WPS7453

Policy Research Working Paper 7453

Jobs and Land Use within Cities

A Survey of Theory, Evidence, and Policy

Arti Grover Goswami
Somik V. Lall

WORLD BANK GROUP
Social, Urban, Rural and Resilience Global Practice Group
October 2015

Policy Research Working Paper 7453

Abstract

Over the last century, the urban spatial structure of cities has transformed dramatically from the traditional monocentric configuration to varying forms of decentralized organization. This paper reviews the theory and empirical evidence to understand the urban morphology of jobs and land use within a city. This survey highlights four broad insights: (i) The evolution of monocentric to polycentric centers has been accompanied by structural changes in the city. (ii) The internal geography of a city is an outcome of the trade-off between the pull from agglomeration economies and the push from congestion. (iii) The presence of externalities implies that the equilibrium spatial organization achieved by profit-maximizing firms may not necessarily be optimal. This justifies the role of public policy in addressing the associated market failures. (iv) The productive edge and competitiveness of a city can be enhanced by introducing policies that increase the overall connectivity to take advantage of economic opportunities across the metropolitan area. The survey also puts together a wide range of policy instruments that are useful in closing the gap between equilibrium urban spatial structure and the optimal outcome.

This paper is a product of the Social, Urban, Rural and Resilience Global Practice Group. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://econ.worldbank.org. The authors may be contacted at agrover@worldbank.org and slall1@worldbank.org.

Jobs and Land Use within Cities:
A Survey of Theory, Evidence, and Policy

Arti Grover Goswami and Somik V. Lall

JEL Classification: R12, R13, R14, R3, R52

Key words: Agglomeration, spatial, monocentric, polycentric, land use regulations.

Acknowledgements: This paper is a part of a World Bank Regional study on the Spatial Development of African Cities. We are grateful to the World Bank Multi-donor Trust Fund on Sustainable Urbanization and UK DFID for financial support. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors, they do not necessarily represent the views of the World Bank Group, its Executive Directors, or the countries they represent and should not be attributed to them. The authors can be contacted at agrover1@worldbank.org and slall1@worldbank.org.

I. Introduction

The internal geographies of cities around the world have been dramatically transformed over the last century. Traditional monocentric organization where jobs are concentrated in the Central Business District of cities has been replaced by a variety of decentralized configurations. In this paper, we survey theory and empirical evidence to understand the pattern of urban morphology of jobs and land use within a city. We focus on studying the evolution of a city's spatial structure, the extent and nature of agglomeration economies and the adjustments in functional specialization of economic centers within cities. Such an investigation is central to answering a range of economic and policy questions on land use, location of jobs, entrepreneurship and innovation, alongside congestion costs. This also has bearing on public policy interventions, such as urban development and taxation policies and transport infrastructure investments.

The theoretical and empirical literature on the spatial distribution of economic activities across cities and regions points to the existence of agglomeration economies, both within sectors (localization economies) and across sectors (urbanization economies) in developed countries and developing countries.[1] Although the distribution of economic activities and jobs within a city has been the focus of a limited set of contributions, nonetheless the literature on the internal structure of cities has made clear that depending on the relative strengths of push and pull factors determining agglomeration, a metropolitan area can take many spatial forms. This includes monocentric, polycentric, and dispersed spatial forms. Theory particularly assigns these push and pull forces in the form of the cost of commuting, population size, cost of congestion and the rate of spatial attenuation of production externalities.[2]

Evidence reveals that the forces that drive agglomeration of firms and jobs at a sub-city level may be quite different from those that drive clustering of activities across different cities. The empirical literature on urban form indicates that most large cities in developed countries have multiple centers. Just as cities prove the existence of agglomeration economies, these multiple nodes of employment within metropolitan areas strongly suggest that agglomeration economies work at different strengths and scopes within a city. The limited research that focuses on smaller spatial scale highlights a series of fundamental issues pertinent to policy makers.[3] These studies elucidate that urban configuration is shaped by the quality and coverage of transport infrastructure and the ensuing cost of commute, population size, proximity to ports, human capital and other inputs critical to production. Most importantly, the literature raises questions on the optimality of allocation of land within a city across residential, commercial and other uses. Our review of economic theory and empirical evidence on how employment clusters and economic specialization emerge suggests that in the presence of externalities, the equilibrium urban form generated by profit maximizing firms is unlikely to be the most efficient spatial structure, often creating a

[1] For surveys on works across cities, see Duranton and Puga (2004) and for that across regions, see Ottaviano and Thisse (2004). For empirical works on developed countries see Rosenthal and Strange (2004), Henderson (2005), Combes et al. (2011) and that on developing countries see Overman and Venables (2005), Henderson (2010) and Venables (2010)

[2] See for example Fujita (1988), Lucas and Rossi-Hansberg (2002), Berliant et al. (2002), Berliant and Wang (2008), Fujita and Ogawa (1982) and Chatterjee and Eyigungor (2014).

[3] See Picard and Tabuchi (2013) for a study that presents a formal model of urban spatial structure based on a microeconomic foundation of forward and backward linkages as it has been discussed in the framework of new economic geography.

wedge between the preference of businesses and broader societal welfare. In this context, we also review urban form shaping policies and infrastructure investment.

We highlight four broad insights:

1. Large cities in today's developed economies have evolved from monocentric structures to polycentric spatial forms. In the late nineteenth-century, for example, four-fifths of Chicago's jobs were compactly located within four miles of State and Maddison streets. By 1970, Chicago had 9 employment subcenters which increased to 15 in 1990 and 32 in 2000.

2. The internal geography of a city is an outcome of the trade-off between the pull from agglomeration and the push from congestion and the cost of commute. The presence of production externalities arising from agglomeration justifies the role of public policy in addressing the associated market failures.

3. Rapidly growing cities need multiple employment sub-centers to harness localized production externalities and reduce congestion. There is a role for government in assisting sub-center formation because social gains from developing new sub-centers exceed the private gains.

4. The productive edge and competitiveness of a city can be enhanced by introducing policies that increase overall citywide connectivity permitting speedier and longer commuting to take advantage of metropolitan-wide economic opportunities.

The rest of this survey is organized as follows. Section II discusses the literature on the internal structure of cities, mainly dealing with the theory and evidence on forces agglomerating firms and workers at the level of a city or finer spatial scales. This section also describes the evolution of historic cities to monocentric forms followed by the emergence of polycentricism or multiple employment centers. Although most evidence in the field relates to developed countries, we introduce examples from developing countries whenever possible. The idea is to draw lessons from both types of experiences for building sustainable cities in Africa. The analytical insights from section II highlight the importance of addressing externalities and hence call for government intervention in aligning the market equilibrium with social welfare maximizing outcomes. These policy options are discussed in section III. The survey ends with concluding comments in section IV with special emphasis on how urban spatial development policies may help accelerate urban transformation in African cities.[4]

[4] While misallocation of other inputs such as labor could be equally important, our survey focuses mainly on the impact of urban form on land use within a city.

