Urban Exceptionalism: Is the social-spatial structure of cities a mark of distinction or commonality?

Andre Comandon
Paavo Monkkonen

ZIMAN DRAFT – August 2021

Urban social-spatial structure – who lives where and at what density - is fundamental to how cities work and how their benefits are shared across social groups. Research shows that it influences innovation and growth, environmental costs, and socioeconomic outcomes. Much of the empirical and theoretical research underpinning these results, however, originated in Anglo-American scholarship at a time when available data was largely limited to a few affluent Anglo countries. This paper introduces a database and accompanying web platform to compare the socioeconomic spatial structure of 140 large cities in six countries – Australia, Brazil, Canada, France, South Africa, and the United States. The goal is to create a set of tools to systematically analyze hypotheses heretofore untested on a global sample of cities. In this first iteration of the database, we focus on putting the cities within the same empirical framework, creating comparable geographic units and measures of socioeconomic structure. We use census data on incomes at the neighborhood level and three geographic variables – area, distance to city center, and bearing – to show the diversity of cities within and between countries and examine the role of inequality and density. We find that higher levels of inequality correlate with more segregation at small and large scales, and that the correlation between density and income is negative or neutral in most cities except for those in Brazil.

Acknowledgements

The authors thank the UCLA Ziman Center for Real Estate's Rosalinde and Arthur Gilbert Program in Real Estate, Finance and Urban Economics for generous support of this research.

1. Introduction

Countless factors shape how cities grow. The settlement's date and timing of growth, colonial occupation, and proximity to natural resources influenced the trajectory and purpose of cities across the globe. The diversity of circumstances surrounding cities gave rise to diverse ways to study urban origins and their implications for development (Bosker & Buringh, 2017; Dorries et al., 2019; Henderson & Venables, 2009; Home, 2013). The multiplication of research emphases has sparked debates about the ability of researchers to generalize our understanding of cities and, therefore, our ability to design urban solutions (Robinson, 2016; Storper & Scott, 2016).

While these debates have been fruitful in opening new areas of studies, international comparative studies tend to lack empirical foundations. That is, the richness of case studies, ethnographies, and historical accounts that make up the bulk of international urban studies outside the Euro-American world is not balanced with representative data that puts cities on the same research plane.

We developed the *Propinquity Project* to both provide a venue for data gathering and to more systematically analyze assumptions about cities in international perspective. The project's aim is to grow the database to be a representative sample of cities and countries that can flexibly be compared using the same metrics. This paper describes the first phases of the project: data collection, comparative methods, and the web platform to explore these data.

The core building blocks of this project are the data collected at a small scale by the national statistical offices, usually as part of their census. Countries have different approaches to census taking and distinct rules for dissemination of data, thus much of our data collection work has focused on standardization. The initial sample of six countries is the result of data availability and type of data. The six countries – Australia, Brazil, Canada, France, South Africa, and the United States –make similar income data available and, therefore, lend themselves to developing robust methods for analysis before extending the sample to countries that have different kinds of data available.[1]

Much of the paper details the method for standardizing the data and fitting them into an empirical framework that enables comparison. We use the methods we developed for these data to create a web application that is available online.[2] The application is connected to the entire database and allows the user to access summary statistics and visualizations about each of the 140 cities in the sample.

The methods for this first phase of the project are purposefully simple. They rely on five variables only: the area of geographic units (neighborhoods), the distance of each neighborhood to the city hall of the metropolitan 'central city', the bearing of the neighborhood relative to city hall, its population, and its income distribution. While the variables are simple, they can be combined in many ways to produce summaries of density patterns and socioeconomic segregation.

The results of our first set of analyses are descriptive. We highlight the differences between cities within and across countries, between countries that have high overall economic inequality – Brazil, South Africa, and the United States – and those where inequality is lower – Australia, Canada, and France, and the relationships between

[1] We currently have data for another nine countries – Chile, Mexico, Belgium, United Kingdom, Ireland, New Zealand, Hong Kong, Denmark, the Netherlands – and Tokyo, Japan.
[2] The website is available at: http://urbanstructure.sites.luskin.ucla.edu/

neighborhood density, proximity to the city center, and incomes. We find that indeed, Australia and Canada have urban structures similar to the United States, with denser centers and sprawling low density suburbs of single-family dwellings. Yet, Australia and Canada lack the strong negative correlation between income and density found in many US cities or the high levels of income segregation at both small and large scale.

The rest of the paper is organized as follows. The literature review highlights some of the main theories underpinning the study of urban spatial structure and provides a brief critique in the context of comparative goals. Then, we introduce the data sources, methodological approaches, and salient hurdles. The fourth section is divided between an overview of the entire sample of cities and detail on the output of the web application. We conclude with a discussion of next steps for the project and comparative urban research.

2. Comparative urban structure

Three questions motivate much urban research: how can cities be made to be more inclusive? How can they provide the highest quality of life possible for their residents? How can they be as efficient as possible (usually in terms of economic production)? These questions are sometimes in tension because the goals of economic efficiency are often seen as inimical to equity. Furthermore, cities have little control over the forces that drive economic development (e.g. where an industry clusters). What cities have more control over is their spatial structure, that is, how they direct the location of economic and residential uses through planning regulations and investment. As such, spatial structure plays a central role in balancing efficiency and equity.

Models based in economics generally begin from the assumption that land use emerges through the decisions of many actors seeking to maximize utility. That is, for firms, the location that will maximize profits and for household the location that maximizes the amount of housing while minimizing distance to work. Such models emphasize transportation and land costs, and the positive and negative externalities of agglomeration as the key factors influencing spatial structure. This framework allows Anas et al (1998) among others to presume that all cities will follow a process similar to U.S. cities with some minor differences because the physical constraints on land are universal.

Scholars base this 'convergence' argument on the observation that cities in Europe and the rest of world have rapidly sprawled and decreased in density. This finding is indeed supported by new measurements of urban population density across a globally representative sample of cities (Angel, 2012). However, in focusing on population density alone, this research misses other, equally significant dimensions. Adding the socioeconomic dimension gives population density profiles of cities added meaning. Not only does an affluent and dense neighborhood differ substantially from a low-income one, it matters to what degree rich and poor can access to employment centers, have access to the same quality public goods and services, and to what degree they can share the same spaces in a city.

On this dimension, the literature has much less to say in a comparative international perspective. The main models of urban structure derive from the monocentric model (Alonso, 1964; Mills, 1967; Muth, 1969). The theory posits that cities grow around a central business district that concentrates economic activity. Rents associated with land diminish with distance because they balance the higher cost of commuting to the economic center. Households allocate their income between housing and commuting expenses. This model's critical insight is that because land prices are higher near the CBD, the density of housing units will be greater to capitalize land values such that land prices and population densities decrease with distance from the center.

The monocentric model has gone through many iterations and refinements over the years, which have demonstrated that its basic insight, the trade-off between accessibility to the city and commuting costs, remains a key component in explaining urban spatial structure (Ahlfeldt, 2011). With each iteration, assumptions were relaxed, and complementary models were added (see Brueckner, 1986 for a review). Two key developments stand out as particularly relevant here.

The first set of assumptions to be relaxed were the homogeneity of income and housing characteristics. Allowing for income to vary, in itself, does not change the basic model, but when adding a housing market with differentiated housing units, income takes on a more important role. The durability of housing means that houses will systematically vary in quality and the monocentric model assumes that a city is built anew in each time period. Some houses will be replaced with higher density construction, others will become prized and command higher prices (Brueckner, 2000; DiPasquale & Wheaton, 1996). In addition, houses are built in different styles, with different amenities and within neighborhoods that are also heterogeneous (Brueckner et al., 1999; Guerrieri et al., 2010). In all locations models, the main conclusion regarding income spatial distribution is that richer households will live further away from the center as long as they value additional housing over additional time spent commuting (Duranton & Puga, 2015).

