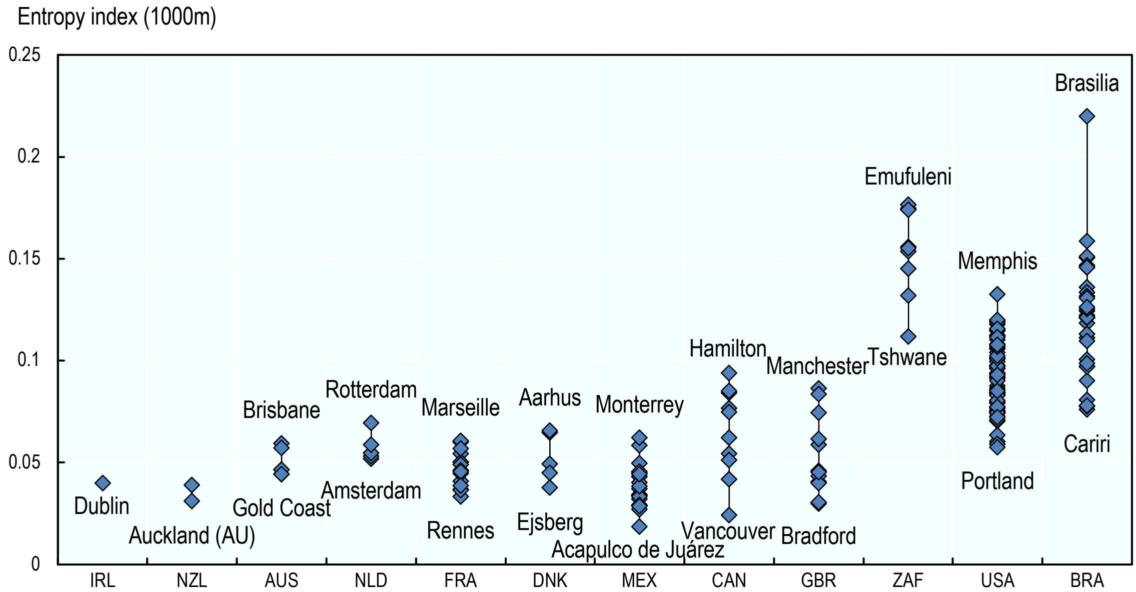


Figure 2: Levels of income segregation in cities, by country (last year available)

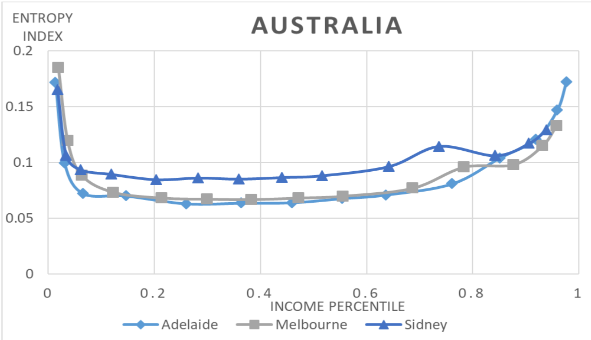
Spatial entropy (1000m scale): higher levels indicate higher segregation



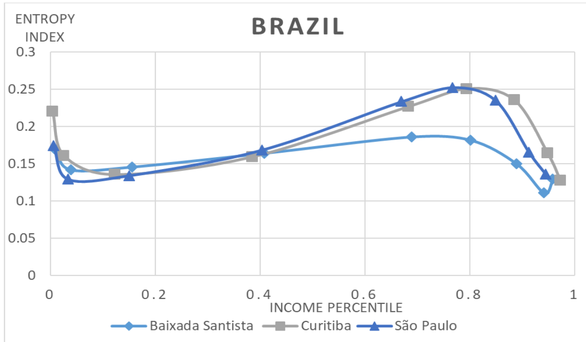
Source: Authors' elaborations based on data detailed in Table 2.

Figure 3: Income segregation in cities by income group

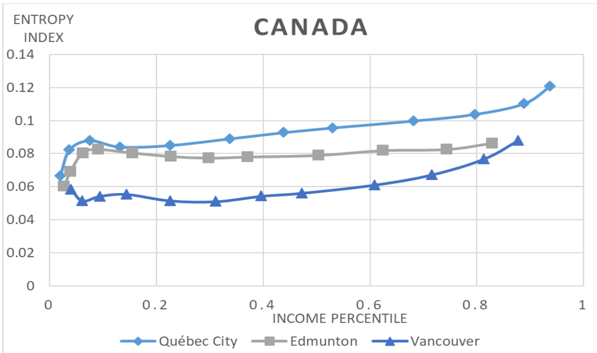
a)



