

Urban Sprawl

How useful is this concept?

Gerald Franz, Gunther Maier, Pia Schröck,

Vienna University of Economics and Business Administration, Vienna, Austria

1. INTRODUCTION

Being at the beginning of a research project on measuring urban sprawl in the Vienna Region with the final goal to find out if urban sprawl exists in Vienna and its Metropolitan Area, this critical literature overview is the first step to shed some light in the debate about sprawl and to clarify the concept in our own understanding. We will discuss what sprawl is, where it originates from, what its causes, symptoms, and consequences are, and how one can measure it in order to diagnose it and maybe find solutions, if there is a problem.

The term urban sprawl is widely used in many disciplines dealing with urban development and urban form. It is moreover used as a current buzzword in urban politics as well as in the public discussion about the growth and form of cities. Since it is used by different disciplines and carries various meanings, the use of the term “urban sprawl” often leads to misunderstandings and confusion. The Michigan Land use leadership council (2005) states: “In the field of land use and planning, no other word spurs more controversy and confusion than sprawl”.

When it comes to writing a scientific paper about sprawl, this controversy and confusion generates major problems. It is necessary to give a broad and detailed theoretical overview about the phenomenon of urban sprawl, in order to sort out the various meanings and components of the terms involved, before one can begin any empirical analysis on this subject. Particularly, when it comes to measuring sprawl, a lack of understanding concerning the concept of sprawl impedes the application of the most appropriate methodological approach, and obtaining significant and valid results.

The paper is divided into six sections. After this brief introduction, the second section deals with the term sprawl in general: where does it come from, and what are its main difficulties. The third section tries to give an overview over types of definitions and definitions of urban sprawl, as there are different ways to conduct this. Section four tries to divide three often confused aspects of urban sprawl: what are its causes, what are its characteristics respectively conditions and finally what are the consequences of sprawl. Section five gives a short insight on how one can measure urban sprawl, including general aspects to consider and different attempts to measure sprawl. In section six we finally draw some conclusions. Within this paper urban sprawl and sprawl will be used as synonyms.

2. THE TERM URBAN SPRAWL

2.1. Origin of the term

The term “sprawl” was first used in 1937 by Earle Draper of the Tennessee Valley Authority in the context of a national conference of planners (cit. in Wassmer 2002). Sprawl was referred to as an unaesthetic and uneconomic settlement form. According to Wassmer (2002) the term “urban sprawl” was first used in the opening paragraph of an article by the sociologist William Whyte in Fortune magazine in 1958. Planners have since then used the term to categorize an urban development, generating undesired social effects. Urban Economists also adopted the term and added to the debate terms like scatter, leapfrogging and ribbon development (explained later). The Real Estate Research Corporation inaugurated in 1974 the controversial debate on positive and negative effects of sprawl (Real Estate Research Corporation, 1974). In the 1990s the phenomenon of sprawl was adopted by other sciences as well as the general public in the US. At this time the Anti-sprawl-movement arose and first measures of urban sprawl were conducted.

Small (2000) argues that the public and policymakers often use the term as a medical analogy. Urban sprawl is seen as a disease, detected by its undesirable symptoms. Many cures are offered for this disease, although we seem to be lacking a solid understanding of the underlying causes and mechanisms. Al Gore (cit. in Wassmer 2002) argues that sprawl has become the concept of the “enemy”, without understanding exactly what it really is.

As the concept of sprawl was “invented” in the US, it was anchored in the US context and discussion for a long time. Europe has the reputation of dense, multifunctional cities with

mixed uses. The image of lively towns, with a strong historic background, is dominant. In recent years, however, the debate has spilled over to Europe. This is justified at least in part by the obvious development problems of many European cities, as the continent presents a very scattered puzzle of territorial conditions.

2.2. Difficulties with the term urban sprawl

As will be discussed in more detail in section 3 of this paper, urban sprawl is notoriously difficult to define. Cervero (2000) claims that sprawl ‘is like pornography. It is hard to define but you know it when you see it.’ Others see sprawl as too general a concept: “the term is so abused that it lacks a precise meaning and defining sprawl has become a methodological quagmire“ (Audirac et al., 1990).

There are various reasons for these difficulties. First, the term is used in a scientific context as well as in public and political discussions. Second, it is used by various scientific disciplines in different manners and from different perspectives. Third, the term is so broad, that it leaves plenty of room for (mis-)interpretations. Fourth, causes, characteristics and consequences of sprawl are often confused. Fifth, it is hard to distinguish sprawl from related terms like suburbanization, urban growth, or suburban development. Sixth, there is no agreed way of measuring sprawl, partly because of the lack of a generally accepted definition. Seventh, the term is used for characterizing a situation as well as a process, which invites further confusion. All these aspects will be discussed in more detail later in this paper.

3. DEFINITIONS OF URBAN SPRAWL

In the previous section, we already gave some indirect definitions of urban sprawl. We mentioned some characteristics as well as consequences of sprawl. But, we also stated that sprawl is an ambiguous term that can be viewed and interpreted in different ways. In this section we try to give an overview of the types of definitions of urban sprawl that exist and filter out, where the different approaches to sprawl originate from. This is not only necessary to understand the conflict between various views of sprawl, but also important for evaluating various types of measuring sprawl that will be discussed in the next chapter.

Siedentrop (2005) mentions five quite different types of definitions of sprawl:

1. Definitions of sprawl according to density attributes of a settlement system: these definitions consider low-density forms of settlement, decreasing density und functional decomposition of cities as sprawl. Representatives of these definitions are for example Glaeser and Kahn, 2003, Fulton et al., 2001.
2. Definitions of sprawl that deal with de-concentration processes of urban functions combined with the spatial expansion of urban uses into rural areas, represented by e.g. Glaeser et al. 2003, Pumain 2003 etc..
3. Definitions of sprawl characterized by structure and form attributes of a settlement system. Sprawl is understood as an urban form building process that transforms a former monocentric compact structure into a discontinuous, polycentric and disperse settlement structure (Galster et al. 2000, Torrens, Alberti 2000, et al.).
4. Definitions based on socially relevant effects of land use, e.g. traffic induced effects, loss of fertile soils, etc. (Ewing, 1997, Downs, 1999).
5. Definitions based on normative planning and order perceptions. Unplanned urban development that runs counter to the objectives of spatial development is identified as sprawl (Gassner 1978 et al.).