The second modification derives from the literature on sorting within cities. This literature focuses both on the preferences of people to live near people like them, and the role of municipal governments and the balance between public expenditure and taxes they offer. It shows that, within large and complex metropolitan areas like the ones we study here, the agglomeration of multiple municipalities creates another dimension along which household make locational decisions. One of the simplifying assumptions of the monocentric model, and one of the longstanding models of urban studies, is to think of cities as differentiated concentric rings, in which the direction one travels does not matter, only distance from the center (Park et al., 2019). A more differentiated model integrates the level and cost of services municipalities offer as a deciding factor for where people choose to live (Tiebout, 1956). While this model is not spatial, its implications lead to spatial differentiation. Municipalities use their authority to regulate entry into their housing market through zoning and can, therefore, exclude lower income people (Fennell, 2006). The exclusionary nature of municipal fragmentation exacerbates segregation (Fowler et al., 2016).

Additionally, the assumption of a single employment center has seen a substantial amount of scholarly attention. The elements of polycentrictity are difficult to define. While many cities more closely resemble the monocentric model, enough cases have a polycentric structure to warrant incorporating theoretical implications into the monocentric framework (Arribas-Bel & Sanz-Gracia, 2014; Lucas & Rossi-Hansberg, 2002; Veneri, 2010). Polycentricity builds on the monocentric framework because it does not change the underlying logic based on the tradeoffs between accessibility and land costs. In line with polycentricity as an extension of the monocentric model, income segregation decreases similarly with the concentration of jobs at either a main center or in sub-centers (Garcia-López & Moreno-Monroy, 2018). Perhaps of greater importance, is the relationship between density and income. Evidence point to density increase income segregation up to a certain point and in more centralized cities (Garcia-López & Moreno-Monroy, 2018; Pendall & Carruthers, 2003).

For each dimension of urban socioeconomic spatial structure – monocentricity or centralization, density, segregation, and inequality – the foregoing overview shows that multiple measurements exist, often leading to different results. In the next section, we explain the choice of methods and data available to summarize and compare data in multiple countries.

3. Data and methods

Sample

Our sample of cities reflects the countries for which income data are available publicly in a format and at a neighborhood scale that makes comparison feasible. In all, we collected data from six countries and a total of 140 metropolitan areas. We include all metropolitan areas with population larger than 500,000.

We use the 2016 OECD's definition of Functional Urban Area (FUA) for the four countries with an OECD definition available. FUAs are metropolitan areas consisting of a central urban cluster (the densest urban center in the region) and all the urban clusters that are estimated to have a least 15% of workers commuting to one of the main regional urban clusters (OECD, 2012). For the two countries outside of the OECD, Brazil and South Africa, we use definitions as similar to the FUA as possible. For South

Africa, the metropolitan municipalities defined by the government match well with the FUA methodology. We make one exception for Johannesburg, the largest city, where the neighboring municipalities of Ekurhuleni to the east and Mogale to the west are included because they are highly integrated with Johannesburg. Brazil is trickier because the most consistent metropolitan area definition relies on grouping municipalities, which often include vast rural areas within their boundaries. This creates regions that are unreasonably large. In the case of Brazil, we therefore use the estimated FUA boundaries of Moreno et al. (2020).

Locating the city center

Measures of urban spatial structure generally depend on a defined city center. We settled for the conventional definition - the city hall of the main city. The main city is the one anchoring the FUA and is usually the oldest and largest urban center in the region, as well as the densest.

We make this choice after ruling out multiple alternative definitions. We experimented with six different approaches, four of which yield similar results, and the other two creating unusable data points. These latter two centers are based on the simple geographical median center of the metropolitan area and the median center weighted by light density (as possible proxy for economic activity). The geographic center is problematic because most metropolitan areas today extend far beyond suburban outer rings to include municipalities whose residents commute to the urban core, but leapfrog over very low-density expanses. The light-based center had limited use because light emission levels reach satellite censors' maximum value for most of the metropolitan area, creating only a slightly better representation than the geographic center.

The three methods that yielded similar centers to the city hall coordinates are the population weighed center, the center of the circle that encloses one standard deviation of the population distribution in the metro area (i.e. standard distance), and the center based on the concentration of employment (Manduca, 2018). For most metro areas, the centers based on each of these methods were proximate enough that results did not significantly change. Cities for which those centers differed markedly would constitute exception worth studying, but fall outside the scope of a paper focused on generalizable patterns.

Summarizing spatial structure

Large-N studies tend to rely on summarizing spatial structure in a single variable that can be compared and analyzed. We use this approach in, for example, calculating a segregation index for each city, but the goal for this project is to allow for case-by-case comparison along several dimensions as well. Indexes have the well-known limitation that they collapse information and many researchers advocate using single case visualization as a crucial supplement to understanding spatial structure (Angel, 2012; Johnston et al., 2014). Our goal is to create a set of flexible measures and visualization that can be quickly modified to accommodate different goals and different data sources.

To describe city sub-units or neighborhoods, we use three spatial variables that can be universally applied and combined: the area, distance, and bearing. Area is the area in square kilometers of the geographic building block for each data source, the smallest unit the census makes available with information about income. Table 1 shows the data sources, geographic unit, and income information. Distance is the straight-line distance (in km) between the city center and the centroid of each neighborhood. Finally, the bearing is the angle at which the straight-line between centroids is located. In other words, if a center were drawn with the city center in the middle, the bearing gives the angle of each centroid on that circle in degrees where 0/360 is north, 90 degree is east, 180 south, and 270 west.

Using these three variables, we create standard geographic units (see Figure 1) by dividing every city into sectors that are either 5 degrees wide and 2km long (small) or 10 degrees wide and 5km long (large). This provides a way to visualize all cities diagrammatically using the same units. It mirrors the classical representation of the city Park and Burgess (2019) introduced in the early 20th century and is therefore a way to examine some of the assumptions embedded in this representation. In addition, we use area, population, and distance to measure how the city changes at different distances from the city center, which can in turn be combined with bearings to examine how these changes depends on the orientation.

Income

The income data come in various forms, as illustrated in Table 1. After reformatting the data to have the same structure – numbers of households within income ranges – we standardize the incomes to the regional median. This removes some of the differences in reporting. Not all differences in census reporting can be overcome. For example, Brazil only reports household income per capita, that is the total household income divided by the number of people in the household. While this is a useful measure, it biases the data downward compared to other countries.

Nonetheless, we can estimate the median income and income quintiles for all geographic units and normalize them with the regional median to obtain an index of income that is readily comparable across countries.

The estimation of income measures from binned data poses challenges (Jargowsky & Wheeler, 2018; von Hippel et al., 2016) because of the lack of information about the tail of the distribution and irregular distributions at small scales. That is, the income distribution for a large region can be expected to follow a distribution that can be modelled (e.g. a Pareto distribution), but small neighborhoods are more likely to have irregular distribution (e.g Westwood, where UC Los Angeles is located, has very high income individuals and students with very low incomes).

Fortunately, these obstacles are less salient when estimating quantiles. Quantiles are invariant to what is happening in the tail of the distribution and a simple linear interpolation will get very close to the true value in most cases where the population is large enough. In all cities, the estimated median was within about 2% of the official census figures where these are available.

4. Comparing Urban Structure

We develop the data infrastructure to be flexible. That is, data can easily be reconfigured to output results at different scales or using different relationships between units. The results we present here are from the first version of a platform we envision will grow to give users greater control to produce different visualizations and comparisons. As the database grows, more systematic analysis of hypotheses will be possible. Meanwhile, we focus on two sets of outputs, a set of global comparisons using the entire sample of cities and an overview of Los Angeles to showcase the outputs available for all cities individually.

International comparison

Much research on urban spatial structure focuses on the relationship between land prices and proximity to the center. Land that is closer to a center of employment is more valuable. Developers balance higher land costs with more intense development, substituting land for capital, leading to higher densities of population and activity near the center of cities. The relationship between density and proximity to city centers holds true globally (Bertaud & Malpezzi, 2003), the relationship to income varies. Figure 2 shows the correlation between density and income for all cities. A higher correlation indicates that higher (if the correlation is positive) or lower (if negative) incomes correlate with higher density.

Conventionally, cities of Europe are often portrayed as having wealthier dense centers and poorer suburbs. However, we see that the poorer suburbs tend to also have high densities, muddling the relationship between density and income. Furthermore, more recent urbanization patterns in European cities have tended towards the production of lower density middle class suburbs (Hirt, 2014). This is clear in France where cities are divided between those that have a strongly negative correlation between density and income and those with no correlation. The lower correlation cities tend to be the largest (Paris and Lyon), in which wealthy and poor neighborhoods are high density, and the smaller cities where density and incomes vary less (e.g. Nantes).