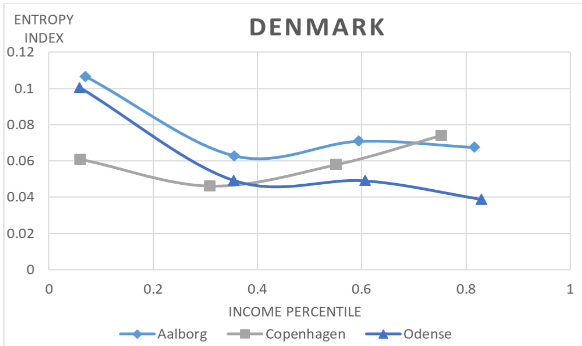
b)



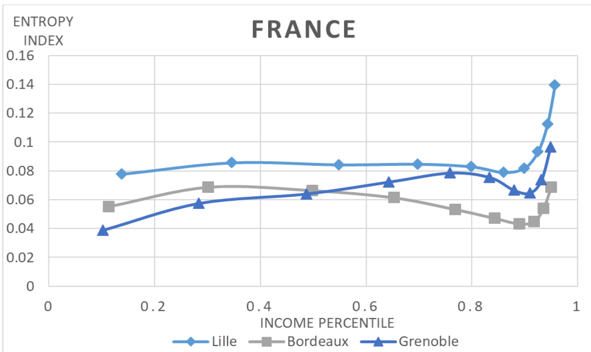
c)



d)



e)



f)

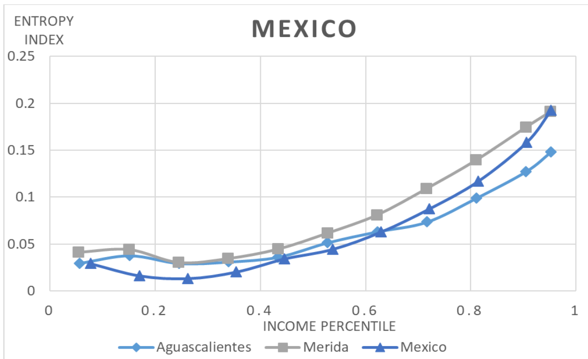
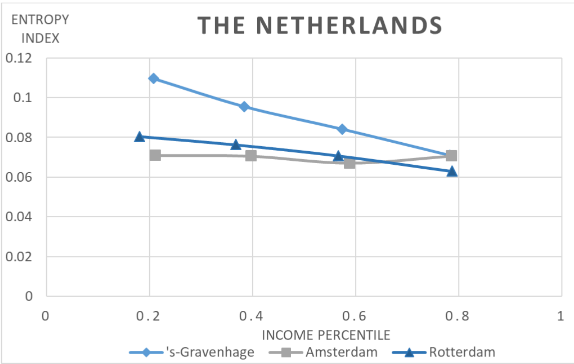
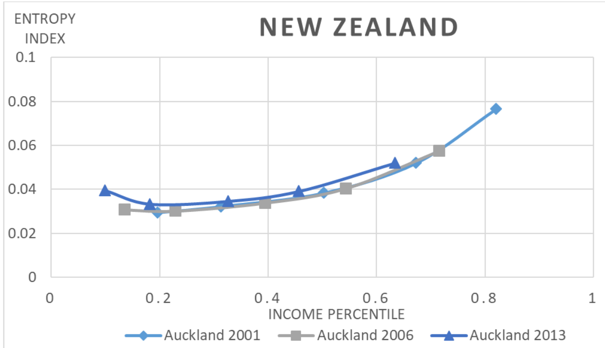


Figure 3 (continued)

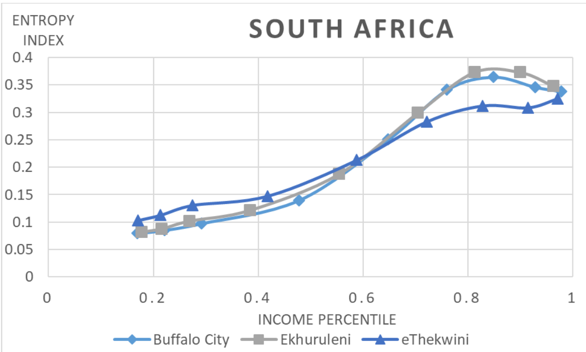
g)