II. The Internal Structure of Cities

The analytic underpinnings of examining the industrial development within cities originate with Harold Hotelling's linear city model where firms agglomerated in the city center to seize the largest possible market. In the 1960s and 1970s models of the monocentric city were augmented by William Alonso, Edwin Mills, and Richard Muth to include endogenous land rents and locations of city limits.[5] The historical location of all US cities can be understood through heterogeneity in land use that these models portray. The basic premise of these models is that the production structure and land use vary significantly between urban and non-urban areas and between urban areas of different size. Since it is more productive for all firms to locate on the piece of land they value the most due to proximity to highways, proximity to their markets and suppliers, the monocentric city is a natural outcome.

Economists have identified a variety of factors that lead to ''agglomeration economies". The most important and best articulated of these factors is increasing returns to scale. This leads to agglomeration, not only of the activity in question, but also of other activities vertically related to it. Increasing returns can be derived from various sources, including sharing a large indivisible good such as a public good in the form of a highway or rail network[6] or a large indivisibility in production activity.[7] These models propose that once the large fixed cost associated with the "indivisible" good is incurred, consumers only have to pay a constant marginal cost per use. [8] The consumer must, however, commute between their home and this indivisible good. This creates a trade-off between the cost of commute and the cost of crowding the land near the location around the public good and facing higher land rent. Equilibrium configuration is an outcome of such a trade-off where households and firms that value the good most locate in proximity to this public good.

Endogenous Urban Configuration[9]

Endogenous economic agglomeration can emerge under three alternative conditions.[10] First, *heterogeneity in space* could result in endogenous CBD formation due to comparative advantage among locations. The rise of Chicago, for example, is attributed to location specific production of lumber, grains, and livestock.[11] Second, imperfectly competitive markets also explain endogenous CBD formation because they generate aggregate increasing returns because of the productive advantages of sharing a wider variety of differentiated inputs. Finally, *externalities from non-market interactions* affect agglomeration

[5] Hotelling (1929), Alonso (1964), Mills (1967) and Muth (1969)

[6] They are the subject of a voluminous literature referred to as club theory (or theory of local public goods when the spatial dimension is explicitly taken into account

[7] This corresponds to the idea of a factorytown, where large fixed costs create internal increasing returns in a production activity that employs the workforce of an entire city whose size is bounded by crowding.

[8] There is in fact a long tradition of modelling cities as the outcome of large indivisibilities in production (Koopmans, 1957, Mills, 1967, Mirrlees, 1972)

[9] The study of monocentric city that assumes a predetermined city center is widely known as "New Urban Economics" (see Fujita and Ogawa, 1982).

[10] These conditions are not mutually exclusive; several models combine two or more features.

[11] See Schweizer et al. (1976) and Berliant and Konishi (2000) for theoretical foundations and Cronon (1991) for Chicago case.

in a central location. Agglomeration theories are based on the assumption that spatial clustering generates positive externalities. These externalities in turn promote greater efficiency, productivity and/or innovation. Positive externalities may result from clustering of firms within the same sector (localization economies) or may even be generated by agglomeration of firms across sectors (urbanization economies).[12] Firms gain from spatial clustering due to the ease of communication, increased knowledge sharing and spillovers, increased scale of markets, access to human capital and other inputs and outputs, and from sharing a common urban infrastructure.[13]

The presence of production externalities implies that a firm's productivity at any site is governed not only by internal factors, but also by the concentration of jobs (and firms) at neighboring sites. The central business district of a city illustrates this phenomenon in the most obvious way where firms agglomerate to take advantage of such externalities. However, agglomeration raises workers' wages and land rents and thus discourages further concentration. In contrast, decentralization of jobs increases commute time and costs for workers and draws them to the center.[14] When the cost of commuting to the city center is high, the proximity of firms and workers is mutually beneficial and therefore the city is more decentralized. This decentralization could be in the form of mixed land use and dispersed jobs all over the city or, alternatively, in the form of few concentrated sub-centers of jobs. However, when the commute costs are low, a traditional monocentric city emerges.[15] In fact, for many models, polycentricity and dispersed land use are the result of the interplay between the CBD's agglomeration diseconomies and transport costs related to transport infrastructure.

Models explaining non-monocentric urban land use were pioneered by Masahisa Fujita and Hideaki Ogawa via the introduction of production externalities.[16] Their models suggest that when commuting costs are low, a Mills-like monocentric city prevails, while a polycentric city with multiple business districts emerges when the cost of commute rises.[17] These models also suggest that the larger is the commuting

[12] Marshall (1890) argued that the clustering of specialized manufacturing firms increased knowledge sharing through processes such as shared labour markets. Such spillovers are even more significant in the service dominated knowledge economy of the 20th century. Knowledge sharing and specialized labor markets are therefore central to productivity. Henderson et al. (1995) show that agglomeration effects for mature industries are related to Marshall scale economies, while newer industries benefit from diversity akin to Jacobs (1969) economies. Externalities associated with sources such as geographic concentration of human capital which explains clustering of firms based on labor pooling is explained in Rauch (1993) and Ciccone and Peri (2006).

[13] Duranton and Puga (2004) summarize the gains from agglomeration in terms of sharing, matching, and learning effects. Sharing effects include the gains from a greater variety of inputs and industrial specialization, the common use of local indivisible goods and facilities, and the pooling of risk; matching effects correspond to improvement of either the quality or the quantity of matches between firms and workers; learning effects involve the generation,
diffusion, and accumulation of knowledge. Agglomeration economies explain the existence of cities. This is particularly important given the growing evidence about the importance of such agglomeration economies. For a more recent survey on the evidence on agglomeration economies, see Combes and Gobillon (2015). For a more detailed exposition of the implications of introducing agglomeration economies in a monocentric city model, see Duranton and Puga (2014) and Behrens and Robert-Nicoud (2015).

[14] See Ogawa and Fujita (1980a, 1980b) and Imai (1982).

[15] Many studies reveal that the degree of decentralization is strongly related with the spread of automobiles (e.g. Anas et al., 1998, Glaeser and Kahn, 2004; White, 1976; Steen 1986 and Sullivan, 1986)

[16] These include Fujita and Ogawa (1982), Ogawa and Fujita (1979, 1980a, 1980b).

[17] The Polycentric City included six main elements: a traditional business core; a secondary business core; a tertiary business core or inner-city edge city; an outer edge city; outermost edge cities; and specialized concentrations. See Fujita and Ogawa (1982), Imai (1982) and Lucas and Rossi-Hansberg (2002) for more details. Also see a more recent paper by Chatterjee and Eyigungor (2014). Duranton and Puga (2015) present an excellent exposition of these models.

cost, the greater is the area devoted to mixed use. The intuition for this result is that a higher cost of commute motivates people to live close to their places of work so as to economize on these costs. Thus, a mixed land use pattern appears.

Recent models suggest that the strength of agglomeration forces also determines the pattern of city development. The higher is the pace of spatial decay of external effects, the greater is the probability of having multiple pure business areas in a city so that firms can reap the benefits of production externalities.[18] In other words, more localized external effects motivate firms to form pure business centers in a city, while a faster pace of spatial decay of external effects promotes a polycentric city. Thus, cities transition from a monocentric to a polycentric structure and then to complete dispersion as production externality weakens.