Nonetheless, the pattern in France contrasts with the United States. The quintessential wealthy, low density suburb/lower income dense center country has a more even distribution, but there, too, cities range from having no relationship between income and density to having a strong negative relationship. The relationship to city size, however, runs the other way. The largest cities have the strongest negative correlation. New York City, despite its famous dense and wealthy neighborhoods has the strongest negative correlation, followed closely by Los Angeles. Highly suburban and lower income regions have the lowest scores (e.g. Raleigh and El Paso). Interestingly, Canada and South Africa have a more consistent negative correlation across their cities.

In the sample, only Australian and Brazilian cities have a positive correlation between income and density. In Australia, the correlation is mostly weak, but Brazil is a mirror image of the United States. In most Brazilian cities, affluent neighborhoods are high density and low-density places are low income on average. The relationship in Brazil holds despite having a similar pattern as French cities where both the wealthiest and poorest neighborhoods tend to be high density. Brazilian cities also vary in this relationship. Rio de Janeiro, the second largest region, has the highest correlation but Sao Paulo, the largest city, has a slightly negative score.

The relationship between neighborhood density and income reveals the type of neighborhoods that people live in but says little about the spatial relationship between them. Figure 3 shows the ratio of income between the first quartile of population by distance to the third quartile. In other words, it compares the median income of the city center to that of the ‘middle suburbs’ for lack of a better description. The normalization of distance by population results in different physical distances for each city. For example, in Lyon it compares people living with 3.5km of the center to those living between 7.5 and 14km away. In Cleveland, which has the same population size, the corresponding distances are 11km and between 17 and 26km.

Brazil and South Africa have the greatest difference between center and suburb. In most cities in these two countries, the city centers have median income 2 to 4 times higher than the suburbs. The difference is consistent in South Africa, but some cities in Brazil, mostly small- to medium-sized, have a lower ration and even the opposite relationship. The largest Brazilian cities have ratios over 2. Australia is the only other country where cities tend to have higher median income in the center, though not by much. Income differences then to also be small in France, but most cities there have higher incomes in the suburbs.

Most cities in Canada and the United States have higher income in the suburbs. In the United States, the magnitude of the difference is comparable to that found in the South Africa. Several cities have income in the central city less than half that of the suburbs (e.g. Indianapolis and Fresno).

In addition to variation over distance, cities segregate at a much finer scale. Figures 4 and 5 show the results for large- and small-scale income segregation. Figure 4 is the ratio of the sectors with the highest and lowest median income. The sectors are broadly defined to include about 3% of the city's total population, roughly the size of a large neighborhood or small city. Therefore, the ratio expresses the degree of large-scale inequality in a city. A high ratio means that large swaths of the city are homogenously high income and low income. Lower ratio indicates that the city has no area that is homogenously high or low income.

Australia, Canada, and France fall in the low ratio category. Few cities have a ratio higher than two, meaning that the sector with the highest median income has income twice as high as that of the sector with the lowest median income. Most Brazilian and South African cities have a ratio higher than two and often as high as four or higher. The large difference between areas of the city reflects the high level of inequality in these countries and the tendency toward developing wholly separate enclaves at the extremes of the income distribution (e.g. gated estate on the scale of small towns). Cities in the United States span the entire range from ratios close to one all the way to nearly four. The cities with higher ratios tend to be declining industrial centers (e.g. Detroit and Dayton) with large gaps between the struggling center and suburbs that benefited from the exodus of people and jobs from those centers.

The traditional segregation index captures the smaller scales separations that happen between neighborhoods (Figure 5). In other words, a city may not have large-scale inequality yet still have smaller scale separation of different income groups. France, for example, does not have large scale inequality but has a much higher segregation index than Australia and Canada. Australia tends to have low levels of spatial separation between higher and lower income residents, while France has high levels of local separation that, when aggregated to larger scale, give the impression of equality between parts of the city. Brazil, South Africa, and the United States in contrast, have consistently high levels of segregation both large and small scale.

Individual cases

The primary goal of the platform is to allow the user to go from high-level comparisons to exploring specific cases. We developed an app that allows the user to browse the sample of cities and compare any pair side-by-side. In this section, we provide an overview of the output of this app to ease interpretation. We use Los Angeles as an example.

For each city, the app outputs a set of summary statistics that form the basis for the international comparison (i.e. population, density, inequality, and segregation). In addition, it produces three figures that expand on the relationships discussed in the international comparison. Rather than a simple correlation between density and income, the first figure (Figure 6) shows how income changes from the least dense areas to the densest. Each geographical unit is ranked by population density to produce percentiles. All units that belong to the same percentile are then aggregated and the median income calculated. As before, there is no spatial relationship in these summaries, it only relates households' urban environments to incomes.

We see in Figure 6 that the median income in Los Angeles is highest in the lower density neighborhoods but varies little in the first quartile. Income begins to decrease past the first quartile and becomes lower than the regional median at around the 70th percentile. In other words, neighborhoods that are less dense than 70% of all neighborhoods are at or above the regional median income while all dense neighborhoods are below. The densest neighborhoods in this case have median income about 75% of the regional median.

The next two figures show how income changes over space and direction in Los Angeles, providing detail to the phenomena demonstrated in Figures 3 and 4, how cities have both large- and small-scale patterns of segregation. Figure 7 shows how income changes over distance from the center of Los Angeles. The overall pattern shows that the center is lower income and income increases gradually before plateauing at around 40% of the cumulative population. The light gray lines, however, show that there is a increasing variation as distance increases. Unlike the main, heavier lines, which are smoothed, the light lines plot the income level for each kilometer band. Therefore, in the last quartile of distance, the median and top quantile of income fluctuate so much that they overlap.

To reduce this variation, we create directional plots. Splitting the city into 90-degree slices centered on the main cardinal directions, we can see how income changes in different parts of the city. The plots in Figure 8 show that the concentration of wealth and poverty is not as simple as the center-suburb dichotomy. The northern and eastern quadrants have increasing income over distance before decreasing again, and variation is middling throughout. In contrast, the southern and western quadrants show greater fluctuations and a much clearer trend of lower income near the center and higher incomes with distance. Notably, the upper quintile line (top yellow line) shows a clear divergence as incomes at the top of the distributions increase fastest.

Figure 9 provides a different way of visualizing these same trends. It uses the sectors defined by distance and bearing to show the concentration of high incomes throughout the city. In the case of Los Angeles, the concentration of higher income in the west and southeast become very clear, as does the absence of upper income residents in much of central Los Angeles and northwest of the region.

5. Conclusion

This paper details the first steps in creating a large international database dedicated to studying urban socioeconomic spatial structure. The results show that through some reformatting, data that are reported in a distinct manner by different national census bureaus can nonetheless be integrated into a single framework. The first version of the database highlights differences in inequality across countries but also important differences across cities within the same country. The contrasting correlation between income and density in Sao Paulo (slightly negative) and Rio de Janeiro (strongly positive), for example, demonstrates the need for a combination of large-scale analysis and deeper case study.

The focus on individual cases through the web application also shows the need for more flexible data infrastructures. Modern applications, like the one we use, have the advantage of putting much of that flexibility in the hands of the user. As we move forward, we plan to increase the number of controls that can be manipulated with ease (e.g. mapping concentration of high/low income or median income).

The main focus of this work, however, will be on increasing the coverage of the database. While the data currently emphasizes incomes, we have many more variables available like education, ethnicity and nationality, and social class. These variables are useful in developing a fuller understanding of inequality, but also in expanding the sample size. Most countries do not record information on income at a small scale in their census, but provide other information that strongly relates to this dimension of inequality. While education is one obvious proxy for income, variables like the construction materials of a house and access to water and electricity can enable us to enrich the database, especially in countries with large informal sectors (e.g. Mexico and Chile). An important next step will be to develop measures that do not depend on income data but capture similar types of inequality. Because we already have countries with both income data and data on many other dimensions (South Africa and Brazil), we can develop such measures and verify that they function as intended.

One of the primary goals of this project is to examine the assumptions and hypotheses about cities that have not been tested thoroughly with international data. The increasing sophistication of data collection efforts that bypass censuses (e.g. remote sensing of urban footprint), has greatly expanded our ability to do so. This project adds an important empirical dimension to the conversation and brings into focus the kind of data that are most useful in shaping urban research and policy.

