© OECD 2018

PART I

Chapter 2

Urban sprawl as a multidimensional phenomenon

This chapter develops a conceptual framework in which the characteristics of urban sprawl are systematically categorised and disentangled from its causes and consequences. Furthermore, it provides a functional definition of urban sprawl that reflects its multidimensional nature and its main differences from other forms of urban development. Using a series of graphical illustrations, the chapter describes various features of urban development that determine urban sprawl: average population density and its variation across urban space; the share of urban footprint and population in areas where population density lies below certain thresholds; fragmentation of urban land; number of peaks of high population density; and the fraction of urban population residing outside these peaks.

2.1. Introduction

This chapter develops a conceptual framework for the phenomenon of urban sprawl, in which its various intrinsic characteristics are isolated from its causes and effects and are disentangled from each other. This effort results in a definition that is based exclusively on different features of a development pattern: urban sprawl manifests itself through various ways population and built environment may be distributed across urban space. This multidimensional definition can help in redirecting part of the policy makers' attention, which in many cases is concentrated on average population density, to other features of the development pattern.

The chapter highlights several cases in which a considerable fraction of the urban fabric may be bound to problematic development patterns, while the average population density of the entire urban area is high. Urban sprawl's potential consequences, such as reduced accessibility, car dependency and environmental degradation, as well as others discussed in Chapter 4, do not constitute part of its definition. Such a value-free definition facilitates policy analysis, as it allows for an objective assessment of the effects of urban sprawl. Importantly, several manifestations of the phenomenon are not necessarily detrimental from an environmental or economic viewpoint. For example, fragmentation of urban fabric and low population density may simply reflect physical limitations, rather than policy constraints affecting urban development. The definition of urban sprawl provided in this chapter is subsequently operationalised in Chapter 3, where the various sprawl dimensions are being measured in more than 1 100 urban areas of 29 OECD countries.

2.2. The multiple dimensions of urban sprawl

Urban sprawl has been defined in various ways in different scientific disciplines: economists, geographers, environmental scientists, urban and transport planners have attached a different meaning to it. For economists, sprawl is the *excessive* spatial growth of cities (Brueckner, 2000). The term excessive refers to the part of the urban growth driven by an increasing land uptake *per capita*, rather than by the increasing population of a city. Therefore, sprawl is a synonym of low population density and an antonym of compactness. The latter is associated with high population density and a small *per capita* uptake of land.

The definition used by economists has been criticised by other disciplines as being too simplistic to describe such a complex phenomenon as sprawl. Geographers and urban planners offer a series of alternative definitions of urban sprawl, whereby it is described as a specific pattern of urban development. In most of the cases, definitions are implicit. In order to extract the sprawl definition implicitly adopted in each study, someone has to examine the measurable attributes used to construct complex sprawl indicators. For instance, Tsai (2005) defines sprawl as a function of the degree to which economic activity is evenly distributed and the extent to which high-density suburbs are clustered in space. Torrens (2008) measures sprawl by combining different metrics that describe the location of artificial areas within a city, the allocation of activity across them and the degree to

which these areas are fragmented, decentralised and accessible. Galster et al. (2001) define sprawl as a condition of land use that is represented by low values on one or more of the following urban form metrics: density, continuity, concentration, clustering, centrality, nuclearity, mixed uses, and proximity. Frenkel and Ashkenazi (2008) imply that urban sprawl is indicated, among others, by the degree of irregularity in the shape of artificial areas and the degree to which land use is mixed, including presence of public facilities and institutions.

Many of these approaches to define and measure urban sprawl are problematic for various reasons. The first reason is that the phenomenon of urban sprawl is confused with its own potential causes or consequences. This leads to tautological definitions that attempt to maximise the correlation between the proposed sprawl indicators and a series of other phenomena including reduced accessibility, car dependency and environmental degradation. Following these approaches, sprawl can be any set of characteristics that may be environmentally and economically detrimental in an urban context. Consequently, a city is sprawled because such negative features are present.

Second, many of the proposed metrics of urban form in the literature seem arbitrary. These include, for instance, measures of porosity and peculiarity in the shape of individual fragments of built area. While the latter may be strongly correlated with undesirable phenomena attributed to urban sprawl, there is limited empirical justification that such correlations could represent causal channels in a valid theoretical framework. In some cases, complex sprawl indicators are constructed to encapsulate some of these alleged aspects of sprawl. The outcome has little policy relevance. Even when these indicators are cause- and effect-free, it is hard to interpret them and even harder to determine policy interventions based on them.