Generally within these types of definitions, sprawl appears to be a multidimensional phenomenon. Measurement concepts concerning these types will be presented later. Galster 2001 understands sprawl as a condition of land use and states that general approaches to sprawl can be by aesthetics, efficiency, equity, and environmental aspects. Furthermore, he describes the following approaches to defining sprawl in more detail:

- Defining sprawl by an example: sprawl is often defined by one or more examples whereby Los Angeles is often seen as the prototype of sprawl: Robert Geddes (1997) calls it “fragmented, incomplete, ad hoc and uncentred.”
- Defining sprawl by an aesthetic judgement: this normative approach describes sprawl as an ugly form of development. Representatives of this approach (e.g. Clawson 1962) used judgmental adjectives when talking about sprawl.
- Defining sprawl as the cause of an unwanted (negative) externality: definitions reach from judgments on the appearance of sprawl to alleged causal links between sprawl and its effects on land use patterns in the form of environmental, social and

economic costs. These definitions are treating more the consequences than the attributes of sprawl (e.g. Downs 1998).

- Definitions of sprawl as the consequence of an independent variable: hereby sprawl is defined as the consequence of something else, e.g. as the result of land control fragmentation over land use in city area (e.g. Black 1996, Burchell et al. 1998).
- Sprawl is also defined via specific land use patterns. Altshuler and Gomez Ibanez, 1993 as well as Harvey and Clark, 1965 define sprawl as continuous low density residential development at the fringe of a metropolitan area, as ribbon low density development along highways and as leapfrogging development within undeveloped land leaving a spatial patchwork.
- Sprawl can also be seen as a process of development. Sprawl is, as we already discussed, a stage in a development process and not a static condition. Some authors believe that sprawl changes into a non sprawl form by densifying and diversifying. In the literature however is not stated, when this metamorphosis will occur.

Most authors agree that sprawl is a multidimensional phenomenon. The preceding definitions, however, leave much space for interpretations and are not very precise. Moreover, defining sprawl as a multidimensional phenomenon raises the question of how to weight these dimensions. Are all dimensions necessary to call some urban structure “sprawl”? Do we need a minimum number? Are trade-offs allowed?

Following this idea of a multidimensional phenomenon, Galster et al. (2001) define sprawl as a pattern of land use in an urban area that shows low levels on eight sprawl dimensions. We will discuss this set of dimensions in section 3 below.

Longley et al. (2002) emphasized that there are various elements “which feed into a definition of sprawl”. According to them sprawl can be defined in the following ways:

- Sprawl definitions of form: according to this definition, sprawl is a matter of degree, not a specific urban form, ranging from compact to dispersed development. These sprawling urban forms reach from contiguous urban growth over linear patterns of strip development to scattered development (Ewing 1994, Pendall 1999). These types of definitions create difficulties because highly diverse phenomena are all classified as sprawl.

- Sprawl definitions based on land use: land use patterns are also used to define sprawl. The Transportation Research Board (1998) lists characteristics such as low density residential development, homogenous single family residential development with scattered units, non residential uses like of shopping malls etc. This is a broad approach – the essence is that sprawl is described as low density development with a segregation of uses. The use based definitions (e.g. Downs 1999) are less common than form based definitions.
- Definitions based on impacts are primarily represented by Ewing (1994) and Johnson (2001). They argue that sprawl and other urban land forms are a matter of degree. As they are difficult to distinguish it is better to define sprawl by its negative impacts. As examples they identified poor accessibility of related land uses and a lack of functional urban space as impacts of sprawl. In other words, sprawl is defined by its costs. The problem is that all negative aspects of urban form may be attributed to urban sprawl, making it the “enemy” Al Gore referred to.
- Definitions of sprawl based on density: many definitions of sprawl are based on the concept of density. Sprawl is associated with low density urban development. However, the term low density is relative and “varies with countries and cultural expectations” (Longley et al. 2002). Usually low-density is not quantified – an overview of how to measure density will be given in the next section. In many definitions, density in terms of sprawl represents the relationship between the number of people living in an area and a given land area, which gives indication of land use intensity. According to Longley et al. (2002) a more precise description of density would include all urban land areas including residential, industrial institutional, service, commercial and vacant land.

Obviously, it is tempting to use a broad set of dimensions and fairly lax criteria in order to capture the various versions and variants of sprawl. With such an approach, practically every urban structure can be identified as sprawl. Such a broad definition makes sprawl ubiquitous and the concept almost useless.

Economists usually use some kind of social cost argument to define urban sprawl. Urban growth and decentralization by themselves are not valued negatively. Sprawl occurs when decentralization creates higher net marginal costs on everyone in the metropolitan area than if the development would have remained in the centre (Harvey and Clark, 1965, Mills 1999, Brueckner 2000, Wassmer, 2005). The net marginal costs result from additional

private costs of other households and businesses for locations in the decentralized area and additional public costs that result from the location decisions. As households and firms take into account only their private costs and benefits when making location decisions, they generate externalities for others. To economists urban sprawl occurs when “the net total private and social costs of decentralization, after accounting for private and social benefits are positive” (Wassmer, 2002). This definition has some fundamental problems. First, it defines sprawl by another term – externality – that is extremely difficult to define. Second, it ignores the fact that agglomerations of activities – cities – can exist only when positive externalities bind them together. So, from an economic point of view, we are dealing with a “second best” situation, which is facing multiple distortions, some concentrating others de-concentrating activities. Third, such a definition requires the comparison of people’s utility levels, which is theoretically impossible. Or, as Levine (1997) points out: “What to one person is sprawl to another is his/her home”.