6. References

Ahlfeldt, G. (2011). IF ALONSO WAS RIGHT: MODELING ACCESSIBILITY AND EXPLAINING THE RESIDENTIAL LAND GRADIENT. *Journal of Regional Science*, *51*(2), 318–338. https://doi.org/10.1111/j.1467-9787.2010.00694.x

Alonso, W. (1964). *Location and Land Use Toward a General Theory of Land Rent*. Harvard University Press.

Anas, A., Arnott, R., & Small, K. A. (1998). Urban spatial structure. *Journal of Economic Literature*, *36*(3), 1426–1464.

Angel, S. (2012). *Planet of cities*. Lincoln Institute of Land Policy.

Arribas-Bel, D., & Sanz-Gracia, F. (2014). The validity of the monocentric city model in a polycentric age: US metropolitan areas in 1990, 2000 and 2010. *Urban Geography*, *35*(7), 980–997. https://doi.org/10.1080/02723638.2014.940693

Bertaud, A., & Malpezzi, S. (2003). *The spatial distribution of population in 48 world cities: Implications for economies in transition* [Unpublished Manuscript].

Bosker, M., & Buringh, E. (2017). City seeds: Geography and the origins of the European city system. *Journal of Urban Economics*, *98*, 139–157. https://doi.org/10.1016/j.jue.2015.09.003

Brueckner, J. K. (1986). The structure of urban equilibria: A unified treatment of the muth-mills model. In P. Nijkamp, E. S. Mills, & P. C. Cheshire (Eds.), *Handbook of regional and urban economics*. Elsevier Science Pub. Co.

Brueckner, J. K. (2000). Urban growth models with durable housing: An overview. In J.-M. Huriot & J. F. Thisse (Eds.), *Economics of cities: Theoretical perspectives*. Cambridge University Press.

Brueckner, J. K., Thisse, J.-F., & Zenou, Y. (1999). Why is central Paris rich and downtown Detroit poor?: An amenity-based theory. *European Economic Review*, *43*(1), 91–107.

DiPasquale, D., & Wheaton, W. C. (1996). *Urban economics and real estate markets*. Prentice Hall.

Dorries, H., Hugill, D., & Tomiak, J. (2019). Racial capitalism and the production of settler colonial cities. *Geoforum*, S001671851930226X. https://doi.org/10.1016/j.geoforum.2019.07.016

Duranton, G., & Puga, D. (2015). Urban land use. In G. Duranton, J. V. Henderson, & Strange (Eds.), *Handbook of Regional and Urban Economics* (Vol. 5). Elsevier Science Pub. Co.

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* (pp. 163–198). Lincoln Institute of Land Policy.

Fowler, C. S., Lee, B. A., & Matthews, S. A. (2016). The Contributions of Places to Metropolitan Ethnoracial Diversity and Segregation: Decomposing Change Across Space and Time. *Demography*, *53*(6), 1955–1977. https://doi.org/10.1007/s13524-016-0517-3

Garcia-López, M.-À., & Moreno-Monroy, A. I. (2018). Income segregation in monocentric and polycentric cities: Does urban form really matter? *Regional Science and Urban Economics*, *71*, 62–79. https://doi.org/10.1016/j.regsciurbeco.2018.05.003

Guerrieri, V., Hartley, D., & Hurst, E. (2010). *Endogenous gentrification and housing price dynamics* (Working Paper No. W16237). National Bureau of Economic Research. http://www.nber.org/papers/w16237.pdf

Henderson, J. V., & Venables, A. J. (2009). The dynamics of city formation. *Review of Economic Dynamics*, *12*(2), 233–254. https://doi.org/10.1016/j.red.2008.06.003

Hirt, S. (2014). *Zoned in the USA: The origins and implications of American land-use regulation*. Cornell University Press.

Home, R. K. (2013). *Of planting and planning: The making of british colonial cities* (Second Edition). Routledge, Taylor & Francis Group.

Jargowsky, P. A., & Wheeler, C. A. (2018). Estimating Income Statistics from Grouped Data: Mean-constrained Integration over Brackets. *Sociological Methodology*, *48*(1), 337–374. https://doi.org/10.1177/0081175018782579

Johnston, R., Poulsen, M., & Forrest, J. (2014). Segregation matters, measurement matters. In C. Lloyd, I. Shuttleworth, & D. Wong (Eds.), *Social-spatial segregation*. Policy Press.

Lucas, R. E., & Rossi-Hansberg, E. (2002). On the Internal Structure of Cities. *Econometrica*, *70*(4), 1445–1476. https://doi.org/10.1111/1468-0262.00338

Manduca, R. (2018). *The Spatial Structure of US Metropolitan Employment: New Insights from LODES Data* [Preprint]. SocArXiv. https://doi.org/10.31235/osf.io/6sxv5

Mills, E. S. (1967). An Aggregative Model of Resource Allocation in a Metropolitan Area. *The American Economic Review*, *57*(2), 197–210. JSTOR.

Moreno-Monroy, A. I., Schiavina, M., & Veneri, P. (2020). Metropolitan areas in the world. Delineation and population trends. *Journal of Urban Economics*. https://doi.org/10.1016/j.jue.2020.103242

Muth, R. F. (1969). *Cities and housing: The spatial pattern of urban residential land use*. University of Chicago Press.

OECD. (2012). *Redefining “Urban”: A New Way to Measure Metropolitan Areas*. OECD. https://doi.org/10.1787/9789264174108-en

Park, R. E., Burgess, E. W., & Sampson, R. J. (2019). *The city*. The University of Chicago Press.

Pendall, R., & Carruthers, J. I. (2003). Does density exacerbate income segregation? Evidence from U.S. metropolitan areas, 1980 to 2000. *Housing Policy Debate*, *14*(4), 541–589. https://doi.org/10.1080/10511482.2003.9521487

Robinson, J. (2016). Comparative Urbanism: New Geographies and Cultures of Theorizing the Urban. *International Journal of Urban and Regional Research*, *40*(1), 187–199. https://doi.org/10.1111/1468-2427.12273

Storper, M., & Scott, A. J. (2016). Current debates in urban theory: A critical assessment. *Urban Studies*, *53*(6), 1114–1136. https://doi.org/10.1177/0042098016634002

Tiebout, C. (1956). A pure theory of local expenditures. *Public Goods and Market Failures, New Brunswick: Transaction Publishers*, 179–192.

Veneri, P. (2010). Urban Polycentricity and the Costs of Commuting: Evidence from Italian Metropolitan Areas: POLYCENTRICITY AND COMMUTING COSTS. *Growth and Change*, *41*(3), 403–429. https://doi.org/10.1111/j.1468-2257.2010.00531.x von Hippel, P. T., Scarpino, S. V., & Holas, I. (2016). Robust Estimation of Inequality from Binned Incomes. *Sociological Methodology*, *46*(1), 212–251. https://doi.org/10.1177/0081175015599807

7. Tables

Table 1. Overview of data sources.