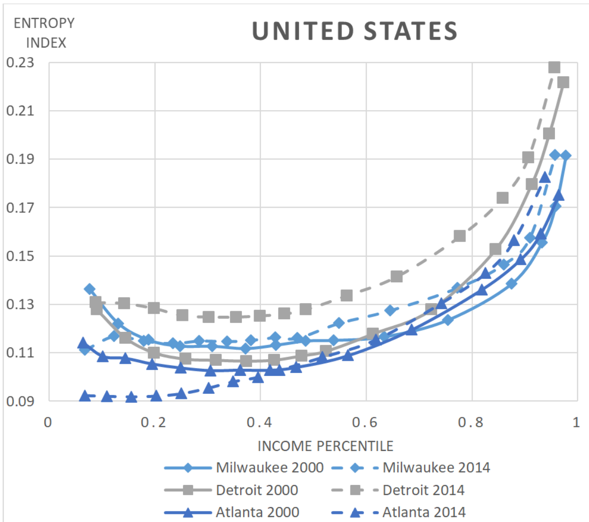
h)



i)



j)

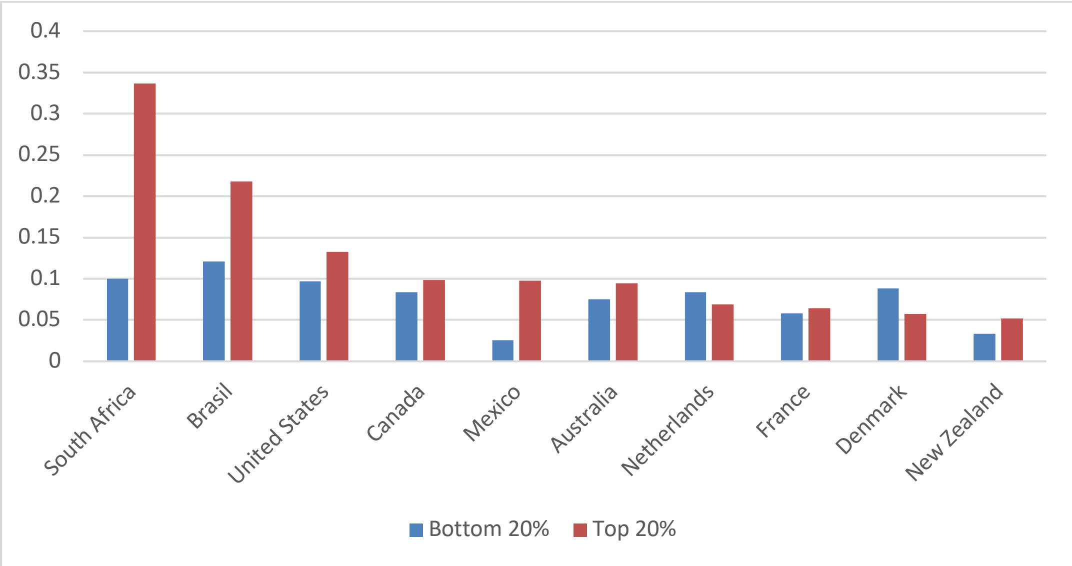


Notes: The curves show levels of segregation across the different income categories. The number of marks on the curves corresponds to the number of income bins available in the data.

Source: Author's elaborations based on data on income distribution at local level (see Table 1).

Figure 4. Income segregation in the bottom and top income groups

In most countries, income segregation is higher among the rich.



Note: Bottom 20% values for Denmark refer to the 6% percentile.

Source: Authors' elaborations based on national data on income distribution (see Table 1).

Appendix I

Technical description of the spatial entropy procedure.

The procedure for calculating the index is derived from a Python script Monkkonen and Zhang (2014) developed and that we modified to handle the larger number of cities and data types in our sample. We also added a step to dasymetrically map population (i.e., we use information about the uninhabited areas from remote sensing data to reallocated census data to populated areas only, see e.g., Dmowska and Stepinski, 2016). This step primarily increases efficiency. It removes the empty parts of sparsely populated tracts at the urban fringe so the next steps can skip them. It also produces more realistic measures of density, which is used to weigh the population in the following steps.

While there is little possibility to resolve this issue completely, the surface-density approach to the measurement of segregation mitigates these effects by estimating the distribution of data within uniformly sized cells. We further minimize the influence of boundaries using dasymmetric mapping. With such a method, data are re-scaled using information about land use (e.g., presence of parks or bodies of water) and related information to provide a more fine-grained gridded population distribution than that available through generic census data. It enables to identify areas that are empty and gain accuracy in some instances where, for example, most of the population of a tract is concentrated near its boundary. Regardless of the method used, segregation measures are only as good as the underlying data. Therefore, in areas where the data is of high quality (i.e., small scale, consistently reliable data), the method matters little. It is only in areas where data are coarser that surface-density yields some advantages, but even then with limitations due to the lack of information about the internal spatial distribution of households within geographical units (Kramer et al., 2010).