Wassmer (2002) refers to some other economically motivated definitions of sprawl, among them the following:

- The Sierra Club (1998) defines sprawl as “low-density development beyond the edge of service and employment, which separates where people live from where they work and therefore requiring cars.”
- Downs (1998) understands sprawl by observable traits, e.g. unlimited outward extension of new development, low density developments in new growth areas, leapfrog development and strip commercial development.
- Ewing (1997) finds that the terms low-density, strip or ribbon, scattered or leapfrog development are most often used to characterize urban sprawl. He speaks of non-compact development and equates sprawl with the degree of residential inaccessibility (to jobs, shopping, etc.) and a lack of functional open space. In a newer study, Ewing defines sprawl as a “low-density development with residential, shopping and office areas that are rigidly segregated, a lack of thriving activity centres and limited choices in travel routes.” (Ewing, 1997)

Some of the definitions that we have summarized above are more frequently used than others and dominate in a way the discussion of urban sprawl. Some typical examples of specific such definitions are the following: Travisi and Camagni (2005) define sprawl as “the uncontrolled spreading out of a given city and its suburbs over more and more

semirural land at the periphery of an urban area.” They further state that the sprawling process of expansion is disordered, unplanned, leading often to inefficient and unsustainable urban expansion patterns. The European Environmental Agency (EEA 2000) defines sprawl as “the physical pattern of low density expansion of large urban areas under market conditions into the surrounding ones.” Based on these definitions, probably every Western city since it came into existence has sprawled. Since such definitions do not leave sufficient room for “non-sprawl”, they are practically useless. A useful definition of urban sprawl has to provide us with a measure that allows us to clearly distinguish when and where the phenomenon exists and when and where not. None of the commonly used definitions fulfils this criterion.

4. CAUSES, CHARACTERISTICS, AND CONSEQUENCES OF URBAN SPRAWL

Due to Galsters (2001) causes, conditions/characteristics and consequences/impacts of urban sprawl are often confused, due to a semantic wilderness of the term sprawl and empirical deficits. He argues that “a thing cannot simultaneously be what it is and what it causes”. Small (2000) point out that it is hard to find solutions against sprawl if we don't fully understand its causes. Given the current state of the literature, it is a complex task to divide these three aspects. It is necessary, however, to understand the underlying factors of sprawl before measuring their effects. In this section we attempt to divide these three aspects of urban sprawl.

4.1. Causes of urban sprawl

According to Siedentop (2005) there are two rivalling explanation patterns for causes of urban sprawl: firstly sprawl is explained by the demand for urban land. Driving forces are land consumption of households, companies, and public uses. Factors such as income, wealth and car use provide the framework and location choices are made based on a comparison of utility effects and costs. Secondly sprawl is explained by specific regulation patterns. The massive public subsidies for low density, suburban forms of living and the publicly financed construction of street networks and local infrastructure reinforce urban sprawl. According to this view, urban planning is the main cause of sprawl.

Conceptually, the arguments based on the demand for urban land relate to the “monocentric model” of the Alonso-Muth-Mills type. In this model, the externally given central business district (CBD) is the center of the city and the location where all relevant interaction takes place. Households – and in some versions also businesses – choose their location in the surrounding area on the basis of microeconomic constrained optimization. They allocate their income optimally between land, consumer goods, and cost of transportation to the CBD. Sprawl-like phenomena can arise from three factors: declining transport costs, increasing income, and increases in total population. The first two factors yield the same effects. Since households demand more land and can afford longer commutes, density declines near the CBD, but increases at the outer parts of the city. The size of the city increases as agricultural land at the urban fringe is converted to urban uses. As far as the footprint of the city is concerned, an increase in population has the same effect. Density increases in all parts of the city as a reaction to population growth. In percentage terms, this increase in density is much larger at the outskirts than near the CBD. While in the monocentric model we are constrained by assumption to only one centre, in polycentric versions of the model, these driving forces lead to the creation of new subcentres, and can explain the rise of edge cities.

Since the causes are two fundamental economic trends – increasing incomes and declining transport costs – the question arises, whether their logical consequences should be called urban sprawl. Particularly, when taking into account the negative connotations of the term. Mieszkowski and Mills (1993) see these factors as driving forces in a “natural evolution” theory of what causes suburbanization. Gordon and Richardson (1996) speak about “natural economic factors”.

Closely related are some social segregation processes that are often related to sprawl. In the monocentric model, households with higher income will locate at higher distance from the CBD than low-income households, since they allocate more money to transportation to the CBD in their optimal allocation than. This effect is supported further by the lower land prices at the urban fringe.

This process is in line with Tiebout’s argument that people sort themselves into different local jurisdictions based on their preferences for local amenities. The income effect in the monocentric model relaxes the constraint for Tiebout-type self selection and can itself be viewed as contributing to the pull factor of the argument. The segregation process is possibly strengthened by some cumulative feedback loops – sometimes called “flight from

the bright” – that push certain social groups from central locations. The loss of high-income population may lead to higher tax rates, higher crime rates, low-performing public schools, the habitation of poor and minorities in the centre etc.; all factors that will push high and middle class population out of the centre. This factor may be strengthened further by the administrative structure of the city and by the fiscal constitution. When urban core and ring belong to different local jurisdictions which finance their public services mainly out of local taxes, the spatial distribution of income generates a corresponding distribution of public services, which reinforces the segregation processes.

How much of this process is “natural” and unavoidable as long as we welcome rising incomes and declining transportation costs? In our view, given the state of discussion it would be severely misleading to attribute all these changes to urban sprawl. On the other hand, there are substantial structural differences between urban areas so that depending on certain factors these processes may work quite differently. This suggests a concept like the one suggested by Mills (1999), who describes sprawl as “excessive suburbanisation”. Using such a definition of sprawl, however, raises the question of where the “natural” and the “excessive” starts.

4.2. Characteristics of urban sprawl

Many of the aspects which characterise sprawl have been already mentioned indirectly. Burchell et al. (1998) characterise sprawl in two ways: on the one hand residential low-density scattered development and on the other hand non-residential scattered commercial and industrial development. Scattered development is a form that is commonly associated with urban sprawl. He further describes 10 points that characterise urban sprawl – these following characteristics are based on a review of research findings:

- Low residential density
- Unlimited outward extension of new development
- Spatial segregation of different types of land uses through zoning regulations
- Leapfrog (discontinuous) development
- No centralised ownership of land or planning of development
- All transportation dominated by privately owned motor vehicles

- Fragmentation of governance authority over land uses between many local governments
- Great variances in the fiscal capacity of local governments because the revenue-rising capabilities of each are strongly tied to the property values and economic activities occurring within their own borders
- Widespread commercial strip development along major roadways
- Major reliance upon the filtering or “trickle-down” process to provide housing for low-income households.