	# of cities	Year	Avg. base unit population	# of income categories	Income definition
Australia	6	2015	430	16	Total weekly household income
Brazil	40	2010	686	9	Fraction of national monthly minimum wage income per household capita
Canada	7	2016	4791	19	Total annual household income
France	15	2013	2913	13	Total annual household income
South Africa	5	2011	700	11	Total annual household income
United States	68	2017	1590	16	Total annual household income

Table 2. National summary statistics

	Avg city population (in million)	Avg density (person per km^2)	Avg population-weighted density	Avg segregation	80/20 inequality ratio	% population in center
Australia	2.6	331	3259	5.2	4.0	12.8
Brazil	2.0	2033	11450	15.2	5.4	37.8
Canada	2.6	384	3338	7.4	3.2	17.8
France	1.6	802	6269	9.5	2.9	42.0
South Africa	2.6	1276	10671	16.4	25.6	6.1
United States	3.9	373	1998	11.7	4.7	11.0

8. Figures

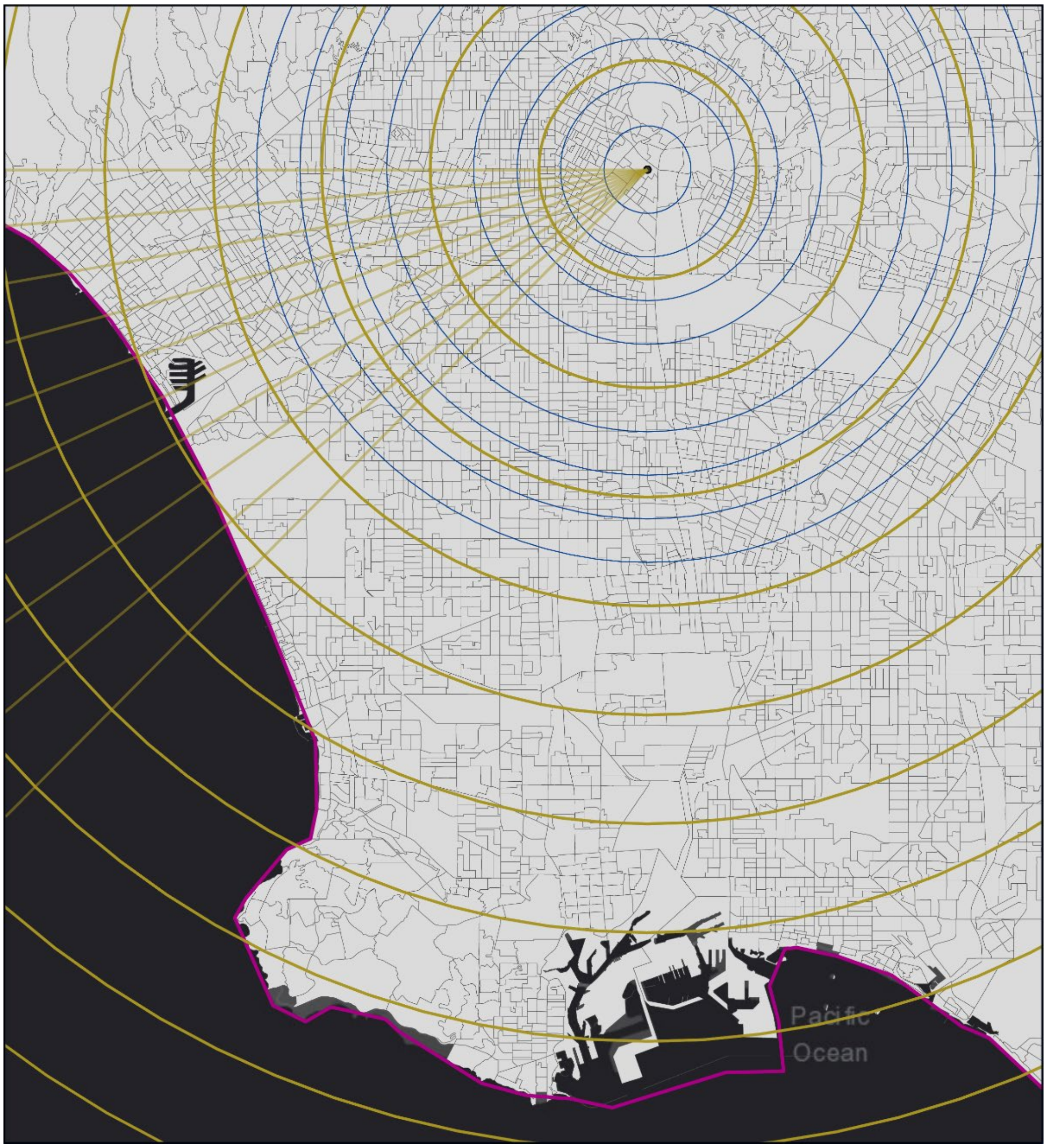

Figure 1. Illustration of the method to create sub-urban sectors

The yellow concentric rings are 5 kilometers apart (larger sectors) and the blue ones 2km apart (smaller sectors). The lines going out from the center are 5 degrees apart.

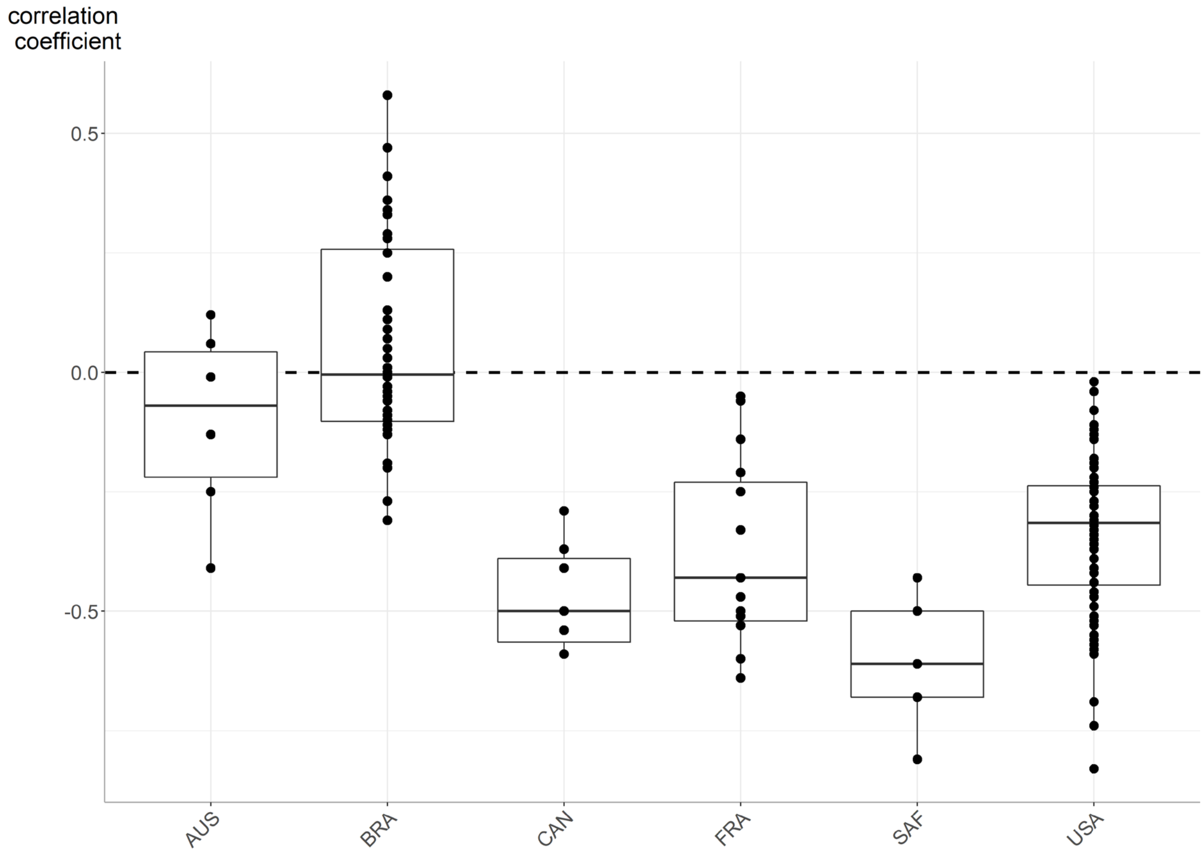

Figure 2. Correlation between neighborhood density and median income

Neighborhoods were ranked to create a percentile of density before calculating the median income for each density percentile. Each dot is the correlation score for one city between the population density at every percentile and median income. The top and bottom lines of each box are the upper and lower quartile of the distribution, and the middle line is the median. Negative correlation indicate that as density increases, income decrease.