Keeping the externality parameter constant, simulation shows that a Mills like monocentric city surrounded by residential land use appears, provided that the rate of spatial decay is within reasonable limits. Figure 1 presents results from such simulations, starting with panel A which shows the results from an extremely low commute cost and going up to panel C representing extremely high cost of commute. Panel B presents the results on urban land use from a moderately high level of cost of commute. In each of these panels, the land devoted to business use is shaded while those used for mixed or residential purposes are not. The panels in the figure clearly suggest that mixed land use begins to appear when commute costs increase to a moderate level (panel B) and in fact the entire city could be devoted to mixed land use if the cost of commute is very high (panel C). In all these figures, alternative *business* land use pattern is shown in dotted lines when the rate of spatial decay of externalities is high. A high rate of spatial decay implies a more concentrated business center that can reap full benefits from agglomeration. As the rates of spatial decay increase beyond the limits shown in figure 1, multiple and concentrated business districts appear while mixed land use disappears. This is because production externalities attenuate fairly rapidly and hence there is no room to accommodate households alongside firms. It would be suboptimal for households to occupy space alongside firms as that would diminish the benefits of production externalities. In equilibrium firms outbid the household for land space to be closer to other firms.

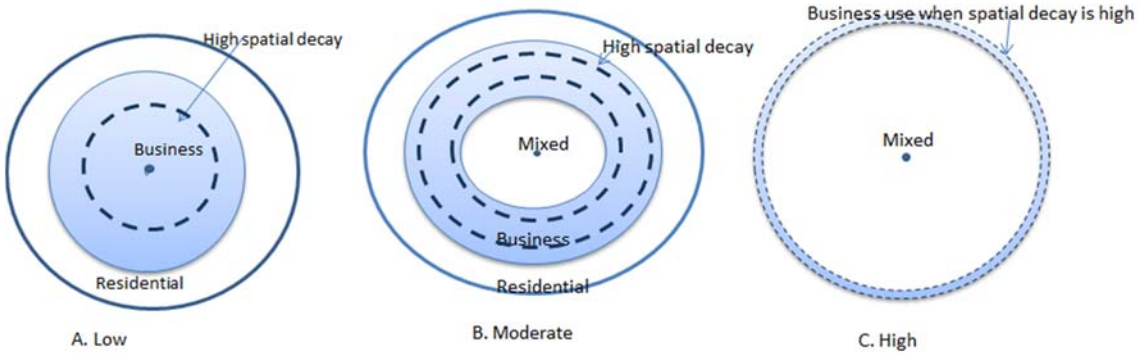

Figure 1: Land Use and the cost of commute

Source: Authors' adaptation from simulations in Lucas and Rossi-Hansberg (2002)

[18] Lucas and Rossi-Hansberg (2002) and Berliant and Wang (2008).

Although theoretical models provide the foundations for examining the equilibrium spatial configuration of polycentric cities, they are not immediately applicable to producing empirically testable comparative static predictions. Recent empirical works at smaller spatial scales show that it is very important to understand a firm's location choice at a given zip code within the same metro city. A general conclusion of these studies is that firms may agglomerate at zip code or county level for usual reasons such as input sharing and labor pooling in a given location. Evidence on the US counties suggests that agglomeration due to input sharing is only present at the state level and not at county or zip code level while the idea of labor pooling makes firms agglomerate at all levels. By contrast, knowledge spillovers significantly affect agglomeration at the zip-code level but not so much at county or state level. Theory suggests that the benefits of agglomeration diminish rapidly with distance and empirical research seems to hint that this decay function varies by the source of agglomeration economies. Particularly, evidence on the US counties reveals that the decay of knowledge spillovers with distance occurs much more rapidly vis-à-vis that of input sharing.[19]

One paper that elegantly ties theory with empirics is that by Daniel McMillen and Stefani Smith.[20] Their study derives tractable predictions from the Fujita and Ogawa (1982) model regarding the relationship between an urban area's population, commuting costs, and its number of sub-centers. Using 1990 data on large US urban areas, the authors test the predictions of Fujita and Ogawa, whose simulations suggest that the equilibrium number of sub-centers is likely to increase with population and the unit cost of commuting. Controlling for other variables affecting urban form, such as the median income in the urban area and the age of the central city, McMillen and Smith find very strong support for theory, with population and commuting costs accounting for nearly 80% of the variation in the number of sub-centers across the urban areas. [21]

[19] Studies at higher spatial level include Glaeser et al. (1992), Henderson et al. (1995), Rosenthal and Strange (2004) while a recent study by Agarwal et al. (2012) makes a descriptive attempt to understand agglomeration at sub-city level in Los Angeles. There are few examples of studies at zip code level by Rosenthal and Strange (2001, 2003); Kolko (2010); De Silva and McComb (2011). Besides, spatial decay in externalities and spillovers, Redfearn (2007b) notes that the stability in subcenter employment in Los Angeles may be associated with historical path dependence. However, path dependence cannot entirely explain the rise of subcenters in the last few decades.

[20] McMillen and Smith (2003) is an empirical application of Fujita and Ogawa (1982). The analysis also brings to light the fact that several theoretical papers have highlighted the role of commute costs (such as Lucas and Rossi-Hansberg, 2002) and population size (e.g. Fujita and Mori, 1997) in determining the urban form.

[21] Much of the recent literature actually suggests there is a lot of persistence in urban patterns. Bleakley and Lin (2012) document the consistent and persistent development of cities in the US along a fall line. Similarly, Redfearn (2009) finds that the most important predictor of clusters of de-centralized jobs in Los Angeles in 2000 in a location is the presence of decentralized jobs in the same location in 1980. Thus, history has much to determine the current state of land use in a given city. For instance, the existence of a number of subcenters for a port city may well be explained by its evolution. Starting from the initial port site, expansion of the port results from technological innovations in maritime engineering and improvements in cargo handling. This marks the beginning of changing spatial relationships between the port and the urban core, as docks are built further away from the CBD due to congestion and space limitation. In the later stages of port development, increased specialization of cargo handling, growing sizes of ships, and ever increasing demands for space for cargo-handling and storage results in port activity being concentrated at sites far removed from the oldest facilities. Such models of de-concentration in port cities have been exemplified by metropolises in the developed countries such those in Europe and the US. In developing countries, on the other hand, studies suggest an increasing level of port concentration as certain hinterland routes develop to a greater extent vis-à-vis others. The urban system evolves from an initial pattern of scattered, poorly connected ports along the coastline to a main network consisting of corridors between gateway ports and major hinterland centres. See Taaffe et al. (1963), Bird (1980), Barke

In general, theory suggests that a city's population and its physical size play a pivotal role in determining the extent of concentration and sub-center formation. Large population leads to high congestion cost that motivates firms to locate outside the CBD. For the USA as a whole, evidence confirms that jobs became increasingly concentrated; however, when considering only metropolitan areas, it appears that there has been de-concentration of jobs.[22] This suggests that the USA witnessed de-concentration across large counties, and increased concentration across smaller-sized counties, a pattern akin to the rise of sub-centers in congested cities.[23, 24]