The first step after data cleaning is to create a grid of equally sized cells (we use 100 m^2). For each cell, the density of each income group is estimated from the underlying census data based on the proportion of the tract or tracts that fall within the cell. In other words, if 50% of the tract is within the cell, 50% of the population will be assigned to that cell. This means that the gridded data can only be as precise as the underlying data and is the main reason we use a scale that is larger than most units in the underlying data. We smooth the data by averaging the counts with neighboring cells while maintaining overall totals. This procedure transforms the layer of discrete geographical units into a more flexible surface-density layer (i.e., all boundaries have been effectively removed to allow for the definition of any scale units).

Aspatial entropy index calculation weighs the subunits by their share of the total population. With this method, the weighing is based on proximity between subunits. Once the data have been transformed, we apply a biweight kernel proximity function to weigh the observations within the defined local environment, 1 km^2 in our case. In other words this creates a local environment around every cell where the composition of close by cells is weighed more heavily than those at the periphery. As such, spatial entropy is based on a large set of observations approximating a continuous distribution. Mathematically, the index is:

$$\tilde{\Lambda} = \int_{p \in R} \frac{t_p}{T} \cdot \frac{v - \tilde{v}_p}{v} \tag{1}$$

Where R is the region for which the index is calculated, T is the total population and t_p is the population of the neighborhood. v and $\tilde{v}_p$ are the entropy for the city and the neighborhood. The latter is calculated as follows:

$$\tilde{v}_p = -\frac{1}{M-1} \sum_{m=1}^{M-1} \tilde{c}_{pm} \, log_2 \tilde{c}_{pm} + \left(1 - \tilde{c}_{pm}\right) log_2 \left(1 - \tilde{c}_{pm}\right) \tag{2}$$

where M is the number of income groups and $\tilde{c}_{pm} = \sum_{k=1}^{m} \tilde{\pi}_{pk}$ is the cumulative income share in the neighbourhood p for each cell in the surface grid, with $\tilde{\pi}_{pk}$ being the weighed share of the population in income group k such that:

$$\tilde{\pi}_{pk} = \frac{\int_{q \in R} \tau_{qk} \emptyset(p,q) dq}{\int_{q \in R} \tau_{q} \emptyset(p,q) dq}$$

Where τ_{qk} is the population density of income group k at point p, τ_q is the population density at point q and $\emptyset(p,q)$ is the kernel proximity function $\emptyset(p,q) = \begin{cases} \left[1 - \left(\frac{d(p,q)}{r}\right)^2\right] & if\ d(p,q) < r \\ 0\ if\ d(p,q) > r \end{cases}$ where r is the defined radii of the local environment.

The same procedure is applied for each neighborhood to obtain v.

Table A1. Variables used for the empirical analysis referred to functional urban areas

Variable	Description
Ln(SE)	Log of the income segregation indicator (Equation 1) within the FUA.
Ln(Population)	Log of total resident population in functional urban areas (FUAs)
Ln(Central population)	Log of resident population in the core of FUAs
Ln(Suburban pop)	Log of resident population in the commuting zone of FUAs
Ln(Land area)	Log of total surface (km^2) of FUAs
Theil concentration Index	Theil entropy on population density in local units within FUAs. It measures the extent to which people are spatially clustered. Source: Veneri (2018)
Average distance to CD	Weighted average distance of the population from the main center of FUAs. Source: Veneri (2018)
Number of centers	Number of population centers within FUAs. Source: Veneri (2018)
Ln (Fragmentation)	Log of the number of local governments per 100 000 inhabitants within FUAs
Ln(Labor productivity)	Log of GDP per worker in FUAs
Ln(Employment rate)	Log of employment rate in FUAs
Ln(Youth dependency ratio)	Log of population aged 0 to 15 over population aged 16 to 64
Ln(Old age dependency ratio)	Log of population aged over 64 over population aged 16 to 64
Dummy Polycentricity	1 if the FUA has more than one core local unit

Notes: Unless differently specified, all indicators are available in the OECD Metropolitan Database (http://dx.doi.org/10.1787/data-00531-en). Last access: February 2019