This report defines urban sprawl as an *urban development pattern characterised by low population density that can be manifested in multiple ways*. That is, an urban area may be sprawled because the population density is, on average, low. Furthermore, urban areas characterised by high average density can be considered sprawled if density varies widely across their footprint, leaving a substantial portion of urban land exposed to very low density levels. Urban sprawl can also be manifested in development that is discontinuous, strongly scattered and decentralised, where a large number of unconnected fragments are separated by large parts of non-artificial surfaces.

This definition departs from the main viewpoint of the economics discipline on the phenomenon, namely that sprawl is described exclusively by the average population density in an urban area. Following the proposition of other scientific disciplines (see, for example, Ewing and Hamidi, 2014), the definition is modified to account for the entire distribution of population density across an urban area, rather than its average value. Apart from the average population density in a given urban area, sprawl is characterised and measured by: i) the variation of that density across space; ii) the percentage of population living below different population density thresholds; iii) the degree to which population density peaks in one or several locations; iv) the extent to which the density distribution is fragmented; and v) the percentage of population living in artificial areas with relatively high population.

The various dimensions of urban sprawl are summarised in descriptive terms in Table 2.1. In Chapter 3, each of these characteristics is operationalised by one or more sprawl indicators that are computed for more than 1 100 urban areas in 29 OECD countries. To ensure cross-country comparability, the analysis is conducted at the level of the

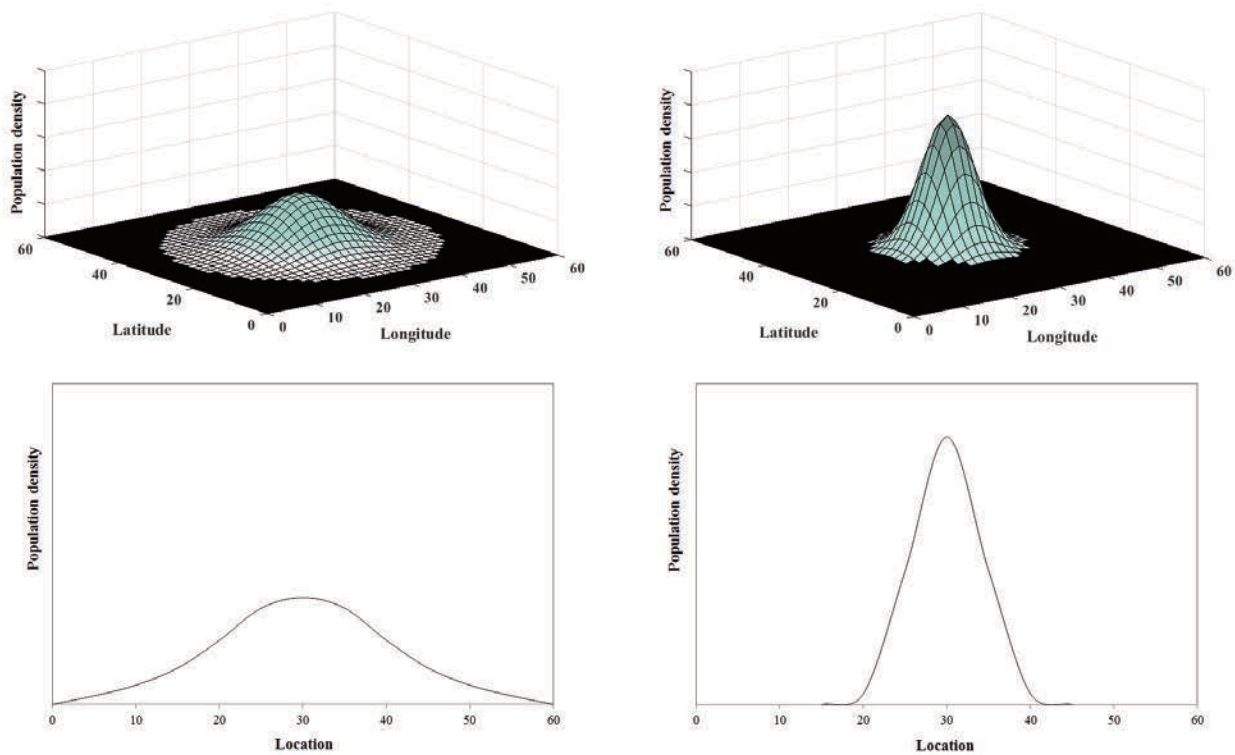
functional urban area (FUA), OECD's economic definition of a city. An FUA comprises a set of contiguous local administrative units, which are characterised by relatively high commuting flows between each other (OECD, 2012).