Developing countries may appear to be different, partly reflecting their early stage of development. For instance, India is observing increased concentration in already dense clusters. Although the costs of congestion in India's mega-cities seems huge, nonetheless, these mega-cities may still benefit from relatively large agglomeration economies, vis-à-vis intermediate-sized cities that may have problems with market access, supply of better quality intermediate inputs and connectivity to basic infrastructure. By contrast, in developed countries, these issues are less of a constraint in intermediate sized cities. Although this phenomenon is not common across all emerging economies, particularly China, in general, however, it seems that constraints such as access to skilled labor[25] and telecom network are curbing the spread of urban growth in second tier cities. [26]

The factor that has received substantially large attention in explaining the decentralization of jobs and sub-center formation is the role of physical infrastructure and transportation.[27] Here, theory suggests that a decline in transport costs tends to untangle the location of business and residential areas and can lead to employment concentration of high value activities at the center.[28] Further, cheaper commute costs also make the CBD accessible to firms located in farther locations. Thus, when transport costs fall, some firms that do not need to be centrally located for their day to day operations can easily move out of CBD because the CBD would still be accessible to them even from farther locations. In sum, a decline in transport costs facilitates the movement of households from central locations and leads to concentration of business activities, while within businesses it aids the movement of non-core activities to remote locations. What do empiricists have to say on the effect of decline in transportation cost on the location of firm activity?

(1986), Hayuth (1981) for general models; see Kuby and Reid (1992), Notteboom (1997), McCalla (1999) and Hayuth (1988) for empirical studies.

[22] Carlino and Chatterjee (2001, 2002)

[23] See Desmet and Fafchamps (2006). This is broadly in agreement with models of a monocentric city with dispersed non-CBD employment, agglomeration economies lead to a small number of employment subcenters in these models (e.g. Brueckner, 1979).

[24] Glaeser and Kohlhase (2004). Other explanations for the decline of central cities in the United States have focused on a variety of social and material ills that have afflicted central cities such as crime (Cullen and Levitt, 1999), the degradation of the housing stock (Brueckner and Rosenthal, 2009), racial preferences (Boustan, 2010), and related changes in the school system (Baum-Snow and Lutz, 2011)

[25] Agglomeration due to human capital and knowledge spillovers is motivated in, for example, Simon (2004). In addition, the existence of human capital externalities is also evident from that fact that the proximity to educated workers is associated with a higher wage. See for example, Glaeser and Mare (2001), Rauch (1993), Moretti (2004).

[26] Desmet et al. (2015).

[27] See Anas et al. (1998), and Glaeser and Kahn (2004), for a review of this literature, and Burchfield et al. (2006) for a recent empirical study of urban sprawl in the United States.

[28] Fujita and Ogawa (1982) and Lucas and Rossi-Hansberg (2002), Glaeser and Kohlhase (2004)

Empirical work supports the theory that decentralization and suburbanization of jobs benefit non-central locations near the transport infrastructure. Evidence supports that suburbanization was encouraged by a well-developed railroad system and/or the freeway network in the US, Spain and Japan.[29] Evidence on developing countries also suggests a somewhat similar pattern. In the case of China, data confirm that each railroad ray displaced 25% of core city industrial GDP and 16% of 1990 core city total GDP to suburbs. For China, however, radial highways do not seem to play an important role in explaining de-concentration of production from the CBD. Highways, on the other hand, seem to matter for India where the movement of formal sector manufacturing from urban regions to rural parts of a district is explained by the Golden Quadrilateral highway. A similar case is observed in Indonesia where the process of decentralization of industrial production from the urban CBD has been facilitated by the building of a highway linking the city to nearby hinterlands.[30]

Although there are no studies on the impact of improvements in transport infrastructure on the spatial organization of economic activity within African cities, limited available evidence available on cross-city work in Africa suggests that there could be considerable potential for an expansion of transit networks. For instance, in Zambia, only 5% of all trips are longer than 10 km, while for the U.S. 5% of trips are longer than 50km, implying that people in African cities are unable to take advantage of a full range of job opportunities.[31] Further, cross-city evidence in Africa indicates that periphery cities with better road connection to port cities grew faster vis-à-vis cities with poor road connection, suggesting that an efficient transport network in Africa is likely to enhance city productivity in the same way as in most cities around the world.[32]

Similar to improvements in physical infrastructure, advances in communication and computing technologies have facilitated the fragmentation of tasks and activities of firms across various sites. Such separation of tasks is expected to occur when the cost of disintegration and the possible loss of agglomeration benefits is more than offset by lower production costs at a cheaper location. Innovations in electronic infrastructure largely explain the migration of back-office jobs to suburbs of a city or to smaller cities while the management or headquarter activity that is a personal or face to face communication intensive task is being retained in the CBD.[33]

[29] For the impact of roads and highways within cities, see, Baum-Snow (2007), Duranton & Turner (2011) for US; Hsu and Zhang (2014) for Japan; Garcia-López et al. (2013) and Garcia-López (2010b) for Spain.

[30] See Baum-Snow et al. (2012), Roberts et al. (2012) and Deng et al. (2008) for China; Ghani et al. (2012, 2014) for India; Henderson and Kuncoro (1996) for Indonesia. Highways in China tend to explain the decentralization in population, Baum-Snow et al. find that with each increment of radial road capacity causing a 3 to 5.5 percent decline in central city population from 1990 to 2010. This finding corroborates with Baum-Snow (2007) results for the US, though the magnitudes are much larger, given the higher income and incidence of auto commuting in the country. In addition, Baum-Snow (2014) finds that the new radial roads on US inter-state highways had a much larger effect on residential decentralization vis-à-vis employment decentralization. For more on the effects of railroads in China, see Banerjee et al. (2012), Baum-Snow et al. (2012) and for India, see, Donaldson (2014).

[31] Schafer (2000).

[32] See Storeygard (2012)

[33] Duranton and Puga (2005), Ota and Fujita (1993), RossiHansberg et al. (2009), Anas et al. (1998)

Finally, research has also unveiled that urban form in several developing country cities has been critically shaped by regulations on land use. [34] For instance, spatial distortion in Moscow is explained by Marxist ideology while a morbid cult of design explains the urban morphology in Brasilia and the policy relating to apartheid or land use regulations discouraging the division of land parcels explains the spatial allocation of land in Johannesburg. All these cities do not fit the "standard" urban model but density gradient in these cities rises as one moves away from the city center. Nonetheless, evidence is almost positive that regulation is only a transmission mechanism and that changing regulations over time can change the form of the city.

Evolution of Urban Spatial Structure: From Monocentricity to Polycentricism and More

In an attempt to understand the internal structure of cities, our review of theory reveals that monocentric cities are untenable. Studies suggest that decentralization is normal and the declining role of central business districts is a fact of life in modern cities (e.g. Mills, 1972). In Alain Bertaud's words, the relationship and evolvement of monocentric and polycentric cities is described as follows:

"*As they [cities] grow in size, the original monocentric structure of large metropolises tends with time to dissolve progressively into a polycentric structure. The CBD loses its primacy, and clusters of activities generating trips are spreading within the built-up area. Large cities are not born polycentric; they may evolve in that direction. Monocentric and polycentric cities are animals from the same specie observed at a different time during their evolutionary process"*.[35]

The evolution of a monocentric city can be traced back to the history of development of transportation routes. Prior to the 1840s, most cities were tied to waterways such as harbors, rivers, and canals or railway networks. Locating closer to these terminals created accessibility advantages and favored the growth of a single manufacturing district. The high cost of within city communication motivated concentration of manufacturing within the CBD of cities such as New York.[36] Similarly, in the late nineteenth-century, four-fifths of the Chicago's jobs were compactly located within four miles of State and Maddison streets.[37] The advent of horse carriage and electric street cars at the end of the 19th century gave rise to a spatial structure consisting of a CBD surrounded by residences concentrated around mass transport spokes. The invention of the internal combustion engine, telephone and assembly line production in the early 20th century eased transport and communication and increased the attractiveness of locations where land was cheap.