2015

2015/1, Foremny, D.; Freier, R.; Moessinger, M-D.; Yeter, M.: "Overlapping political budget cycles in the legislative and the executive"
2015/2, Colombo, L.; Galmarini, U.: "Optimality and distortionary lobbying: regulating tobacco consumption"
2015/3, Pellegrino, G.: "Barriers to innovation: Can firm age help lower them?"
2015/4, Hémet, C.: "Diversity and employment prospects: neighbors matter!"
2015/5, Cubel, M.; Sanchez-Pages, S.: "An axiomatization of difference-form contest success functions"
2015/6, Choi, A.; Jerrim, J.: "The use (and misuse) of Pisa in guiding policy reform: the case of Spain"
2015/7, Durán-Cabré, J.M.; Esteller-Moré, A.; Salvadori, L.: "Empirical evidence on tax cooperation between sub-central administrations"
2015/8, Batalla-Bejerano, J.; Trujillo-Baute, E.: "Analysing the sensitivity of electricity system operational costs to deviations in supply and demand"
2015/9, Salvadori, L.: "Does tax enforcement counteract the negative effects of terrorism? A case study of the Basque Country"
2015/10, Montolio, D.; Planells-Struse, S.: "How time shapes crime: the temporal impacts of football matches on crime"
2015/11, Piolatto, A.: "Online booking and information: competition and welfare consequences of review aggregators"
2015/12, Boffa, F.; Pingali, V.; Sala, F.: "Strategic investment in merchant transmission: the impact of capacity utilization rules"
2015/13, Slemrod, J.: "Tax administration and tax systems"
2015/14, Arqué-Castells, P.; Cartaxo, R.M.; García-Quevedo, J.; Mira Godinho, M.: "How inventor royalty shares affect patenting and income in Portugal and Spain"
2015/15, Montolio, D.; Planells-Struse, S.: "Measuring the negative externalities of a private leisure activity: hooligans and pickpockets around the stadium"
2015/16, Batalla-Bejerano, J.; Costa-Campi, M.T.; Trujillo-Baute, E.: "Unexpected consequences of liberalisation: metering, losses, load profiles and cost settlement in Spain's electricity system"
2015/17, Batalla-Bejerano, J.; Trujillo-Baute, E.: "Impacts of intermittent renewable generation on electricity system costs"
2015/18, Costa-Campi, M.T.; Paniagua, J.; Trujillo-Baute, E.: "Are energy market integrations a green light for FDI?"
2015/19, Jofre-Monseny, J.; Sánchez-Vidal, M.; Viladecans-Marsal, E.: "Big plant closures and agglomeration economies"
2015/20, Garcia-López, M.A.; Hémet, C.; Viladecans-Marsal, E.: "How does transportation shape intrametropolitan growth? An answer from the regional express rail"
2015/21, Esteller-Moré, A.; Galmarini, U.; Rizzo, L.: "Fiscal equalization under political pressures"
2015/22, Escardíbul, J.O.; Afcha, S.: "Determinants of doctorate holders' job satisfaction. An analysis by employment sector and type of satisfaction in Spain"
2015/23, Aidt, T.; Asatryan, Z.; Badalyan, L.; Heinemann, F.: "Vote buying or (political) business (cycles) as usual?"
2015/24, Albæk, K.: "A test of the 'lose it or use it' hypothesis in labour markets around the world"
2015/25, Angelucci, C.; Russo, A.: "Petty corruption and citizen feedback"
2015/26, Moriconi, S.; Picard, P.M.; Zanaj, S.: "Commodity taxation and regulatory competition"
2015/27, Brekke, K.R.; Garcia Pires, A.J.; Schindler, D.; Schjelderup, G.: "Capital taxation and imperfect competition: ACE vs. CBIT"
2015/28, Redonda, A.: "Market structure, the functional form of demand and the sensitivity of the vertical reaction function"
2015/29, Ramos, R.; Sanromá, E.; Simón, H.: "An analysis of wage differentials between full-and part-time workers in Spain"
2015/30, Garcia-López, M.A.; Pasidis, I.; Viladecans-Marsal, E.: "Express delivery to the suburbs the effects of transportation in Europe's heterogeneous cities"
2015/31, Torregrosa, S.: "Bypassing progressive taxation: fraud and base erosion in the Spanish income tax (1970-2001)"
2015/32, Choi, H.; Choi, A.: "When one door closes: the impact of the hagwon curfew on the consumption of private tutoring in the republic of Korea"
2015/33, Escardíbul, J.O.; Helmy, N.: "Decentralisation and school autonomy impact on the quality of education: the case of two MENA countries"
2015/34, González-Val, R.; Marcén, M.: "Divorce and the business cycle: a cross-country analysis"
2015/35, Calero, J.; Choi, A.: "The distribution of skills among the European adult population and unemployment: a comparative approach"
2015/36, Mediavilla, M.; Zancajo, A.: "Is there real freedom of school choice? An analysis from Chile"