Table 2.1. **Characterisation of urban sprawl**

Dimension	Explanation
Average urban population density	The average number of inhabitants in a square kilometre of land of an urban area.
Population-to-density allocation	The share of population living in areas where population density is below a certain threshold.*
Land-to-density allocation	The share of urban footprint occupying areas where population density lies below a certain threshold.*
Variation of urban population density	The degree to which population density varies across the city.
Fragmentation	The number of fragments of urban fabric per km^2 of built-up area.
Polycentricity	The number of high-density peaks in an urban area.
Decentralisation	The percentage of population residing outside the high-density peaks of an urban area.

* The thresholds used in the corresponding calculations are 1 500, 2 500 and 3 500 inhabitants per km^2 (for details, see Section 3.4).

Figure 2.1. **Geographic distribution of population in two monocentric cities with different total land uptake**



Note: The left panels display a city where built-up area per capita is high (low average population density), whereas the right panels display a city where built-up area per capita is low (high average population density).

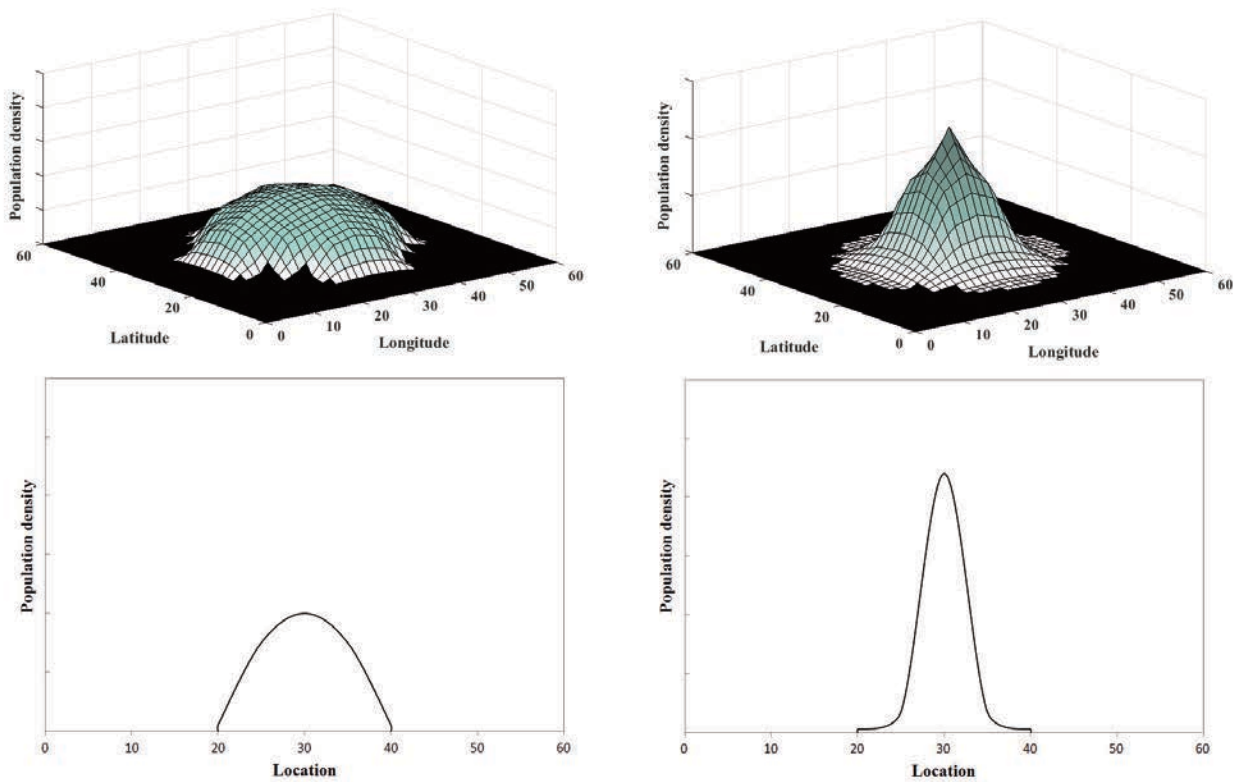
The characteristics of urban sprawl considered in this report are illustrated in Figures 2.1-2.7. To simplify discussion and avoid any loss of generality, the footprints of all urban areas are depicted as squares of equal surface that accommodate the same population size. The coloured parts of the footprints (squared pattern) designate artificial areas, i.e. areas covered by buildings, roads and other structures, while the black parts denote undeveloped land within the urban area, where population density is zero by definition. In each figure, the upper panels depict the *geographic distribution of population*.

This function associates every location within an urban area with a local population density. To simplify the exposition, lower panels in each figure provide illustrative, two-dimensional examples of the population density.[1]

Figure 2.1 depicts two urban areas of equal population that differ fundamentally with respect to their total artificial footprint. As a result, the *per capita* land uptake and its inverse, average population density, i.e. the average number of people per square kilometre, differ widely between the two cases. The left panels of the figure display an urban area that is sprawled as average population density is low. In this case, low average density results from the expansion of the urban fabric into periurban areas.