After the World War II, large scale construction of inter-state highways and the creation of suburban rail terminals reduced the cost of trucking and travel. These changes weakened the necessity of locating near transport networks and enabled manufacturing to move out from the CBD to the suburbs. Manufacturing

[34] Bertaud and Malpezzi (2014)
[35] Bertaud (2003)
[36] Chinitz (1960)
[37] Fales and Moses (1972)

in CBD was rapidly replaced by service and office centers. Edge cities or polycentric spatial structures thus evolved in North America as a result of developments in transportation and communications technology.[38]

Cities such as Chicago, New York and Los Angeles have evolved from monocentric structures to polycentric spatial forms while smaller cities such as Milwaukee in the mid-west or Buffalo in upstate New York remain largely monocentric. Chicago started as a monocentric city way back in 1850s but had 9 sub-centers in 1970. The number of sub-centers in Chicago increased to 13 in 1980, 15 in 1990 and 32 in 2000 while more than 30 sub-centers were identified in New York and Los Angeles in a 2003 study.[39] Although there are not many studies that identify sub-centers in a developing country city, some that do present mixed results. For instance, Mexico City was found to have 35 sub-centers mainly in or near the core of the metropolis[40] while in the case Lima (in Peru) 10 centers of employment concentration were identified in 2008.[41]

Table 1 documents the evidence of declining concentration of jobs in five mid-western cities. The table suggests that across all these cities, 35 to 52 percent of suburban residents worked in the central city in 1960, whereas only about 6 to 17 percent of city residents worked in the suburbs. However, by 1990, the percent of suburban residents working in the central city approximately equaled the percent of city residents working in the suburbs. Thus, table 1 clearly shows that the CBD is not the dominant employment site in any of these cities and that other centers of employment besides the CBD have emerged in the last three decades.[42]

	City residents working in the suburbs				Suburban residents Working in the cities			
	1960	1970	1980	1990	1960	1970	1980	1990
Chicago	6.6	16.1	18.4	22.5	34.6	27.1	22.5	25.6
Cleveland	7.7	24.4	28.6	30.3	52.4	43.5	34.8	32.5
Detroit	17.3	32.1	34.3	36.4	33.5	24.6	16.9	19.4
Minneapolis-St. Paul	6.6	19.7	24.5	29.8	52.1	43.5	31.2	30.5
St. Louis	8.3	21.1	24	35.9	36.7	30	25.4	27.9

Table 1: Employment de-concentration in Cities, 1960-1990

Source: McMillen (2001b)

There is extensive evidence today to support the idea that polycentrism is the reality of urban settlements in many countries.[43] Urban empiricist, however, started off with an estimation of a monocentric city, primarily focusing on the prediction that gradients of land and housing prices, population density, and intensity of construction decline exponentially from the city center. A comparison of urban density

[38] Being much older, European cities have evolved somewhat differently. Central parts of many of such cities are allocated to mixed use. Nevertheless, these cities have also witnessed massive suburbanization and the emergence of edge cities.

[39] McMillen and Lester (2003); McMillen and Smith (2003)

[40] To identify these subcenters, Aguilar and Alvarado's (2004) used metropolitan-wide data from the Mexican 1999 Economic Census, broken down to the so-called AGEB-level (Areas Geo-Económicas Básicas) which are administrative units comparable to the U.S. census tracts and applied a minimum of 5,000 localised jobs as cut-off point.

[41] Gonzales et al. (2012)

[42] For the 100 largest metropolitan areas in the US, Baum-Snow (2014) also finds that the employment share of central cities declined from 61% in 1960 to 34% in 2000.

[43] See Cervero and Wu (1997), Giuliano and Small (1991), McDonald and Prather (1994), McMillen and Smith (2003) and McMillen and McDonald (1998). Particularly, in the case of US, Glaeser and Kahn (2001) find that in 1996 only about 25% employees in US metropolitan areas worked within 5 km of their CBD.

gradients across developed and developing countries suggests that (i) density functions tend to be flatter in countries with higher incomes.[44] For example, Japan and the Republic of Korea had similar population densities in 1970s, yet the average 1970 gradient for Korean urban areas was 0.670, whereas the average for Japanese urban areas was 0.391. (ii) Although the tendency for gradients to decline as time passes is present in both groups, nonetheless it is more pervasive and persistent in the case of developed countries. For instance, the average density gradient for Japanese urban areas declined to 0.46 in 1965 to 0.39 in 1970 while for urban areas in Korea it declined from 0.70 to 0.67 during the same period. In the case cities in Ghana, this gradient did not decline at all over the mentioned period. In fact, in developing countries, there are also some exceptions to this downward trend in density gradients, such as the smaller urban areas in Mexico and in India during the early years of the post-World War II period.[45, 46]

The standard monocentric model predicts that land prices fall with distance from the CBD. This has been found to be true, for instance, in the case of New York starting from 1835 when the land price gradient was -0.65 implying that land values declined by 65% per mile from the CBD of New York. The fit of this semi-log standard Mill's type monocentric urban land values model was 0.71. Contrarily, by 1900, the land price gradient had declined to -0.14 and the fit of the monocentric model had fallen to 0.06.[47] In the case of Chicago, evidence on land values reveals a similar story suggesting that the explanatory power of the standard monocentric model has declined over time.[48] In general examples from cities in the developed world suggest that the shape of cities evolves over time where a monocentric shape is replaced over time by a polycentric form with multiple sub centers. In the case of developing countries, this transition could be slower or even perverse as observed in the case of Moscow, Brasilia and Johannesburg.[49]

Although the monocentric model is very useful, it is difficult to apply it to suburban regions of cities such as New York, Chicago and San Francisco with multiple nodes of employment and where distance from CBD does not monotonically determine employment density and land values.[50]

[44] See Mills and Tan (1980) for earlier evidence presented here and also see Bertaud and Malpezzi (2014) for more recent analysis. The latter study suggests that cities such Shanghai, Guangzhou, Singapore, Toulouse show the steepest slope (negative) of the density gradient while at the other extreme are cities such as Mumbai, Johannesburg, St. Petersburg and Moscow that show a positive gradient.

[45] Spatial development of African cities in Ghana also witnessed a similar evolution of urban form as those in other developing countries (see Asabere and Owusu-Banahene, 1983). Work on 5 cities in Ghana reveals that although the average density gradients in these cities are relatively higher than those found for western and Latin American cities, they are nonetheless lower than those found for Korean and Indian cities.