2015/37, Daniele, G.: "Strike one to educate one hundred: organized crime, political selection and politicians' ability"
2015/38, González-Val, R.; Marcén, M.: "Regional unemployment, marriage, and divorce"
2015/39, Foremny, D.; Jofre-Monseny, J.; Solé-Ollé, A.: "'Hold that ghost': using notches to identify manipulation of population-based grants"
2015/40, Mancebón, M.J.; Ximénez-de-Embún, D.P.; Mediavilla, M.; Gómez-Sancho, J.M.: "Does educational management model matter? New evidence for Spain by a quasiexperimental approach"
2015/41, Daniele, G.; Geys, B.: "Exposing politicians' ties to criminal organizations: the effects of local government dissolutions on electoral outcomes in Southern Italian municipalities"
2015/42, Ooghe, E.: "Wage policies, employment, and redistributive efficiency"

2016

2016/1, Galletta, S.: "Law enforcement, municipal budgets and spillover effects: evidence from a quasi-experiment in Italy"
2016/2, Flatley, L.; Giulietti, M.; Grossi, L.; Trujillo-Baute, E.; Waterson, M.: "Analysing the potential economic value of energy storage"
2016/3, Calero, J.; Murillo Huertas, I.P.; Raymond Bara, J.L.: "Education, age and skills: an analysis using the PIAAC survey"
2016/4, Costa-Campi, M.T.; Daví-Arderius, D.; Trujillo-Baute, E.: "The economic impact of electricity losses"
2016/5, Falck, O.; Heimisch, A.; Wiederhold, S.: "Returns to ICT skills"
2016/6, Halmenschlager, C.; Mantovani, A.: "On the private and social desirability of mixed bundling in complementary markets with cost savings"
2016/7, Choi, A.; Gil, M.; Mediavilla, M.; Valbuena, J.: "Double toil and trouble: grade retention and academic performance"
2016/8, González-Val, R.: "Historical urban growth in Europe (1300–1800)"
2016/9, Guio, J.; Choi, A.; Escardíbul, J.O.: "Labor markets, academic performance and the risk of school dropout: evidence for Spain"
2016/10, Bianchini, S.; Pellegrino, G.; Tamagni, F.: "Innovation strategies and firm growth"
2016/11, Jofre-Monseny, J.; Silva, J.I.; Vázquez-Grenno, J.: "Local labor market effects of public employment"
2016/12, Sanchez-Vidal, M.: "Small shops for sale! The effects of big-box openings on grocery stores"
2016/13, Costa-Campi, M.T.; García-Quevedo, J.; Martínez-Ros, E.: "What are the determinants of investment in environmental R&D?"
2016/14, **García-López, M.A; Hémet, C.; Viladecans-Marsal, E.:** "Next train to the polycentric city: The effect of railroads on subcenter formation"
2016/15, Matas, A.; Raymond, J.L.; Dominguez, A.: "Changes in fuel economy: An analysis of the Spanish car market"
2016/16, Leme, A.; Escardíbul, J.O.: "The effect of a specialized versus a general upper secondary school curriculum on students' performance and inequality. A difference-in-differences cross country comparison"
2016/17, **Scandurra, R.I.; Calero, J**.: "Modelling adult skills in OECD countries"
2016/18, **Fernández-Gutiérrez, M.; Calero, J**.: "Leisure and education: insights from a time-use analysis"
2016/19, Del Rio, P.; Mir-Artigues, P.; Trujillo-Baute, E.: "Analysing the impact of renewable energy regulation on retail electricity prices"
2016/20, Taltavull de la Paz, P.; Juárez, F.; Monllor, P.: "Fuel Poverty: Evidence from housing perspective"
2016/21, Ferraresi, M.; Galmarini, U.; Rizzo, L.; Zanardi, A.: "Switch towards tax centralization in Italy: A wake up for the local political budget cycle"
2016/22, Ferraresi, M.; Migali, G.; Nordi, F.; Rizzo, L.: "Spatial interaction in local expenditures among Italian municipalities: evidence from Italy 2001-2011"
2016/23, Daví-Arderius, D.; Sanin, M.E.; Trujillo-Baute, E.: "CO2 content of electricity losses"
2016/24, Arqué-Castells, P.; Viladecans-Marsal, E.: "Banking the unbanked: Evidence from the Spanish banking expansion plan"
2016/25 Choi, Á.; Gil, M.; Mediavilla, M.; Valbuena, J.: "The evolution of educational inequalities in Spain: Dynamic evidence from repeated cross-sections"
2016/26, Brutti, Z.: "Cities drifting apart: Heterogeneous outcomes of decentralizing public education"
2016/27, Backus, P.; Cubel, M.; Guid, M.; Sánchez-Pages, S.; Lopez Manas, E.: "Gender, competition and performance: evidence from real tournaments"
2016/28, Costa-Campi, M.T.; Duch-Brown, N.; García-Quevedo, J.: "Innovation strategies of energy firms"
2016/29, Daniele, G.; Dipoppa, G.: "Mafia, elections and violence against politicians"
2016/30, Di Cosmo, V.; Malaguzzi Valeri, L.: "Wind, storage, interconnection and the cost of electricity"