In contrast to Figure 2.1, Figure 2.2 depicts two cities of equal population and urban footprint. That is, the average population density is identical in the two cases. However, population density displays a larger *variation* across urban space in the latter case (right panels), with a substantial part of the urban footprint being characterised by either very high or very low population density.

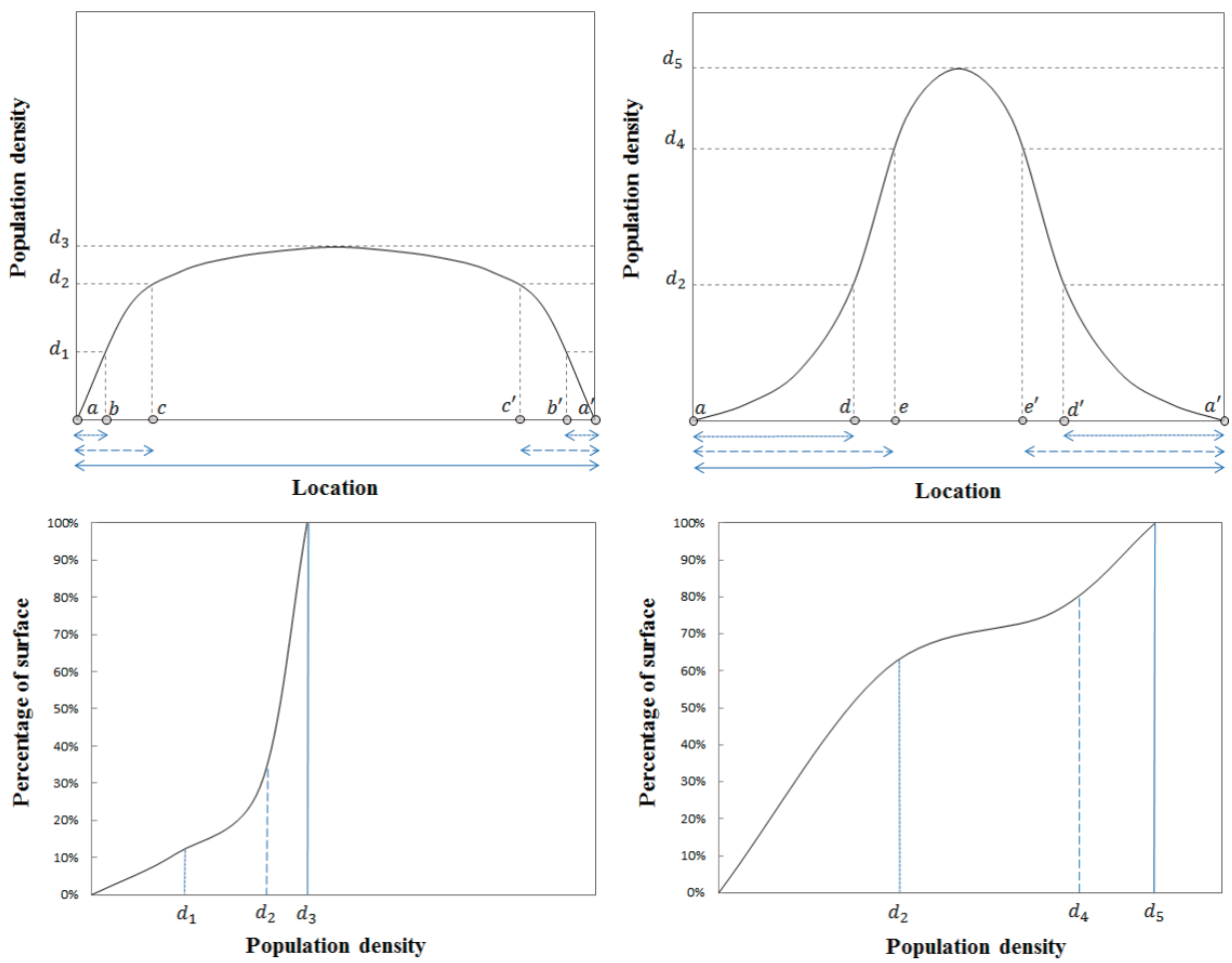
Figure 2.2. **Geographic distribution of population in two monocentric cities with identical population and land uptake**