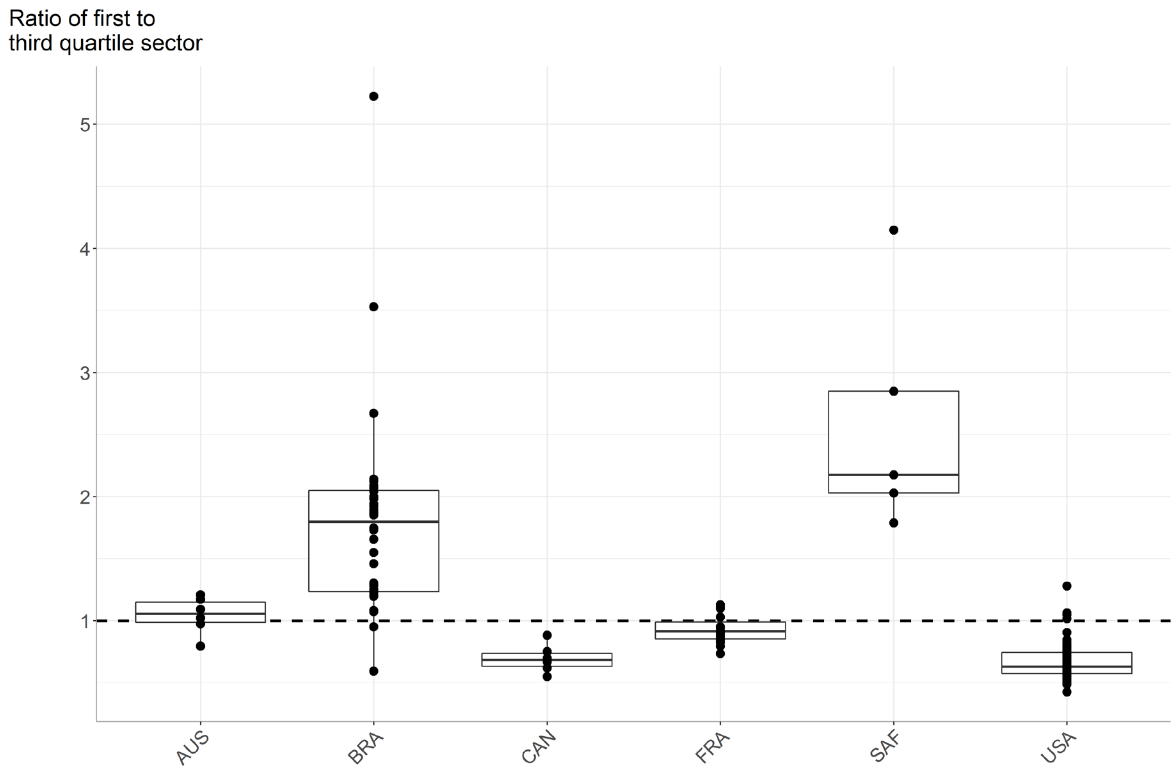

Figure 3. Ratio of median income of first distance quartile to third distance quartile (as measured by cumulative population)

All neighborhoods were ranked by distance before calculating the cumulative population as the distance from the center increased. As such the physical distance is different for all cities, but in every case, we are comparing the median income of people who live where 25% of the population lives closest to the center to where 50 and 75% of the population lives.

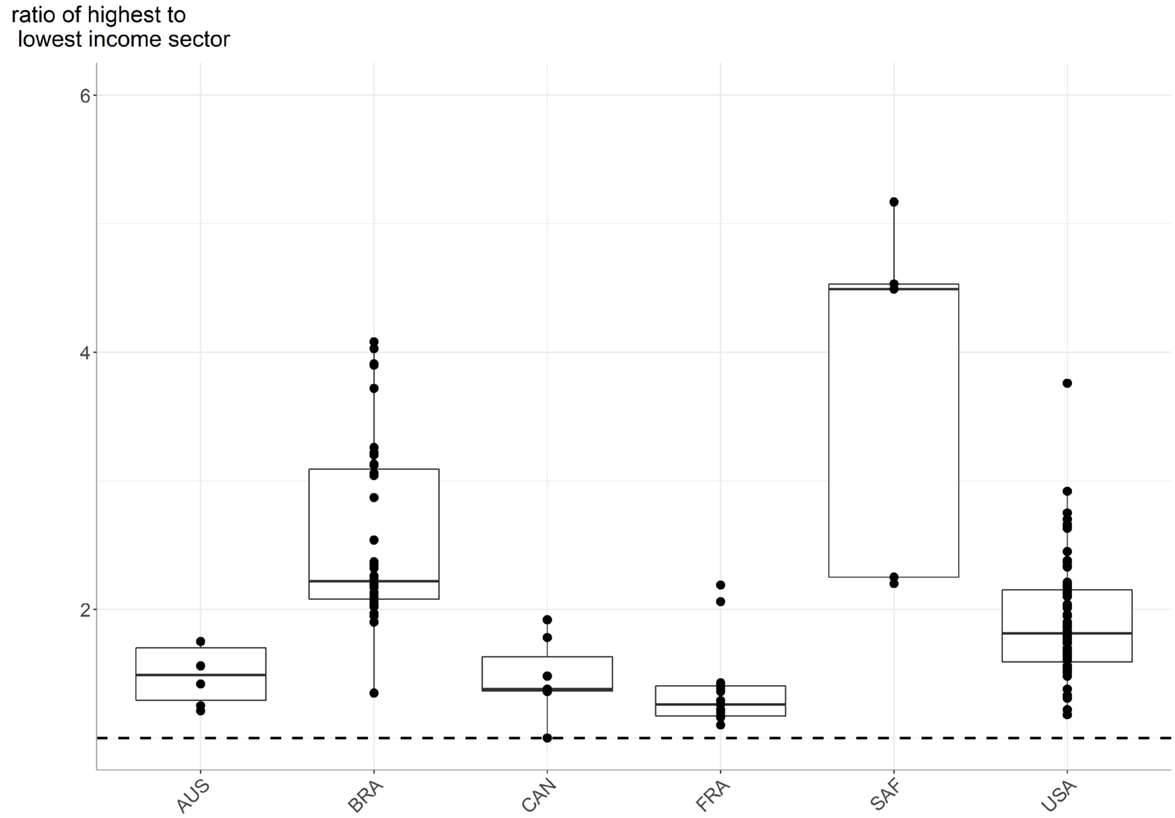

Figure 4. Neighborhood inequality

Each city was divided into 32 sectors, each representing about 3% of the population. The median income of each sector was estimated and the ratio of the highest to lowest value recorded.

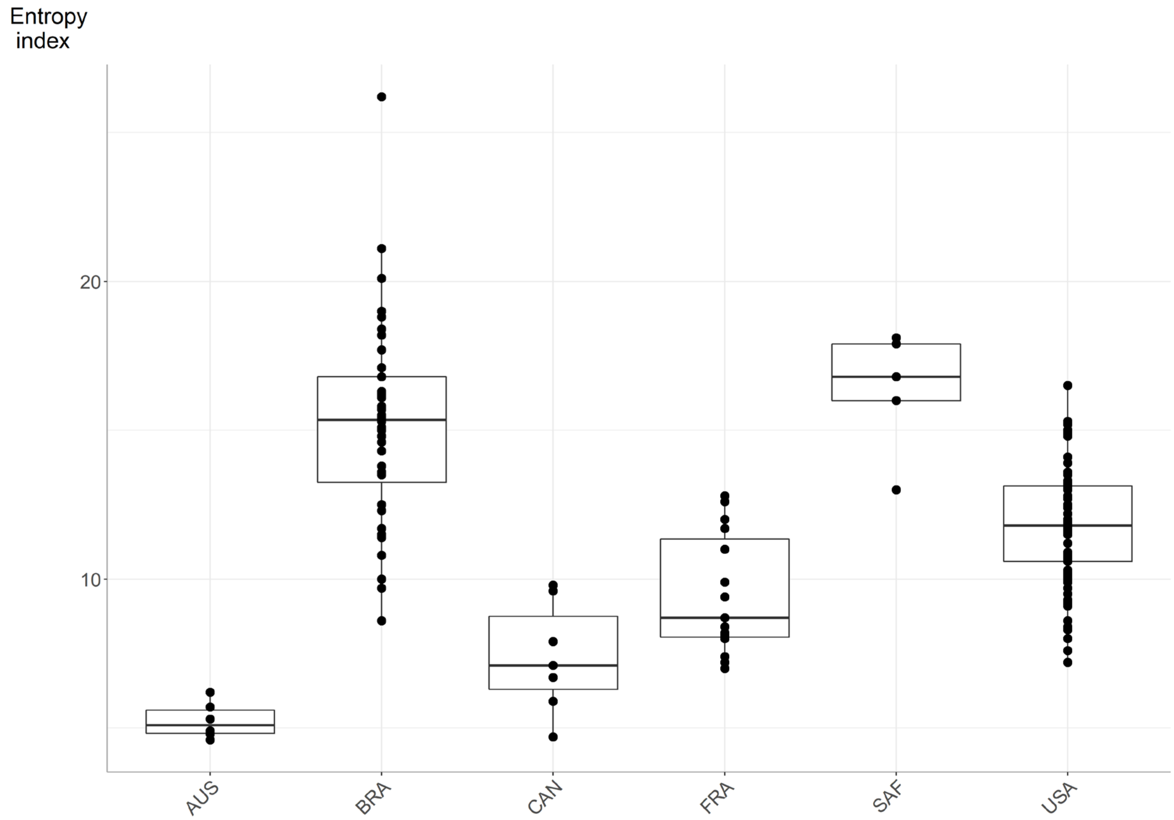

Figure 5. Segregation measured using the entropy index

The entropy index measures segregation by comparing how evenly the distribution of income in a neighborhood matches the income distribution of the city. A lower score means that neighborhoods tend to mirror the distribution of the city (low segregation) and a higher score means that neighborhoods tend to overrepresent specific income groups (high segregation).