[46] Clark (1951) was perhaps the first to show negative association between population density and distance from CBD. However, a high R-square in such models is not necessarily supportive of the monocentric model because it results mainly due to the inherent smoothing associated with the ring approach. In fact, the R-square associated with the regression of density in small areas within a city, such as tracts, on their distance to the CBD is much lower because there indeed are areas of fairly high density that are located relatively far from the main centre (see Duranton and Puga, 2014, 2015)

[47] Atack and Margo (1998)

[48] Mills (1969)

[49] Bertaud and Malpezzi (2003, 2014)

[50] Duranton and Puga (2015) describe that the monocentric city model explains patterns of residential land use and commuting within a city but not why individuals wish to be in a city to begin with. Neither does this model explain the complex pattern of division of land between residential and commercial uses within a city. Nonetheless, Bertaud and Malpezzi (2003) argue that the negative-exponential parameterization of the monocentric model can hold so long as there is "a higher density of employment at the center than elsewhere".

Empirically, the definition of a sub-center is often complex, case dependent and ambiguous because suburban employment tends to be highly dispersed. The literature has applied two criteria to identify sub-centers. One, a sub-center is defined as an area with significantly higher employment densities than surrounding areas. Two, a sub-center should be large enough to have a significant effect on the overall spatial structure of the urban area, leading to local rises in population density, land prices, and perhaps housing prices. Early urban empiricists applied these definitions and extended the usual empirical expression of a monocentric city to its multi-nodal counterpart but not without drawbacks.[51] Alternative empirical methods for capturing the internal distribution of economic activity in a polycentric setting include clustering models and nonparametric models.[52] Clustering method bunches together adjacent tracts to build sub-centers based on ad hoc definitions of employment density cut-offs and total job thresholds. For example, a cut-off that has been used in the literature defines two types of centers: "10–10" and "20–20." The former is an agglomeration of contiguous tracts having a minimum job density of 10 jobs per acre combined with a total number of jobs of at least 10,000 while the latter is a tract with at least 20 jobs per acre and 20,000 total jobs in each sub-center.[53] These models are criticized for the arbitrary nature of the density and total employment cutoffs implying that the number of sub-centers can be increased by decreasing the threshold and vice versa.[54]

The technique proposed in the above research have been applied to identify sub-centers in a number of large US cities, including Atlanta, Baltimore-Washington, Boston, Dallas, Houston, Indianapolis, Los

[51] E.g. the model requires a priori specification of the candidate subcenters implying that the gradient from each center contributes to employment density at every point in the metropolitan area. This complicates estimation especially if there are a larger number of subcenters. (See, Gordon et al.,1989 and Gordon and Richardson,1996).

[52] The most critical challenge associated with studying urban morphology is the difficulty in obtaining data at fine spatial scales. Cities chosen to study this model have based on the ease of collecting this data and these are notably the ones with strong priors on their urban forms. For instance, not many cities have population and employment data available at the level of a tract. Land registry or any other source that is able to identify parcels is far less available. Thus, issues such as cross-city heterogeneity within a country or the behavior of density gradients over time have been largely ignored (See Combes et at., 2012 for a comparison of land price gradients across cities in France and McMillen, 1996 for the dynamics of price gradient over 150 years in Chicago). Secondly, there have been measurement issues relating to not just identifying subcenters but also when making predictions for land parcels based on area density gradient. Thirdly, what we want to ask the literature is not whether cities are monocentric or not, a hypothesis that can be rejected in nearly all cities, but exactly how monocentric is a city. The key difficulty with this analysis is that the R2 of gradient regressions do not provide a good metric by which to measure this and reasonable methods in this field are yet to be developed (Duranton and Puga, 2015). Similarly, the predictions of the monocentric model on various aggregate need not be taken literally, but generally, empirical methods must be developed to measure how different is reality from such predictions. Monocentric model makes a number of predictions about some aggregate quantities and land rent at the centre (e.g. proportionality between total land rent and total commuting costs). One study that makes such an assessment is by Combes et al. (2012) who examine the proposition that land rent at the centre is proportional to city population.

[53] Giuliano and Small (1991), Small and Song (1994), Bogart and Ferry (1999), Anderson and Bogart (2001), Giuliano et al. (2007).

[54] Daniel McMillen along with his co-authors has sought to address the misspecification of monocentric models and the lack of statistical tests for the clustering methods through nonparametric regression-based models (McMillen, 2001a; McMillen and Smith, 2003; McMillen and McDonald, 1997, 1998). These models look at the deviations from general spatial trends in employment density and identify outliers relative to a spatial average that uses half of all the observations. This, however, substantially obscures the topography of local employment. Recently, methods have been devised such that the estimation of nonparametric employment density surface occurs locally— with a subsample that is kept sufficiently small to keep intact the structure of local employment density. This is necessary in order to identify subcenters, as the general spatial trends in employment density using a larger window size largely mask independent local concentrations of employment. Local maxima on the density surface become candidate centers (Redfearn, 2007a).

Angeles, New Orleans, New York, Philadelphia, the San Francisco Bay Area, and St. Louis.[55] Depending on the exact methodology used, the exact number of sub-centers varies but, in general all studies agree on multiple sub-centers for a relatively large city.[56] For instance, the city of Los Angeles is found to have 41 sub-centers when the more flexible non parametric technique of relying on the local maxima of each candidate sub-center is used while there are 48 sub-centers using the "10–10" definition and 10 sub-centers using the "20–20" thresholds for density and total employment in the clustering method.[57] In contrast, some small cities such as Austin in Texas, Buffalo in New York and Nashville in Tennessee do not have any sub-centers at all.[58] Similar evidence exists for cities in other developed countries such as Barcelona or other metro cities in Spain as well as Paris in France. Even in the case of a developing country, large cities such as Bogota (Colombia), has evolved to generate 10 sub-centers.[59]

Besides the monocentric and polycentric models, recent advances in urban economics propose three other theoretical forms of spatial organization of cities. (i) Maximum Disorder model, where workers' homes and their jobs are randomly distributed in the city. (ii) The Constrained Dispersal model, where a small number of sub-centers exist outside the CBD and attract workplaces to each other or to shared public infrastructure and amenities. In this form, both workers and firms adjust their locations to remain within a tolerable commuting range of each other. (iii) The Mosaic of Live-Work Communities, where workers and firms are all within walking or bicycling distance of each other. Of the 40 American cities examined in a recent study, evidence suggests that an average of 25% of jobs is located in sub-centers and CBD while the remaining 75% are scattered throughout metropolitan areas. This suggests that while the polycentric model dominates over the monocentric one, it is the constrained dispersal model that appropriately describes the urban spatial structure among American cities.[60]

Structural Changes with Urban Spatial Evolution: From Manufacturing to Tradable Services

As urban spatial structures have evolved from a monocentric organization to that of multiple centers of economic activity, so has the structural composition of the CBD. Since the World War II, manufacturing activity is de-concentrating while services are increasingly agglomerating in central locations.[61] Research

[55] Anderson and Bogart (2001) identify subcenters in Cleveland (9 sub centers), Indianapolis (11), Portland (11) and St. Louis (10); Bogart and Ferry (1999) also study Cleveland and found 9 subcenters; Cervero and Wu (1997, 1998) analyze the San Francisco Bay Area; Giuliano and Small (1991) identify 35 subcenters, Redfearn (2007a) study Los Angeles and found 41 subcenters in 2000; and McMillen and McDonald (1998) identify 15 subcenters in the Chicago metropolitan area. McMillen (2001a) identifies 13 subcenters in Chicago, and reasonable subcenters in Dallas, Houston, Los Angeles, New Orleans, and San Francisco. This analysis is extended in McMillen and Smith (2003) to cities such as Atlanta (4 subcenters), Baltimore (5) Washington DC (10), Boston (11), New York (38), and Philadelphia (4). Others have focused on identifying points of peak density and estimating density gradients to formally test for polycentricity, for example, Small and Song (1994) for Los Angeles and Craig and Ng (2001) for Houston.