The implications of density variation across urban space are further examined in Figure 2.3 and Figure 2.4. In both figures, the upper panels illustrate the distribution of population density over space in two hypothetical cities with the same land uptake and population. Figure 2.3 illustrates the derivation of a *land allocation function* from the geographic distribution of population. The land allocation function associates any level of population density in an urban area with the percentage of artificial surface where

population density lies beyond that level. For example, approximately 80% of the urban area depicted in the right panels of Figure 2.3 hosts residential densities below d_4 (lower right panel). In the upper right panel, this percentage corresponds to the ratio of the length of the long-dashed arrows (i.e. the sum of intervals ae and $e'a'$, which represent urban land hosting density below d_4) to the length of the solid arrow (i.e. interval aa', which represents the total populated area). Similarly, more than 60% of the same urban area hosts density levels below d_2. This corresponds to the ratio of the length of the short-dashed arrows (i.e. the sum of intervals ad and $d'a'$) to the length of the solid arrow (i.e. interval aa') in the upper right panel.

Figure 2.3. **Derivation of the land-allocation function from the geographic distribution of population in two cities with identical land uptake and population**

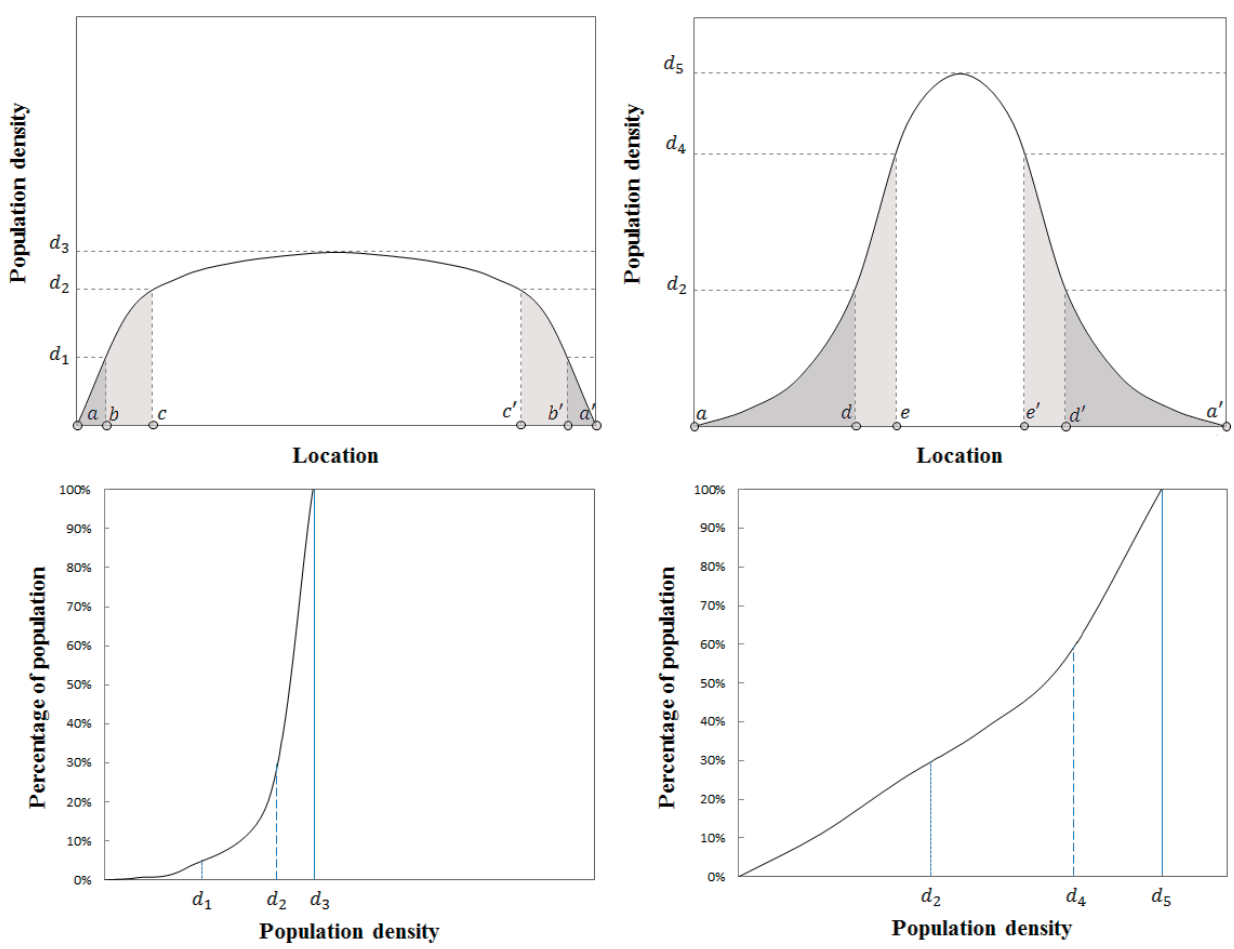


The land allocation function encapsulates important information about the internal structure of cities and the social cost of providing local public goods, such as roads, and services, such as public transport and water supply. In the city depicted in the right panels of Figure 2.3, a substantially larger share of surfaces are exposed to low population density than in the city depicted in the left panels.[2] Hence, public goods and services subject to economies of density, as those mentioned above, may be significantly more expensive to provide in the city of the right panels.