2017

2017/1, González Pampillón, N.; Jofre-Monseny, J.; Viladecans-Marsal, E.: "Can urban renewal policies reverse neighborhood ethnic dynamics?"
2017/2, Gómez San Román, T.: "Integration of DERs on power systems: challenges and opportunities"
2017/3, Bianchini, S.; Pellegrino, G.: "Innovation persistence and employment dynamics"
2017/4, Curto-Grau, M.; Solé-Ollé, A.; Sorribas-Navarro, P.: "Does electoral competition curb party favoritism?"
2017/5, Solé-Ollé, A.; Viladecans-Marsal, E.: "Housing booms and busts and local fiscal policy"
2017/6, Esteller, A.; Piolatto, A.; Rablen, M.D.: "Taxing high-income earners: Tax avoidance and mobility"
2017/7, Combes, P.P.; Duranton, G.; Gobillon, L.: "The production function for housing: Evidence from France"
2017/8, Nepal, R.; Cram, L.; Jamasb, T.; Sen, A.: "Small systems, big targets: power sector reforms and renewable energy development in small electricity systems"
2017/9, Carozzi, F.; Repetto, L.: "Distributive politics inside the city? The political economy of Spain's plan E"
2017/10, Neisser, C.: "The elasticity of taxable income: A meta-regression analysis"
2017/11, Baker, E.; Bosetti, V.; Salo, A.: "Finding common ground when experts disagree: robust portfolio decision analysis"
2017/12, Murillo, I.P; Raymond, J.L; Calero, J.: "Efficiency in the transformation of schooling into competences: A cross-country analysis using PIAAC data"
2017/13, **Ferrer-Esteban, G.; Mediavilla, M.:** "The more educated, the more engaged? An analysis of social capital and education"
2017/14, Sanchis-Guarner, R.: "Decomposing the impact of immigration on house prices"
2017/15, Schwab, T.; Todtenhaupt, M.: "Spillover from the haven: Cross-border externalities of patent box regimes within multinational firms"
2017/16, Chacón, M.; Jensen, J.: "The institutional determinants of Southern secession"
2017/17, Gancia, G.; Ponzetto, G.A.M.; Ventura, J.: "Globalization and political structure"
2017/18, González-Val, R.: "City size distribution and space"
2017/19, García-Quevedo, J.; Mas-Verdú, F.; Pellegrino, G.: "What firms don't know can hurt them: Overcoming a lack of information on technology"
2017/20, Costa-Campi, M.T.; García-Quevedo, J.: "Why do manufacturing industries invest in energy R&D?"
2017/21, Costa-Campi, M.T.; García-Quevedo, J.; Trujillo-Baute, E.: "Electricity regulation and economic growth"

2018

2018/1, Boadway, R.; Pestieau, P.: "The tenuous case for an annual wealth tax"
2018/2, Garcia-López, M.À.: "All roads lead to Rome ... and to sprawl? Evidence from European cities"
2018/3, Daniele, G.; Galletta, S.; Geys, B.: "Abandon ship? Party brands and politicians' responses to a political scandal"
2018/4, Cavalcanti, F.; Daniele, G.; Galletta, S.: "Popularity shocks and political selection"
2018/5, Naval, J.; Silva, J. I.; Vázquez-Grenno, J.: "Employment effects of on-the-job human capital acquisition"
2018/6, Agrawal, D. R.; Foremny, D.: "Relocation of the rich: migration in response to top tax rate changes from spanish reforms"
2018/7, García-Quevedo, J.; Kesidou, E.; Martínez-Ros, E.: "Inter-industry differences in organisational eco-innovation: a panel data study"
2018/8, Aastveit, K. A.; Anundsen, A. K.: "Asymmetric effects of monetary policy in regional housing markets"
2018/9, Curci, F.; Masera, F.: "Flight from urban blight: lead poisoning, crime and suburbanization"
2018/10, Grossi, L.; Nan, F.: "The influence of renewables on electricity price forecasting: a robust approach"
2018/11, Fleckinger, P.; Glachant, M.; Tamokoué Kamga, P.-H.: "Energy performance certificates and investments in building energy efficiency: a theoretical analysis"
2018/12, van den Bergh, J. C.J.M.; Angelsen, A.; Baranzini, A.; Botzen, W.J. W.; Carattini, S.; Drews, S.; Dunlop, T.; Galbraith, E.; Gsottbauer, E.; Howarth, R. B.; Padilla, E.; Roca, J.; Schmidt, R.: "Parallel tracks towards a global treaty on carbon pricing"
2018/13, Ayllón, S.; Nollenberger, N.: "The unequal opportunity for skills acquisition during the Great Recession in Europe"
2018/14, Firmino, J.: "Class composition effects and school welfare: evidence from Portugal using panel data"
2018/15, Durán-Cabré, J. M.; Esteller-Moré, A.; Mas-Montserrat, M.; Salvadori, L.: "La brecha fiscal: estudio y aplicación a los impuestos sobre la riqueza"
2018/16, Montolio, D.; Tur-Prats, A.: "Long-lasting social capital and its impact on economic development: the legacy of the commons"