[56] For example, using non-parametric technique, Redfearn (2007a) finds 7 subcenters in orange county of Los Angeles city. Although the number of subcenters in Los Angeles county of the city is 28, orange county itself is also polycentric.

[57] See Redfearn (2007a) and Giuliano et al. (2007)

[58] McMillen and Smith (2003)

[59] See Garcia-López (2010a, 2010b), Holl and Viladecans-Marsal (2011) Duarte et al. (2011), Mu˜niz et al. (2008) for cities (particularly Barcelona) in Spain; Gilli (2005) for Paris in France; Avendaño and Enríquez (2012) for Bogota in Colombia;

[60] See Angel and Blei (2015a)

[61] Dumais et al. (2002), Kim (1995). Kolko (1999) finds that there has been a growing concentration of manufacturing in small and medium-sized cities and of business services in larger cities.

on the US suggests that most services, particularly, wholesale retail, finance, insurance and real estate (FIRE) and other services have become more concentrated while most other sectors, such as 'manufacturing' and 'farming', have exhibited de-concentration and are located on lower cost urban periphery or in small cities and towns, perhaps to take advantage of lower rents and wages. Evidence suggests that nearly all agglomerated services industries are highly urbanized, whereas many manufacturing industries agglomerate in smaller cities or rural areas. Studies on US and French services firms suggest that service industries are localized at very short distances and are primarily situated in the heart of a few big cities and strongly benefit from highly localized technological spillovers.[62]

What explains this sort of structural change in urban form where central locations are increasingly being replaced by services rather than manufacturing firms? Some services require close proximity with consumers (e.g. professional services) while others can be transmitted over the internet. Services that rely on ICT can substitute for proximity, and are less likely to agglomerate even in the same county. Services and manufacturing agglomerate for different reasons: while manufacturing agglomerates for reasons related to labor pooling due to occupational specialization, services agglomerate because of knowledge spillovers. The important difference in explaining urbanization across the sectors, however, is the share of output going to consumers, which is positively correlated with urbanization for manufacturing industries and negatively correlated with urbanization for services industries. Thus, the cost of transporting goods to customers could affect location decisions differently for manufacturing and for services. [63] Furthermore, innovations in communication technologies fragments a firm's activities – when the front-unit activity needs to be in touch with other firms' headquarters, it will agglomerate in the CBD, while the back-unit activity will be located in the suburbs.[64,65] The shift from sectoral to functional specialization also implies that service focused business centers are few and large whereas manufacturing centers are more numerous and smaller in size. Such structural changes are not unique to the US; German and Chinese data also reveal a similar pattern of spatial development of activities.[66,67]

Other reasons explaining such structural changes include the maturity of the industry. In a mature industry the localization economies have been mostly exploited, and knowledge spillovers have lost much of their

[62] See Glaeser and Kohlhase (2004), Desmet and Fafchamps (2006), Arzaghi and Henderson (2008) and Desmet and Rossi-Hansberg (2014) for the case of US and Barlet et al. (2013) for the study on France. The theoretical argument for this observation is explained in Desmet and Ross-Hansberg (2009, 2014).

[63] See Glaeser and Kohlhase (2004) and Kolko (2010) for more details on the effect of these factors as well as the differences in the two sector with respect to co-agglomeration forces.

[64] Ota and Fujita (1993), Duranton and Puga (2005). Henderson and Ono (2008) provide a very comprehensive evidence related to the spatial separation of firm headquarters and production facilities.

[65] See Duranton and Puga (2005). Using data from micro data set on auxiliary establishments in the US from 1977 to 1997, Davis and Henderson (2008) find that headquarters of firms agglomerate in the CBD of large metro areas to take advantage of services and production externalities (which decline with distance) in the form of the availability of differentiated local service input suppliers and the scale of other headquarter activity nearby.

[66] We know of only one study, that is, by Rossi-Hansberg et al. (2009), that examines the within city sectoral pattern of urban development. They document an increase in the share of employment, and establishments in the suburbs or city edges. They find that the shift in employment from CBD to suburbs in the case of US was not driven by any one particular industry. During 1980-1990, the average employment share at the center declined from 0.42 to 0.38 in manufacturing and from 0.47 to 0.43 in services. That is, average employment shares declined by about the same percentage in both sectors. This information is, however, not indicative of whether the share of services relative to manufacturing increased at the city center.

[67] Example of such accounts on San Francisco bay area is found in Walker (2004) and on Pittsburg is described in Muller (2001).

importance.[68] This makes agglomeration in such an industry expensive if land is a critical factor of production. Hence, as spillovers weaken, mature industries such as those in the manufacturing sector tend to move out of CBD. Contrarily, services industry has relatively recently been affected by a major GPT in the form of ICT and many firms in the industry have not fully adapted to the new technological paradigm.[69, 70]

A recent study by Jed Kolko suggests that services do not agglomerate as much at a county level vis-à-vis manufacturing but more so at a zip code level (with zip codes being much smaller than counties). This trend is in contrast to that of manufacturing firms that tend to agglomerate in the same county to save on transport costs and also access to similar resources (natural or labor) but they do not need to co-locate within the same the zip code (and thus economize on land rents).[71]

It is noteworthy that the observed trends in the developed world may not necessarily be mirrored in a developing country context. In fact, there are similarities between the current developing country trends and that of the U.S.A. in the late 19th and early 20th centuries. Manufacturing facilities in developing countries often locate in large cities, perhaps because learning and adaptation are critical to the successful transfer of technology from abroad. However, as the industry matures and technology standardizes, higher land rents are likely to push the firms out to peripheral locations. Such a process is in the beginnings in China and Indonesia. However, in India, both services, and to a certain extent manufacturing, are increasingly concentrating in high-density clusters.[72] Based on the theoretical explanations, this suggests that services in India are “young", whereas manufacturing is not as “mature" as in the developed world. This should not be surprising though. Nonetheless, there are differences even across manufacturing. Research shows that organized manufacturing is moving to rural areas, away from city centers while the reverse is true for unorganized manufacturing. Thus, it seems that unorganized sector seems to be driving the overall results for Indian manufacturing.[73]

III. Is the Market Equilibrium Optimal? Some Policy Options

In the presence of externalities, equilibrium urban form generated by profit maximizing firms and utility maximizing workers may not be the most efficient spatial structure. For instance, in a polycentric environment, where urban sub-centers are formed from the tension between centripetal and centrifugal forces in the form of production externalities versus congestion or commute costs, there are distinct