The derivation of a *population allocation function* from the geographic distribution of population is illustrated in Figure 2.4 (lower panels). That function indicates the percentage of population residing in all locations of the city where density lies below a certain number. For example, roughly 60% of the urban area depicted in the right panels of Figure 2.4 hosts residential densities below d_4 (lower right panel). In the upper right panel, this percentage corresponds to the ratio of the population residing in areas of density below d_4 (i.e. the sum of the light and dark grey shaded areas) to the total population of the city (i.e. the total area below the density curve). Similarly, less than 30% of the population resides in areas where density is lower than d_2, which is the portion of the total area under the density curve shaded in dark grey.

While the land-allocation function and population allocation function are strongly correlated, the information they provide is complementary. For example, while both cities displayed in the illustration host roughly 30% percent of their population in areas of density below d_2 (lower panels of Figure 2.4), the percentage of urban land occupied by these areas differs by approximately thirty percentage points (35% versus 65%), as suggested by the lower panels of Figure 2.3.

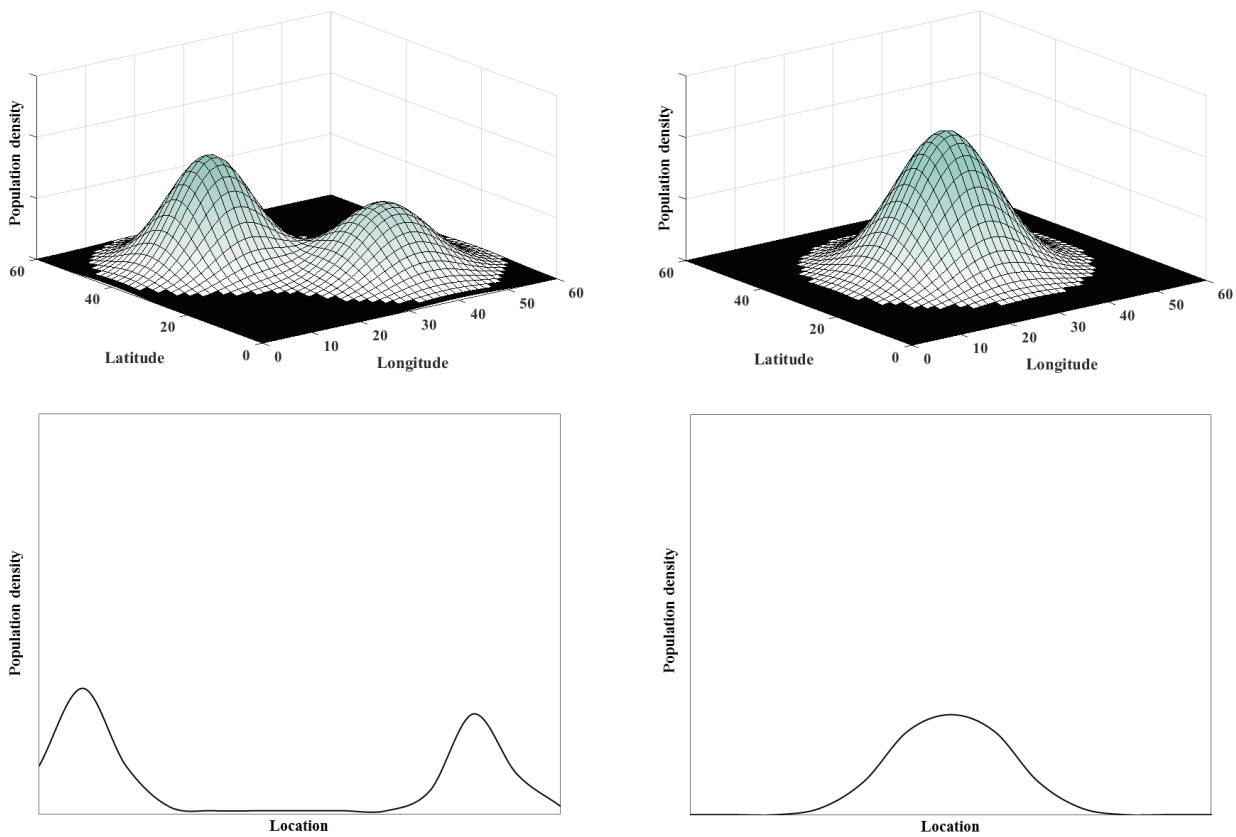
Figure 2.4. **Derivation of the population-allocation function from the geographic distribution of population in two cities with identical land uptake and total population**