This categorisation brings a lot of points into the discussion – the problem is that within this list, the limits between causes, characteristics and consequences of sprawl are ambiguous and a clear distinction between these categories is not entirely possible. The 10 points stated can be subdivided in spatial patterns, main causes and main consequences of sprawl.

One of the most elaborated characterisations of urban sprawl is given by Galster et al. (2001). Galster contends that sprawl is characterised by 8 dimensions. We will find these dimensions again, when we talk about measuring sprawl, because he orientates along these dimensions when quantifying the degree of sprawl. Within this section we present these 8 dimensions and their meaning:

- Density: is a widely used indicator of sprawl whereby different types of density can be described.
- Continuity: is the degree to which the unused land has been built densely in an unbroken fashion. Sprawl can be continuous or discontinuous in other places.
- Concentration: describes the degree to which development is located disproportionately rather than spread evenly.
- Clustering: sprawl is frequently clustered what means that it only occupies a small portion of the respective land area.
- Centrality: the loss of centrality is one of the most serious concerns about sprawl.
- Nuclearity: describes the extent to which an urban area is characterised by a mononuclear pattern of development.
- Mixed uses: sprawl is seen as a process that separates the different kinds of land uses (separation of homes, workplaces, conveniences, income segregation along residential communities).

- Proximity: proximity is the degree to which land uses are close to each other (housing, work, shopping, etc.). The lack of proximity contributes to many of the externalities attributed to urban sprawl.

In section 5 when we will discuss the question of how to measure urban sprawl, we will come back to Galster's list of characteristics of urban sprawl.

4.3. Consequences of urban sprawl

According to OECD (2000), urban sprawl has a range of negative consequences. Frequently mentioned consequences are: green space consumption, high costs of infrastructure and energy, an increasing social segregation and land use functional division. Furthermore, the need to travel, dependence on the private car and as a consequence increased traffic congestion, energy consumption and polluting emissions are associated with sprawl.

Due to Wassmer (2005) a lot of negative urban consequences can be attributed to sprawl, but sprawl also has positive effects. When it comes to negative effects he mentions: the car and its polluting effects, a lack of functional open space, air and water pollution, a loss of farmland, tax dollars spent on duplicative infrastructure, concentrated poverty, racial and economic segregation, a lack of employment accessibility etc. Talking about positive effects of sprawl there have to be considered increased satisfaction of housing preferences, the convenience of car travel, the filling in of leapfrogging land, lower crime rates and better public schools in suburban local governments.

Glaeser et al. (2003) analyse the impacts of sprawl in form of traffic congestion, environmental consequences, infrastructure costs and social consequences. They conclude that cars are producing externalities in form of congestion and pollution. However because of the decentralisation of jobs, the pollution problem is reduced. As people move to edge cities, commutes are getting shorter. Sprawl uses up formerly undeveloped land. But, on the other hand only a small portion of (US) landscape is built-up land, implying that there is no scarcity of land. He further argues that externalities decreased over time per miles travelled. Moreover urban agglomeration economies may be reduced by sprawl and deter overall productivity. However, this must not necessarily be the case. Sprawl cities differ substantially in productive, as a simple comparison of e.g. Detroit and Silicon Valley shows. The only true negative consequences of sprawl are social. The segregation

processes that we have discussed above lead to a sharp social separation: Those who can afford cars live in the suburbs, those who can't in the inner city.

Siedentrop (2005) takes a counter position and identifies the following impacts of sprawl:

- Ecological impacts: Building and sealing of land, as well as indirectly loss of natural potential of soils and the expulsion of endangered animal and plants. According to him the problem is not that agricultural space is used, but the fact that connected agricultural land is destroyed.
- Traffic impacts: It is argued that there is a negative correlation between built density and traffic costs. Inhabitants of densely built cities have to bear lower traffic costs. Efficiency of public transport is higher than in urban areas with lower density. However, critics say that density has little influence on traffic behaviour. Since households and firms suburbanize, radial commuting to the city centre is more and more replaced by cross-commuting within the urban area. With jobs nearby, transportation costs may actually be lower, even in a more decentralized structure. The time cost of commuting would have increased even more without suburbanization.
- Social and health impacts: Sprawl leads to an erosion of functioning urban cores. This has not only social and infrastructural consequences but also impacts on innovation capacity of regional economies – in formless space, creative milieus may develop worse (Cervero et al. 1997). There is a significant connection between broadening of settlements and concentration of poverty in city cores. The degree of social interaction in sprawled areas has decreased (Putnam 1994). On the other hand suburbia is not urban in form, but can be in terms of functions. Critics argue that social heterogeneity and cultural diversity in suburbs is higher than alleged.

The Transportation Research Board (1998) defines consequences of sprawl in the form of costs. The report divides effects of sprawl into five types of costs: public and private capital and operating costs, transportation and travel costs, land/natural habitat preservation, quality of life, and social issues. They further argue that empirical or quantitative data is available in more or less detail concerning these aspects. Benefits of sprawl are often ignored.

Concerning the costs of sprawl there are different debates in the literature: Ewing (1997) supports a compact city form with development through planning while Gordon and Richardson (1997b) are supporting the dispersed pattern of development with market led

development. Another debate concerns the consumer preference for low density living: Gordon and Richardson (1997b) claim that consumers prefer to live in low density development, while Pendall (1999) claims that land use controls and fiscal arrangements can influence the density despite consumer preferences.

As we can see from our discussion, the confusion concerning the definition of urban sprawl is not resolved by looking at the causes, characteristics, or consequences of urban sprawl. Therefore, we will analyse the attempts to measure urban sprawl in the next section.

5. Measuring Urban Sprawl

Numerous attempts have been made to measure urban sprawl. Different scientific methods and approaches were used for this task. Yet it is a difficult thing to do, as the concept of urban sprawl is broad and confusing. It is hard to measure something accurately that you cannot define properly. There are various aspects which have to be taken into account when measuring the degree of urban sprawl. The result, however, is not guaranteed to be widely accepted by the scientific community as different interpretations of causes, conditions and consequences of urban sprawl exist.

5.1. What to consider when measuring urban sprawl

When attempting to measure urban sprawl, a number of general aspects have to be considered. They will be discussed in this subsection of the paper.