2018/17, Garcia-López, M. À.; Moreno-Monroy, A. I.: "Income segregation in monocentric and polycentric cities: does urban form really matter?"
2018/18, Di Cosmo, V.; Trujillo-Baute, E.: "From forward to spot prices: producers, retailers and loss averse consumers in electricity markets"
2018/19, Brachowicz Quintanilla, N.; Vall Castelló, J.: "Is changing the minimum legal drinking age an effective policy tool?"
2018/20, Nerea Gómez-Fernández, Mauro Mediavilla: "Do information and communication technologies (ICT) improve educational outcomes? Evidence for Spain in PISA 2015"
2018/21, Montolio, D.; Taberner, P. A.: "Gender differences under test pressure and their impact on academic performance: a quasi-experimental design"
2018/22, Rice, C.; Vall Castelló, J.: "Hit where it hurts – healthcare access and intimate partner violence"
2018/23, Ramos, R.; Sanromá, E.; Simón, H.: "Wage differentials by bargaining regime in Spain (2002-2014). An analysis using matched employer-employee data"

2019

2019/1, Mediavilla, M.; Mancebón, M. J.; Gómez-Sancho, J. M.; Pires Jiménez, L.: "Bilingual education and school choice: a case study of public secondary schools in the Spanish region of Madrid"
2019/2, Brutti, Z.; Montolio, D.: "Preventing criminal minds: early education access and adult offending behavior"
2019/3, Montalvo, J. G.; Piolatto, A.; Raya, J.: "Transaction-tax evasion in the housing market"
2019/4, Durán-Cabré, J.M.; Esteller-Moré, A.; Mas-Montserrat, M.: "Behavioural responses to the re)introduction of wealth taxes. Evidence from Spain"
2019/5, Garcia-López, M.A.; Jofre-Monseny, J.; Martínez Mazza, R.; Segú, M.: "Do short-term rental platforms affect housing markets? Evidence from Airbnb in Barcelona"
2019/6, Domínguez, M.; Montolio, D.: "Bolstering community ties as a means of reducing crime"
2019/7, García-Quevedo, J.; Massa-Camps, X.: "Why firms invest (or not) in energy efficiency? A review of the econometric evidence"
2019/8, Gómez-Fernández, N.; Mediavilla, M.: "What are the factors that influence the use of ICT in the classroom by teachers? Evidence from a census survey in Madrid"
2019/9, Arribas-Bel, D.; Garcia-López, M.A.; Viladecans-Marsal, E.: "The long-run redistributive power of the net wealth tax"
2019/10, Arribas-Bel, D.; Garcia-López, M.A.; Viladecans-Marsal, E.: "Building(s and) cities: delineating urban areas with a machine learning algorithm"
2019/11, Bordignon, M.; Gamalerio, M.; Slerca, E.; Turati, G.: "Stop invasion! The electoral tipping point in anti-immigrant voting"

2020

2020/01, Daniele, G.; Piolatto, A.; Sas, W.: "Does the winner take it all? Redistributive policies and political extremism"
2020/02, Sanz, C.; Solé-Ollé, A.; Sorribas-Navarro, P.: "Betrayed by the elites: how corruption amplifies the political effects of recessions"
2020/03, Farré, L.; Jofre-Monseny; J., Torrecillas, J.: "Commuting time and the gender gap in labor market participation"
2020/04, Romarri, A.: "Does the internet change attitudes towards immigrants? Evidence from Spain"
2020/05, Magontier, P.: "Does media coverage affect governments' preparation for natural disasters?"
2020/06, McDougal, T.L.; Montolio, D.; Brauer, J: "Modeling the U.S. firearms market: the effects of civilian stocks, crime, legislation, and armed conflict"

ieb@ub.edu
www.ieb.ub.edu

Cities