By taking into account the entire distribution of population density, it becomes clear that urban sprawl can occur even in cities of high average population density. As illustrated in Figure 2.3 and Figure 2.4, urban areas of identical footprint and average population density may be characterised by completely different internal structures. That aspect has either been neglected by a large number of earlier studies on urban sprawl, or its importance has not been articulated in a clear way. The implications of residential density variation are clarified further in Chapter 4, where the impacts of urban sprawl are discussed in detail.

The difference between monocentric and polycentric urban settings is depicted in Figure 2.5. In a polycentric setting, residential density peaks at multiple locations that lie at considerable distance from each other. Polycentric settings are likely to require a larger urban footprint to accommodate a given population, even though this will not necessarily be the case in all types of urban structure. The environmental implications of adjusting residential density in monocentric and polycentric urban areas are discussed in Chapter 4.

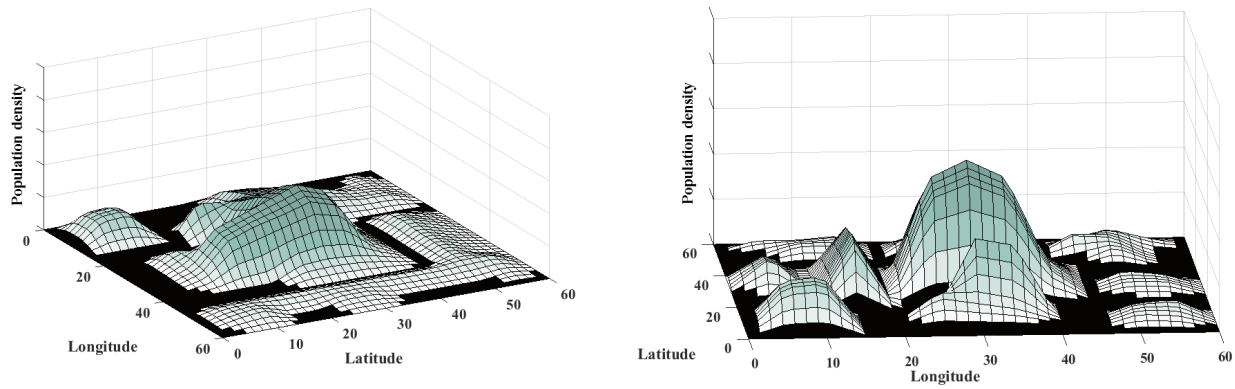
Figure 2.5. **Polycentric versus monocentric cities of identical population**



Note: Left panels display a bicentric city, i.e. a city with a pair of peaks in population density. Right panels display a monocentric city, i.e. a city with a single peak in population density.

The concept of fragmentation is illustrated in Figure 2.6. A fragmented urban fabric is not contiguous. Instead, it is split into a set of *patches* that are scattered across the urban area. While fragmentation is used to describe the extent of discontinuity in a city's built environment, it contains no information on how population is distributed on the various patches of urban fabric.

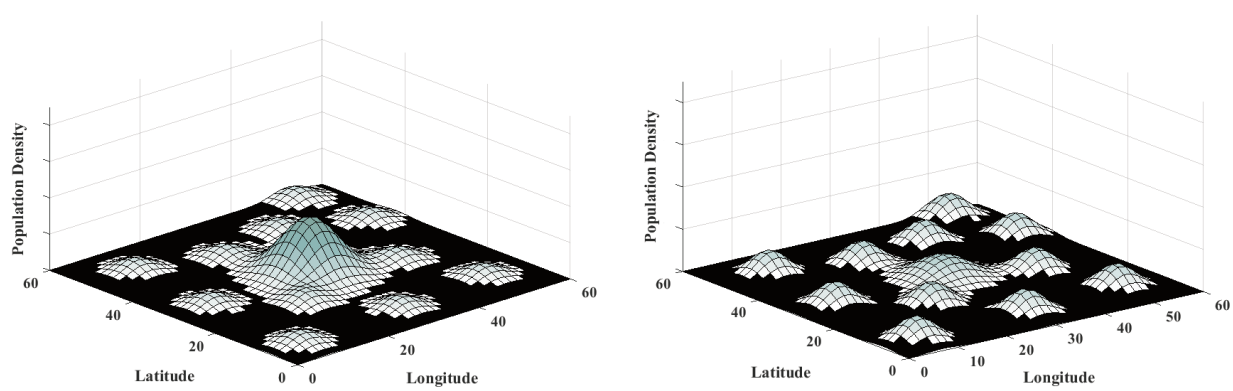
Figure 2.6. **Urban areas of low and high fragmentation**



Note: Left panel displays a city with relatively low fragmentation, while right panel displays a city with relatively high fragmentation.