5.1.1. The spatial delineation – what area to measure?

First of all, the question occurs where to delineate the fringes of the urban area or in other words what area to observe. Various definitions exist concerning the urban area. The metropolitan area for example comprises the central city and the surrounding counties economically integrated. The definition of urban areas talks about densely settled central places in an urban area plus the less densely settled territory (urban fringe) that surrounds these places (Wassmer 2005). Central places are dominant employment and residential

centres in each urbanized area. Metropolitan areas in the USA are characterized by rising population, rising real incomes and declining transport costs.

Glaeser et al. (2003) argues that for measuring urban sprawl one has to know where the centre is. Often, the CBD is the urban area with the highest density. In a polycentric city, however, there may be competing centres and the highest density with respect to one function may differ from that with respect to another.

5.1.2. The time factor: static vs. dynamic development

Another aspect is the fact that measuring sprawl normally has a static component. An urban area is measured at a given point in time. In order to compare the evolution of sprawl, an urban area has to be measured at different points in time. Then the dynamic aspect of urban sprawl can be observed. This raises the question, however, whether the spatial delineation should be kept constant or should be adjusted between the observations. When we adjust the delineation, we most likely have to add new areas to the analyzed city, so that we actually compare differently defined cities. When we keep the spatial delineation constant, we will either have to include areas that are functionally outside the city, or exclude parts of the functional city in one of the observations.

Another approach which is chosen in many studies is to measure and compare different urban areas within one or more countries. In this case, the issue of compatible definitions and spatial delineations is even more challenging.

5.1.3. Quantitative and qualitative measures on urban sprawl

It is essential for measuring sprawl to select the appropriate scientific approach. One step within this process comprises the decision whether to choose a qualitative or quantitative approach. Qualitative approaches are for example conducted within the URBS PANDENS project. They are nonetheless as complicated as quantitative approaches. More often quantitative approaches are used and shall be the focus of this paper. In a number of studies sprawl is not only measured according to one indicator but on various indicators building a so-called sprawl index.

5.2. Multi-disciplinary measures of urban sprawl

As we already discussed, many measures of sprawl are possible. In this paper we do not just want to promote one measure, but to categorize the different attempts of measuring sprawl. Knaap, Song, Ewing and Clifton (2005) give an overview of various dimensions of measuring sprawl and how different disciplines are connected to these dimensions:

- Metropolitan structure: measuring the metropolitan structure is the oldest form of measuring urban form. It focuses on spatial patterns at the metropolitan scale. Especially economists have used these measures in order to test the varieties of urban form. It contains measures of population density, population density gradients and recently also employment density and distributions. Social scientists are using census data to analyze spatial patterns at the metropolitan scale. Economists have generally found that density gradients behave as predicted by economic theory and conclude that markets are important in structuring urban form.
- Sub metropolitan structure: this approach to measure sprawl is generally conducted by transportation planners. They use census data to explore transportation behaviour. This model recommends less automobile use among high-density mixed use census tracts in form of public policies. It is all about transportation infrastructure which is very costly. Planners try to answer the question: „how to best provide access from one part of the metropolitan area to another“.
- Community design: land use planners use local GIS data to analyze spatial patterns and urban sprawl at a community scale. This approach is relatively new because of technical conditions. They evaluate the merits of plans by their levels of density and mixture of use. The focus is on spatial patterns at very high levels of resolution (the tax lot parcel).
- Urban design: architects and urban designers conduct audits and visual preference surveys (Interviews, etc.) to analyze environments at the neighbourhood and building scale. It is important for them to find preferences for textured urban environments and to give recommendations. This approach is drawn from a rich and well developed body of theory. In this approach scientists observe phenomena

such as crime, social capital, residential satisfaction, transport behaviour etc. The approach contains objective and subjective measures.

- Landscape ecology: this dimension of measuring sprawl used land cover data to analyze landscapes at varying scales. According to them urban growth causes habitat and farmland fragmentation, greater impervious surfaces and thus they recommend that urban growth should be contained. According to landscape ecologists all problems of urban form are scale dependant, place specific and interdependent with similar problems. The approach is based on natural landscapes. The interest is usually on the patch of land not developed for urban use and not the urban uses surrounding the undeveloped patch. This method contains soil classification, assessment of forest conditions etc.

5.3. Measuring sprawl and its impacts

The smart growth movement represented e.g. by the Sierra Club (1998) and Ewing, Pendall, Chen (2002) draws a picture of measuring sprawl from a planning perspective. They also work with different factors that can be analyzed and measured. These measures are based on older attempts to measure sprawl. The authors try to give an overview of various former studies measuring sprawl and their different methodological approaches:

- *Studies using simple measures of sprawl*

USA today (2001) measured sprawl based on two density-related measures: the percentage of a metro area's population living in urbanized areas and the change in the percentage of metropolitan population living in urbanized areas between 1990 and 1999. The advantage of this measure is its simplicity, the disadvantage its total reliance on density as indicator of sprawl.

The Sierra Club (1998) measured sprawl as low-density development beyond the edge of service and employment, which separates where people live from where they shop, work, recreate and educate. It assumes that cars are required to move between zones. In this study sprawl is defined by effects.

Galster et al. (2001) developed the most complex multi-faceted sprawl index to date, where sprawl is characterized by eight indicators. Causes and consequences are excluded. This approach will be analyzed in more detail later on.

- *Studies measuring sprawl and relating it to its outcomes*

Kahn (2001) explored one potential benefit of urban sprawl – the increased housing affordability and greater equality of housing opportunities across racial lines. The author measured sprawl by the degree of employment decentralisation in a metropolitan area. Specifically, he measured the proportion of metropolitan employment located more than 10 miles from the CBD. The study fails to consider residential development patterns as well as multi-centred employment patterns.

Downs (1999) focussed his research on sprawl and its effects on urban decline. He explores the relationship between suburban sprawl and urban decline. The author failed to distinguish causes and consequences from characteristics of sprawl.

- *Studies measuring sprawl and exploring causes*

Glaeser et al. (2003) related sprawl to the degree of decentralisation of employment. In order to explain the differences across metro areas the authors related the measure of sprawl to the age of the metropolitan area and to the degree of political fragmentation within the metropolitan area.

Pendall (1999) relates the measure of sprawl strictly to density. The incidence of sprawl is given in terms of land values, metropolitan political organization, local government spending, traffic congestion and various local land use policies. Adequate public facilities requirements were found to discourage sprawl, while low-density zoning and building capacities were associated with more sprawl.