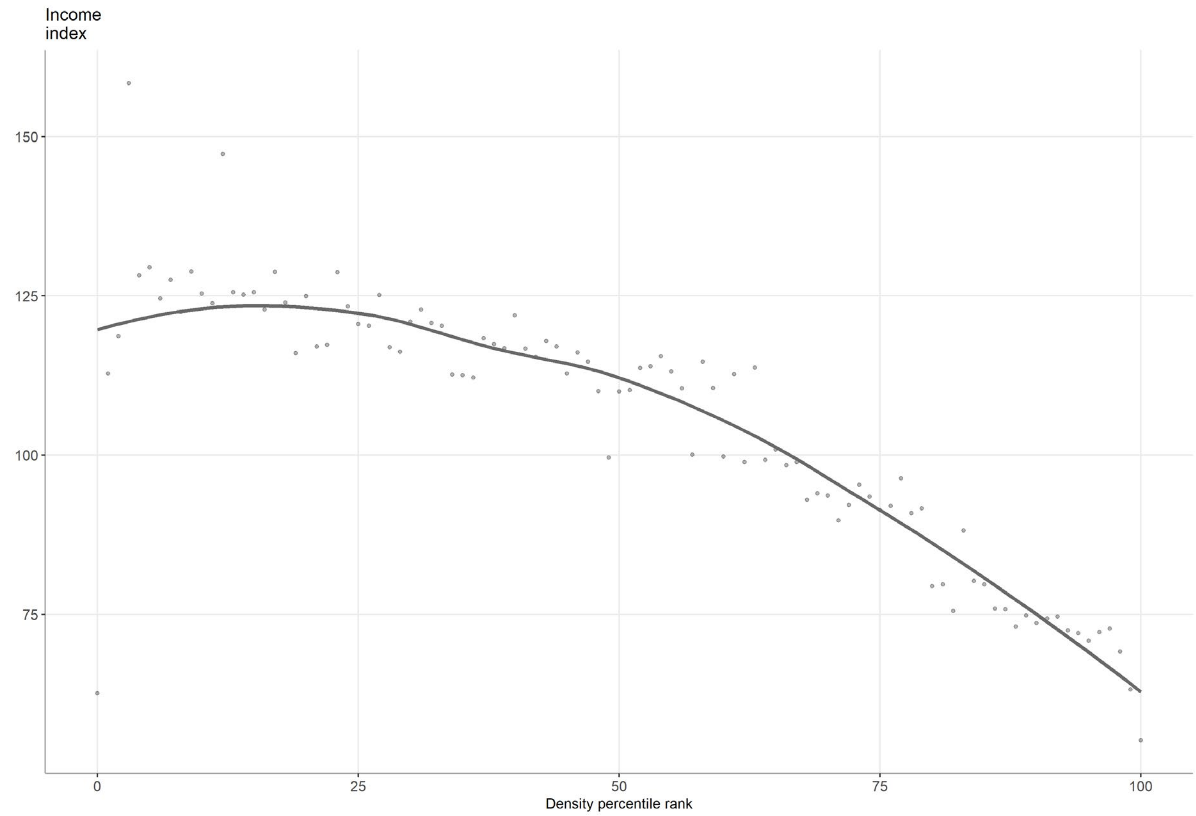

Figure 6. Neighborhood Density and Income in Los Angeles

Using the rank of density of every neighborhood in Los Angeles, we calculate the median income for each percentile and plot them agains the percentiles. The lowest percentile is the lowest density. Median income is normalized by the regional median income. That is, 100 is the point at which the percentile median income is the same as the regional median income.

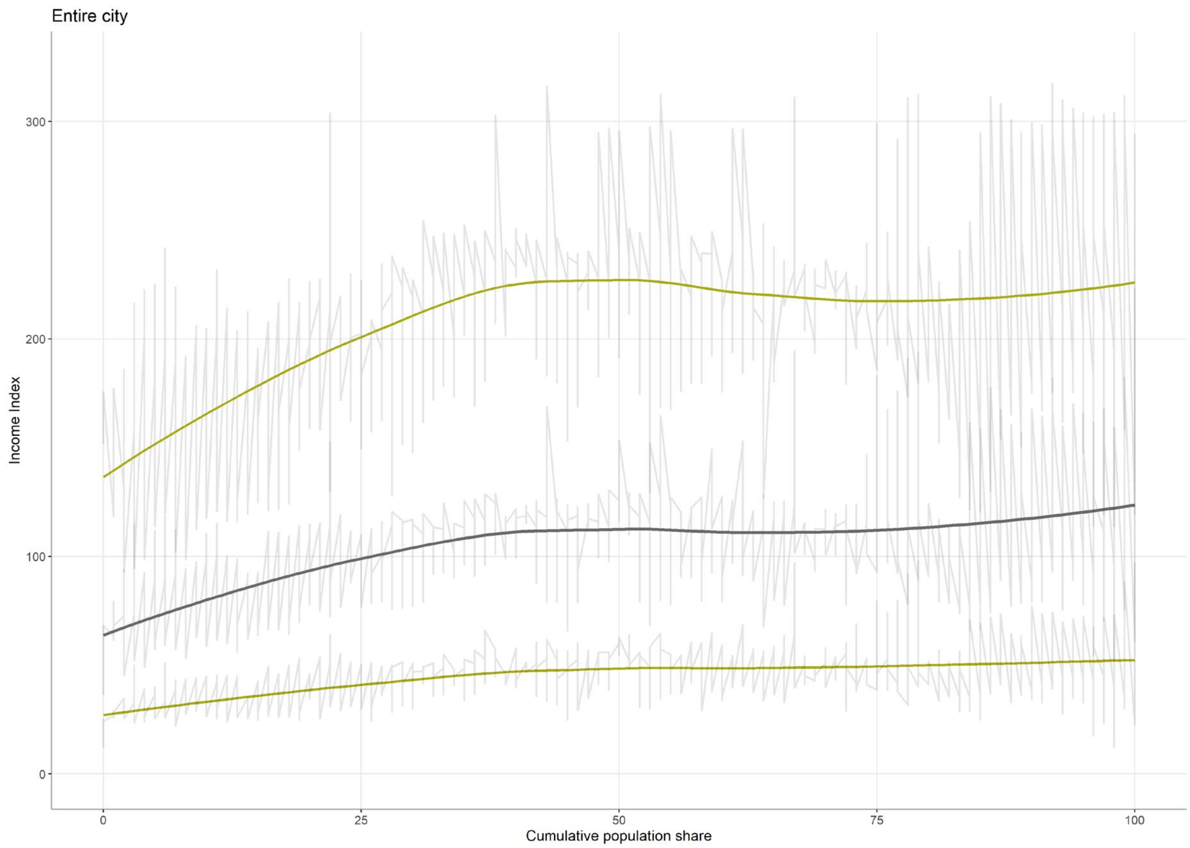

Figure 7. Income over distance for each of the main cardinal directions

Every neighborhood was ordered by their distance from the center. Then, the cumularive population was calculated so that distance is not represented in kilometers but in terms of population. In other words, 50 is the distance where 50% of the popualtion is closest to the city center. The heavy grey line is the median income and yellow lines are the top and bottom quintiles of income. The light grey lines are the fluctuations by cumulative population percentile.