Decentralisation reveals how unequally population is distributed between areas of peak density and the rest of the urban area. The concept is graphically shown in Figure 2.7.

Figure 2.7. **Centralised and decentralised urban areas**



Note: Left panel displays a centralised city, while right panel displays a decentralised one.

2.3. Summary

This chapter provided a new definition of urban sprawl that is based entirely on the intrinsic characteristics of an urban development pattern, rather than on its causes or effects. Sprawl is defined through several features characterising the distribution of population and built surfaces across an urban area. Therefore, the definition allows different manifestations of the phenomenon beyond low average population density: high variation of population density across space; a high share of urban footprint and population occupied by areas in which population density lies below certain thresholds; fragmentation of the urban fabric; a large number of high-density peaks within a city; and a low fraction of urban population residing in them. Some of these sprawl dimensions, which are further operationalised and measured in Chapter 3, are not always detrimental from an environmental or economic viewpoint. Chapter 4 provides an extensive literature review that attempts, whenever possible, to map the causes and effects of urban sprawl to the different sprawl dimensions presented in this chapter.

Notes

1. Each density peak in the lower panels is represented by a single location in the horizontal axis, e.g. x^*, and the density in a certain location, e.g. x_A, can be considered as an approximation of the average population density in any location at distance $|x^* - x_A|$ from the density peak. For example, the density in location $x_L = 20$ can be considered as the average population density in locations whose longitude is smaller than 30 and whose distance from the density peak point $x^* = (30,30)$ is equal to 10 kilometres. Similarly, the density in location $x_R = 40$ can be considered as the average population density in locations whose longitude is bigger than 30 and whose distance from the density peak point $x^* = (30,30)$ is 10 kilometres.
2. For example, for density level d_2 the difference is large: roughly 35% for the city in the left versus more than 60% for the city in the right.

References

Brueckner, J.K. (2000), "Urban sprawl: Diagnosis and remedies", *International Regional Science Review*, Vol. 23/2, pp. 160-171.

Ewing, R. and S. Hamidi (2014), *Measuring Sprawl 2014*, Report prepared for *Smart Growth America*, available online at: *www.smartgrowthamerica.org/app/legacy/documents/measuring-sprawl-2014.pdf*.

Frenkel, A. and M. Ashkenazi (2008), "Measuring urban sprawl: How can we deal with it?", *Environment and Planning B: Planning and Design*, Vol. 35/1, pp. 56-79.

Galster, G. et al. (2001), "Wrestling sprawl to the ground: Defining and measuring an elusive concept", *Housing Policy Debate*, Vol. 12/4, pp. 681-718.

OECD (2012), *Redefining "Urban": A New Way to Measure Metropolitan Areas*, OECD Publishing, Paris, *http://dx.doi.org/10.1787/9789264174108-en*.

Torrens, P. (2008), "A toolkit for measuring sprawl", *Applied Spatial Analysis and Policy*, Vol. 1/1, pp. 5-36.

Tsai,Y. (2005), "Quantifying urban form: Compactness versus sprawl". *Urban Studies*, Vol. 42/1, pp. 141-161.

From:

Rethinking Urban Sprawl

Moving Towards Sustainable Cities

Access the complete publication at:
https://doi.org/10.1787/9789264189881-en

Please cite this chapter as:

OECD (2018), "Urban sprawl as a multidimensional phenomenon", in *Rethinking Urban Sprawl: Moving Towards Sustainable Cities*, OECD Publishing, Paris.

DOI: https://doi.org/10.1787/9789264189881-4-en

This work is published under the responsibility of the Secretary-General of the OECD. The opinions expressed and arguments employed herein do not necessarily reflect the official views of OECD member countries.

This document and any map included herein are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

You can copy, download or print OECD content for your own use, and you can include excerpts from OECD publications, databases and multimedia products in your own documents, presentations, blogs, websites and teaching materials, provided that suitable acknowledgment of OECD as source and copyright owner is given. All requests for public or commercial use and translation rights should be submitted to rights@oecd.org. Requests for permission to photocopy portions of this material for public or commercial use shall be addressed directly to the Copyright Clearance Center (CCC) at info@copyright.com or the Centre français d'exploitation du droit de copie (CFC) at contact@cfcopies.com.