Fulton et al. (2001) also use a strictly density related concept. They studied urban land consumption relative to population change. If land is consumed at a faster rate than population is growing, sprawl is said to be increasing.

5.4. Multi-factor approach to measure sprawl – the sprawl index

As we can conclude from the previously described methods of measuring sprawl, density is one of the most commonly used indicators. However, the theoretical arguments and

definitions of sprawl that we have discussed above also point to other important factors that – in combination with density – may be better characterizing urban sprawl. This suggests the combination of various factors in some weighted fashion to form a so-called sprawl index. In this sub-section we will present some sprawl indices used currently in the literature.

- *Ewing, Pendall, Chen (2002)*

These authors develop further former studies of measuring sprawl. They define sprawl as low density development with residential, shopping and office areas that are rigidly segregated, a lack of thriving activity centres and limited choices in travel routes. Therefore according to them four factors need to be measured and analyzed:

- 1) Residential density
- 2) Neighbourhood mix of homes, jobs and services
- 3) Strength of centres such as business districts
- 4) Accessibility via the street network

The authors try to combine these four indicators and to measure them empirically. This is done in the following way:

- 1) Density factor: average density is taken for the urban sections collectively. Central diversity is just the intercept of a negative exponential density function.
- 2) Mix factors: is representing the relative balance between jobs and population.
- 3) Centres factor: can exist with respect to population or employment and with respect to a single dominant centre or multiple sub-centres.
- 4) Streets accessibility factor: the block size captures not only the length of block faces but also the extent to which streets are interconnected.

The most difficult issue thereby is how to adjust the sprawl index for the size of the metropolitan area. As metropolitan areas grow, so do their labour and real estate markets and the land prices. As a consequence density gradients shift upward and other measures of compactness follow suit. The main aim for the authors is to find a method of transforming the sum of the four sprawl factors into an index that is neutral to population size.

- *Frenkel et al. (2005):*

The authors try to measure sprawl from a landscape perspective by use of an urban land use survey. According to them the alleged negative impacts of sprawl are a lack of scale

economies (sprawl reduces the level of public services in suburbs and weakens the base of central cities), increased energy consumption, damage to ecosystems, etc. They define sprawl as “a form of spatial development, characterized by low densities, scattered and discontinuous leapfrog expansion, segregation of land uses, all encouraging a massive use of private vehicles and strip malls in open rural lands at the edge of metropolitan areas.” According to Frenkel et al. (2005) sprawl can be measured along the following indicators:

- 1) Growth rates: are measured by sprawl quotients, i.e. the ratio between the growth rate of built-up areas and the population growth rate.
- 2) Density: is the most popular sprawl measure – there are many types of densities and many ways to measure it. Here it is seen as ratio between a certain urban activity and the area in which it exists. Urban activity is residential units, number of residents and employees.
- 3) Spatial geometry contains the largest group of measures. It was adopted from ecological research and fractal geometry. Two main characters of urban landscape are configuration (i.e. geometry of a built up area) and composition (i.e. the level of heterogeneity). Common measures are leapfrog and continuity – they quantify scatter and fragmentation of the urban landscape.
- 4) Accessibility: according to Ewing the condition of poor accessibility is followed by the massive use of private vehicles. Measures are for example road length, road areas, travelling times of households.
- 5) Aesthetic measures: sprawl is a boring, homogenous form of development from an aesthetic point of view, what is difficult to measure.

According to Frenkel et al. density cannot be the sole parameter of sprawl and further indices are needed to quantify the phenomenon. The authors use an integrated sprawl index containing configuration (density and scatter) as well as composition (mixed land use).

- *Torrens and Alberti (2000):*

The authors argue that sprawl is a heavily discussed phenomenon, but often not empirically founded. Many measures try to find out consequences of sprawl rather than to measure its characteristics. When measuring sprawl Torrens and Alberti (2000) orient themselves on the “Ewing-Gordon-Richardson-debate” and the related sprawl characteristics: density,

scatter, fragmentation and leapfrogging, aesthetics, ecology and accessibility. According to these indicators the following ways of measuring are possible:

- 1) Measuring density: can be measured by population density gradients. According to them density declines with growing distance from the CBD to the fringe along the gradient.
- 2) Measuring scatter: measured by weighed mean distances. The fractal dimension describes space-filling abilities of sprawl.
- 3) Measuring aesthetics: there exist architectural and photogrammetric approaches. The correlation between pixel “signatures” and sprawl is measured.
- 4) Measuring ecology: measured are the landscape composition (Shannons diversity, evenness Index), landscape configuration, etc.
- 5) Measuring accessibility: possible are opportunity based measures (gravity, how far do I get with a limited time budget?) and utility based measures (spatial choice and decision theory).

According to the authors, there are some aspects to be worried about – among others the following: data availability, scale dependency and sensitivity, measures are static – sprawl is dynamic. Another problem is how to integrate affected people into the equations of measuring sprawl.

- *Galster et al. (2001)* :

We already mentioned the eight dimensions of urban sprawl developed by Galster et al. (2001). In this sub-section, we present how these aspects can be used for measuring sprawl in the form of indicators. The dimensions are defined as follows (Galster et al. 2001):

- 1) Density: is measured as the average number of residential units per square mile of developable land in a urban area.
- 2) Continuity: measured as the degree to which developable land has been built upon at urban densities in an unbroken fashion.
- 3) Concentration: measured as the degree to which development is located disproportionately in relative few square miles of the total urban area rather than spread evenly throughout.

- 4) Clustering: measured as the degree to which development has been tightly concentrated to minimize the amount of land in each square mile of developable land occupied by residential or non-residential uses.
- 5) Centrality: measured as the degree to which residential or non-residential development (or both) is located close to the central business district (CBD) of an urban area.
- 6) Nuclearity: measured as the extent to which an urban area is characterized by a mononuclear pattern of development.
- 7) Mixed Uses: is measured as the degree to which two different land uses commonly exist within the same small area, and this common across the urban area.
- 8) Proximity: is measured as the degree to which different land uses are close to each other across an urban area.