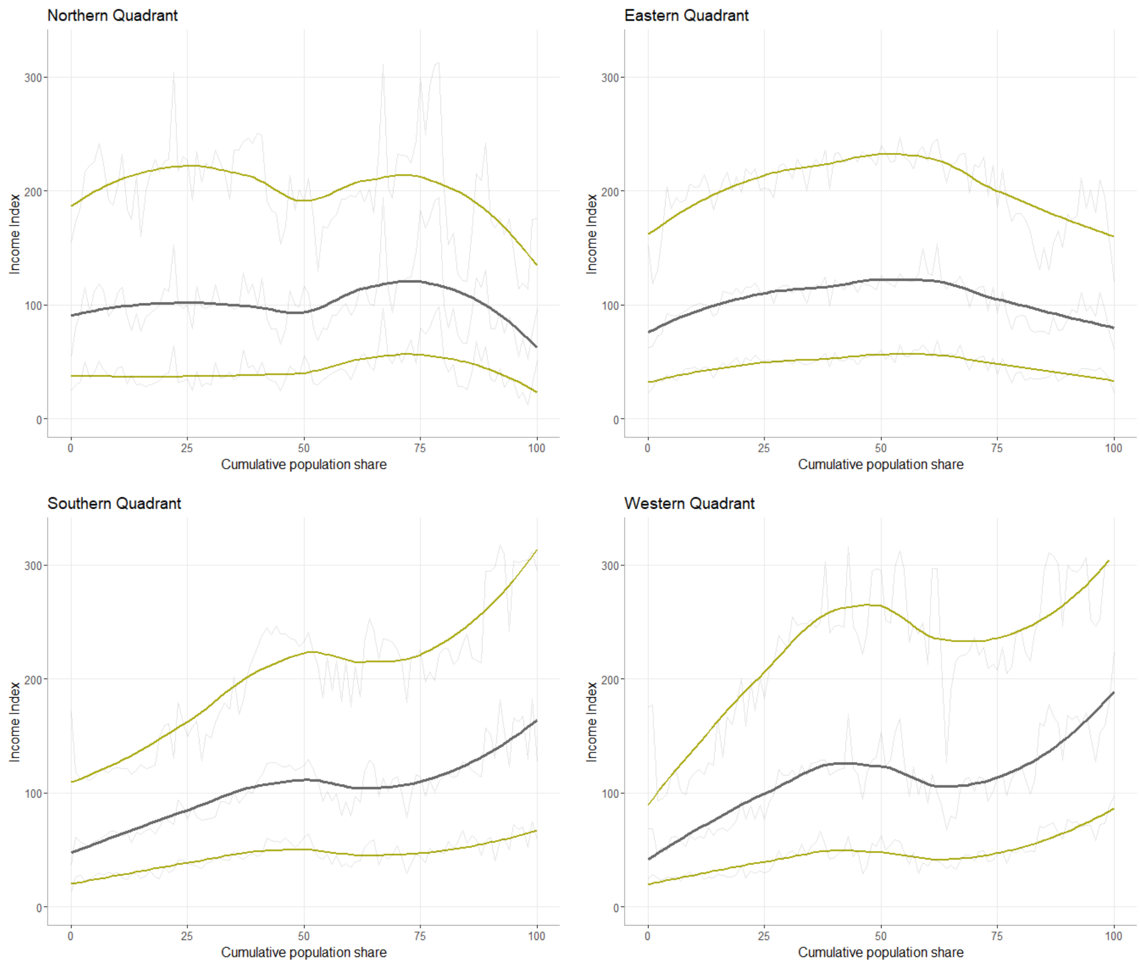

Figure 8. Income over distance for each of the main cardinal directions

Every neighborhood within each quadrant was ordered by their distance from the center. Then, the cumularive population was calculated so that distance is not represented in kilometers but in terms of population. In other words, 50 is the distance where 50% of the popualtion is closest to the city center. The heavy grey line is the median income and yellow lines are the top and bottom quintiles of income. The light grey lines are the fluctuations by cumulative population percentile.

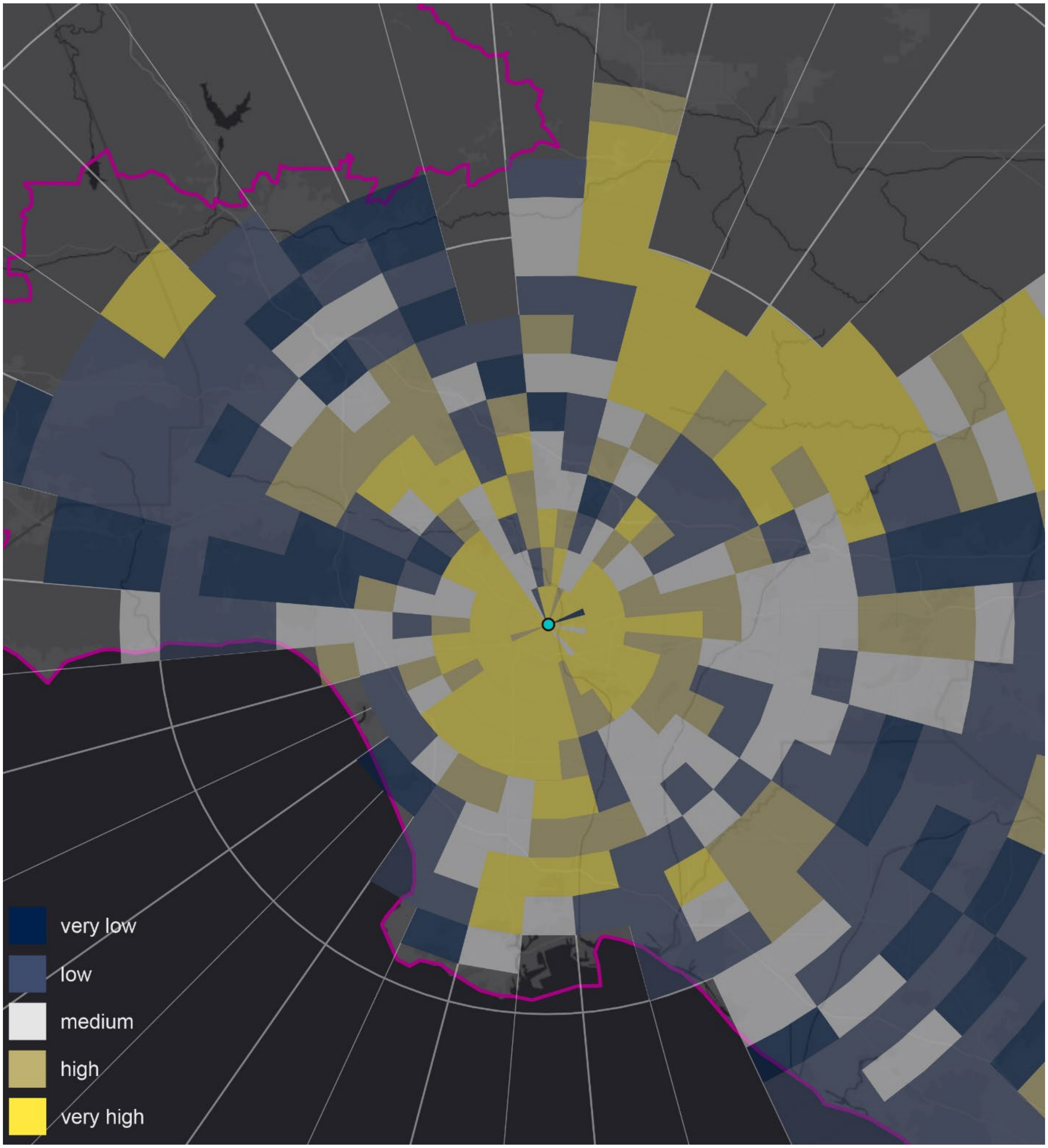

Figure 9. Variation in Neighborhood Income by Sector in Los Angeles

This overlays the diagram of sectors onto a map of Los Angeles. For each sector we calculate the share of the popualtion with income above the regional top quintile. If the distribution were even, every sector would have 20% of their population above that level. A low or high level indicates that there is a deviation of less between 5 and 10 percentage point. Very high or low are sectors with a deviation of more than 10 percentage points.