In order to conduct the operationalization of the sprawl dimensions, land has to be divided into 3 types: residential land, non-residential land and nondevelopable land (because of natural features, public use, regulatory barriers etc.). The aim is to obtain a composite sprawl index across all these dimensions that can be calculated mathematically. According to Galster et al. (2001) the following problems occur: first, the appropriate geographic scale – the extended urban area has to be measured. Second, the problem of non-developable land occurs and the question how to ingrate it into the sprawl measure. Galsters dimensions has been partly tested – the problem is that due to its complexity and its high data requirements, only few urban areas have been tested up to now.

All the measures that we have discusses, measure the state and specific characteristics of urban development. If we understand sprawl as an excessive form of urban development and of suburbanization in particular (Mills, 1999), then none of them really measures sprawl, because they do not distinguish the “natural” from the “excessive” part of change. So, the measures and indices that we have discussed will allow us to place a city of a specific point in time on a scale. But, the question, whether a certain point on this scale indicates sprawl or not, can only be answered by theory, not by the measure itself. Even the most sophisticated measure or index cannot overcome the problem of a vague definition of the phenomenon.

6. Conclusions

In preparation for an empirical study of measuring urban sprawl for the city of Vienna, Austria, we have dealt with the concept of urban sprawl in this paper. We trace the history of the term and various problems associated with it in section 2 of the paper. In section 3 we discuss definitions of urban sprawl that exist in the literature. As it turns out, most these definitions are vague. They mix together causes, characteristics, and consequences and are in most cases so general that they cannot distinguish clearly between sprawl on the one hand and non-sprawl on the other. In many cases, definitions are based on implicit value statements that make the concept questionable for use in scientific research.

Section 4 deals with characteristics, causes and consequences of sprawl as discussed in the literature. In addition to the vagueness of the term which spills over as a problem from section 3, this section shows that in many contributions we can find a remarkable lack of understanding of the underlying urban development processes. The cumulative nature of these processes also makes it difficult to distinguish clearly between causes, characteristics, and consequences. A consequence in one relation is often the cause of decentralization tendencies in another.

In section 5 we deal with quantitative measures of urban sprawl. Various attempts of measuring urban sprawl can be found in the literature. All of them face serious problems as far as the spatial and temporal definitions are concerned. As with every multidimensional phenomenon, the measurement of sprawl is faced with the problems of what indicators to use and how to weight them in the creation of an index. Moreover, once such a measure is computed, we face the problem of interpreting it. For none of the measures there exists a theoretically well defined threshold that distinguishes sprawl from non-sprawl. So, the decision whether a certain empirically derived value characterizes sprawl or not is in the hands of the researcher, policy maker, or policy activist. From a scientific point of view, this is clearly unsatisfactory.

References

Altshuler, A., Gomez-Ibanez, J.A. (1993), *Regulation for Revenue: The Political Economy of Land Use Exactions*, Washington D.C.

Anas, A. Arnott, R. & Small, K.A. (1998), "Urban Spatial Structure," *Journal of Economic Literature*, American Economic Association, vol. 36(3), pages 1426-1464.

Audirac, I., Shermeyen A.H., Smith M.T. (1990), "Ideal Urban Form and Visions of the Good Life," In: *Journal of the American Planning Association*: 47083.

Black, T.J. (1996), "The Economics of Sprawl", In: *Urban Land* 55 (3), 6-52.

Brueckner, Jan (2000), *Urban Sprawl, Diagnosis and Remedies*, *International Regional Science Review* 23, 17-171.

Brueckner, Jan and Fansler, David (1983), *The Economics of Urban Sprawl: Theory and Evidence on the Spatial Sizes of Cities*, *The Review of Economics and Statistics* 65 (3), 479-482.

Burchell, R.W., Lowenstein, G., Dolphin W.R., Galley, C.C., Downs, A., Seskin, S. , Grey Still, K., Moore, T. (1998), "The Costs of Sprawl – Revisited", TRCP Report 39, Washington D.C.

Camagni, R. (2002), "Urban mobility and urban form. The social and environmental costs of different patterns of urban expansion." In: *Ecological Economics* 40 (2), 199-216.

Cervero, R., Wu, K. (1997), "Polycentrism, Commuting, and Residential Location in the San Francisco Bay Area", In: *Environment and Planning*, 29 (5), 865-886.

Cervero, R., 2000. *Shapeless, Spread Out, Skipped Over and Scattershot – Sprawl Sweeps the Globe*. In: *The World Paper*, March/April, 2000, S. 5–6.

Clawson, M. (1962), "Urban Sprawl and Speculation in Suburban Land", In: Land Economics 38 (2), 94-111.

Dangschat, J., Kratochwil, S., Schwantner, U. (2004), "Individuell ok, aber in Summe problematisch! AkteurInnen rund um Sprawl(ing)", STEP Workshop, Wien.

Downs, A. (1999), "Some Realities about Sprawl and Urban Decline", In: Housing Policy Debate 104 (4), 955-974.

Downs, A. (1998), "How America's Cities are growing: The big Picture", In: Brookings Review 16 (4), 8-12.

EEA - European Environment Agency 2000, "Down to earth: soil degradation and sustainable development in Europe : a challenge for the 21st century", Environmental issue series no. 16, Copenhagen.

Empirica (1996), Wanderungsverhalten in Relation zur Wohnbauentwicklung in Leipzig und Umland. Typoskript. Berlin.

Ewing, R., Pendall, R. and Chen, D. (2002), Measuring Sprawl and its impact, Smart Growth America, Washington D.C.

Ewing, R. (1997), "Is Los Angeles Style Sprawl Desirable?" In: Journal of the American Planning Association 63 (1), 107-127.

Ewing, R. (1994), "Characteristics, Causes and Effects of Sprawl", A literature Review, Environmental and Urban Issues: 1-15.

Frenkel, A. & Ashkenazi, M. (2005), "Measuring Urban Sprawl - how Can we Deal with It?", ERSA conference papers ersa05p50, European Regional Science Association.

Fulton, W. (2001), "Who sprawls most? How growth patterns differ across the U.S.", The Brookings Institution, Washington D.C.

Galster, G., Hanson R., Wolman, H., Coleman, S., Freihage, J. (2001), "Wrestling Sprawl to the Ground: Defining and Measuring an Elusive concept", Housing Policy Debate, Volume 12, Issue 4, Fannie Mae Foundation, 681-717.

Gassner, E. (1978), Zersiedelung im Umland, Problematik und Abwehrstrategie, In: Deutscher Rat für Landespflege (Hrsg.): Verdichtungsgebiete, Städte und ihr Umland, Heft 30, 645-653.

Geddes, R. (1997), "Metropolis Unbound: The Sprawling American City and the Search for Alternatives", In: American Prospect 8 (35), 40.

Glaeser, E. L., Kahn, M. E. (2003), Sprawl and Urban Growth, NBER Working Paper Series No. 9733.

Glaeser, E.L. (2001), "Job Sprawl: Employment location in U.S. Metropolitan Areas", The Brookings Institution, Washington D.C.

Gordon, P., Richardson, H. (1997b), "Where is the Sprawl?" In: Journal of the American Planning Association 63 (1), 107-127.

Gordon, P., Richardson, H. (1996), "Employment Decentralisation in U.S. Metropolitan Areas: Is Los Angeles an Outlier or the Norm", Environment and Planning A 28: 1727 – 1743.

Gottlieb, P. D. (1999), "Do Economists have anything to contribute to the Debate on Urban Sprawl?" Forum for Social Economics 28 (2): 51-64.

Harvey, R. O., Clark, W.A.V. (1965), „The Nature and Economics of Urban Sprawl“, Land Economics 41 (1), 1-9.

Hesse, M., Kaltenbrunner, R. (2005), Zerrbild "Zersiedelung": Anmerkungen zum Gebrauch und zur Dekonstruktion eines Begriffs. In: DISP 160 (1/2005) Urban Sprawl, 16-22.

Jacobs, J. (1969), *The Economy of Cities*, New York: Random House.

Johnson, M.P. (2001), Environmental Impacts of Urban Sprawl: a Survey of the Literature and proposed Research Agenda. In: *Environment and Planning A*, Vol. 33 (4), 717–736.

Knaap, G J, Y. Song, R. Ewing, and K. Clifton (2005), “Seeing the Elephant: Multi-disciplinary Measure of Urban Sprawl.” *Journal of Planning Literature*, submitted, 1-46.

Levine, N. (1997), *American Planning Association Journal* 63 (1), 279-282.

Leontidou, L. Afouxenidis, A., Kourliouros, E. (2002), *Causes of Urban Sprawl in Athens & East Attica, 19981-2001, First Greek Annual Report*, Hellenic Open University.

Longley, P., Batty, M., Chin, N. (2002), *Sprawling Cities and Transport: preliminary findings from Bristol, UK*, Abstract: As part of the Fifth Framework Programme, SCATTER, 1-24.

Michigan Land Use Leadership Council (2005), “Defining Sprawl and Smart Growth” (Working Paper prepared by Public Sector Consultants Inc), 1-9.

Mieszkowski, P., Mills, E. S. (1993), “The Causes of Metropolitan Suburbanization”, In: *Journal of Economic Perspectives*, Volume 7, Number 3, 1993, 135-147.

Mills, E. S. (1999), The Brawl over so-called Sprawl, In: *Illinois Real Estate Letter*, 1-7.

Nelson, A.C. (2001), *Urban Containment Policy*, Atlanta GA.

Nuissl, H., Rink, D.: *Sprawl und Schrumpfung: das Beispiel Leipzig* In: Nuissl, H., Rink, D. (Hrsg.), *Schrumpfung und Urban Sprawl - analytische und planerische Problemstellung*. UFZ-Diskussionspapiere 3/2004, UFZ Leipzig-Halle GmbH, Leipzig, S. 19-35.

OECD (2000), *Managing Urban Growth*, DT/Tdcp, Paris.

Pendall, R. (1999), "Do Land Use Controls Cause Sprawl?", *Environment and Planning*, B 26 (4), 555-571.

Pumain, D. (2003), *Urban Sprawl: Is there a French Case?* In: Richardson, H.W.; BAE, C.C. (Eds.): *Urban Sprawl in Western Europe and the United States*. London: Ashgate, S.137–157.

Putnam, R.D. (1995), "Bowling Alone: America's Declining Social Capital." In: *Journal of Democracy* 6:1, Jan 1995, 65-78.

Real Estate Research Corporation, (1974), "The Costs of Sprawl: Environmental and Economic Costs of Alternative Residential Development Patterns at the Urban Fringe", Real Estate Research Corporation (US Environmental Protection Agency, Washington, DC)

Siedentop, S. (2005), *Urban Sprawl – verstehen, messen, steuern*. DISP 160. Zürich, 23-35.

Sierra Club (1998), "Sprawl: The dark side of the American Dream".

Small, K. (2000), "Urban Sprawl: A Non-Diagnosos of Real Problems" In: *Metropolitan Development Patterns: 2000 Annual Roundtable*, Cambridge MA.

Sudhira, H.S. et al. (2003), "Urban sprawl pattern recognition and modelling using GIS", Indian Institute of Science, Bangalore.

Torrens, P. M., Alberti, M. (2000), *Measuring Sprawl*, Paper presented to the Association of Collegiate Schools of Planning (CSP), Atlanta, GA, 1-13.

Torrens, P. M., Alberti, M. (2000), *Measuring Sprawl*, CASA Paper 27, Centre for Advanced Spatial Analysis, London.

Transportation Research Board, National Research Council (1998), "The Costs of Sprawl – Revisited", National Academy Press, Washington D.C.

Travisi, C. M. & Camagni, R. (2005), 'Sustainability of Urban Sprawl: Environmental-Economic Indicators for the Analysis of Mobility Impact in Italy,' Working Papers 2005.102, Fondazione Eni Enrico Mattei

Wassmer, R. W. (2002), An economic perspective on urban sprawl. Working Paper for the California Senate Office of Research: 1-21.

Wassmer, R.W. (2005), Causes of Urban Sprawl (Decentralization) in the United States: Natural Evolution, Flight from Blight, and the Fiscalization of Land Use, Working Paper, 1-34.

Whyte, W. H. Jr. (1959), Urban Sprawl, Fortune, 103-109.

Wolman, H. Galster, G., Hanson, R., Ratcliffe, M., Furdell, K. (2002), Measuring Sprawl: Problems and Solutions, Paper prepared for presentation at the 2002 meeting of the Association of Collegiate Schools of Planning , Baltimore MD.