[68] Duranton and Puga (2001)
[69] Hobijn and Jovanovic (2001)
[70] See Desmet and Ross-Hansberg (2009). The increased geographic concentration of services has also been studied by, Glaeser and Ponzetto (2010), who suggest that this phenomenon may be driven by lower transport and communication costs.
[71] Specifically, the average Ellison-Glaeser county level agglomeration index for services industries at the six- digit North American Industry Classification System (NAICS) level is 0.0068, while that for manufacturing is 0.0132.
[72] Desmet et al. (2015). India's experience is not common to all fast-growing developing economies. The spatial growth pattern of China looks more similar to that in the U.S. than to that of India. See Baum-Snow et al (2012) for China and Rothenberg (2013) for Indonesia.
[73]See Ghani et al. (2012). One of the factors that could be leading to de-concentration of Indian manufacturing includes delicensing that took place in the 1990s (Fernandes and Sharma, 2012).

challenges. Although production externalities encourage clustering but if private incentives to join an agglomeration is inadequate, an equilibrium with excessive dispersion may emerge. On the other hand, if production externalities are very strong, then excessive concentration leads to excess congestion. Since different externalities operate at different scales, the spatial pattern of economic activity may turn out to be too concentrated in the CBD while too decentralized at the sub-center level.

In sum, the presence of externalities creates a wedge between the market equilibrium outcome and the socially optimal one. This section discusses the policy issues in the urban spatial development and the related instruments that can help bring the market equilibrium closer to an optimal outcome.[74] Annex table A.1 presents a summary of the research discussed in this section.

The few models that compare the equilibrium configuration with an optimal one[75,76] contend that business land and employment should be more concentrated in the optimal allocation relative to their equilibrium structure. In the optimum allocation, mixed areas tend to disappear because workers' location decision also takes into account the effect of employing an extra worker on the production of other firms in the city. This, in turn, changes the gains from using a given location for business purposes. Thus, for no value of cost of employing an extra worker in a given location are the conditions necessary for the existence of a mixed area satisfied.

There are two policy issues for cities in developing countries in the current context: One, how can cities transform into productive hubs by integrating labor markets with productive firms and channelizing resources into tradable activities? Two, not all density profiles are compatible with an efficient transit system. Specifically, when residential densities are dispersed throughout the city, there is a role for government in creating conditions for the mass public transit system to work efficiently. This section discusses some of the available policy options, such as zoning, urban growth boundary, regulations on density (e.g. floor area ratio) and so on for aligning the equilibrium allocation of land use with the optimal outcome.

One of the most important policies known to influence land use corresponds to zoning laws whereby a central planner can assign land use in different sections of the city. For example, some areas can be

[74] By optimal outcome we mean the one that maximizes welfare.

[75] Early models that studied the optimal allocation of workers within a monocentric city include: Borukhov and Hochman (1977), Dixit (1973) and Hochman and Pines (1982) while Stull (1974) and Helpman and Pines (1977) obtain an endogenous size of business district. These models exogenously impose the land use structure and assume that workers commute to the CBD rather than their actual work places. This is sort of a restrictive condition because the problem of firm and household location cannot be determined endogenously.

[76] Rossi-Hansberg (2004); this section mainly discusses the results of this model. More recently, Kyriakopoulou and Xepapadeas (2013) present a model incorporating not only agglomeration economies but also pollution externality. Their results are throughout consistent with Rossi-Hansberg (2004). Arnott et al. (2008) build a model with pollution externalities where the spatial layout results from a balance between two opposing forces: one is the repulsion of households from industry that is the source of the pollution and the other is the attraction between households and industry caused by increasing-with-distance commuting costs. Other papers that compare market outcomes with socially optimal outcomes on the basis of land use mixing include - Wheaton (2004) and Zhang and Kockelman (2014). The former study suggests that agglomeration in CBD is associated with longer commute and worsened congestion, while employment decentralization entails shorter commuting distance and lower congestion levels while unlike Rossi-Hansberg (2004), the latter study models congestion externalities as well.

designated as purely business or entirely residential areas while others can be assigned for mixed land use. Although zoning is inevitable for providing the essential public facilities such as roads, public buildings and so on, it is also implemented to mitigate the distortive effects of externalities generated by the proximity of "incompatible" uses like a residence and a noxious factory, the need to contain urban sprawling via green belts as well as curbing congestion in central areas of the city. In the early half of the 20th Century, zoning was perceived to be a first-best policy instrument that could mitigate the effects of distortive externalities.[77] The problem with such zoning restrictions, however, is that it does not internalize the externality, and hence is not optimal. Nonetheless, it could still take advantage of higher external effects caused by proximity of firms to each other which would push the overall output of the city.[78]

Another policy option available to governments is the use of urban growth boundary (UGB), which is a regional boundary circumscribing an urban area, fixed with the objective of controlling urban sprawl and land use. For instance, the area inside a UGB can be mandated for use of higher density urban development, while the area outside could be reserved for lower density development. UGB could be an effective guide to zoning and land use legislation for controlling the extent of mixed land use among cities in developing countries. Some researchers argue that regulations on urban growth boundary (UGB) and building size, if properly chosen, may be an effective second-best substitute for the first-best land use policy. This is because the shadow price of land is less than its market value at the region's edge in a congested and monocentric setting.[79] Nevertheless, some other studies in a monocentric framework suggest that UGBs may achieve far lower welfare levels vis-à-vis the strategies of first-best tolling.[80] By contrast, the efficiency and welfare effect of the UGB policies in a polycentric framework is found to be mixed. UGBs seem to reduce travel costs and distances, and city size, but without significantly changing equilibrium job levels.[81]

A series of recent works in urban economics explores density regulations, such as lot size zoning as well as floor area ratio (FAR) regulation to control urban sprawl and the population density over space. Floor area ratio refers to the ratio of the total floor area of the building relative to the aggregate size of the plot on which they are built. In contrast, lot size zoning fixes the minimum or the maximum size of the lot in a given part of the city. Lot size zoning is imposed in the peripheral area of the city to control the size of each lot containing one household, whereas FAR regulation is imposed in the urban area to control the

[77] Pines and Sadka (1985); Wheaton (1998)

[78] The problem with the equilibrium with zoning restrictions is that land rents are not be continuous at the boundary between business and residential sectors which motivate some land owners to lobby for changes in the zoning restrictions. Zoning has the advantages that it entails relatively low transactions costs and is well-established. But, it is inflexible and unresponsive to market signals, and the private gains from rezoning invite rent-seeking behavior and corruption. Alternatives include pricing policies (such as differential property tax rates or impact fees on developers), government creation of previously absent markets (such as that for transferable development rights), or the redefinition of property rights (such as the use of restrictive covenants in Houston).

[79] Arnott, (1979); Pines and Sadka (1985)

[80] Brueckner (2007); Kono et al. (2012)

[81] See Zhang and Kockelman (2014). For similar results on US cities see Anas and Rhee (2007); Anas and Pines (2008); Staley et al. (1999); Cho et al. (2008).

size of buildings containing many households, such as apartment buildings, condominiums, etc.[82] While lot size zoning can be a first-best policy as it directly impacts population density, FAR regulation is a second-best policy because it only controls the total floor supply of a building without controlling the per-capita floor space consumption.[83] To manage population density at the desired amount, an extremely low FAR may have to be implemented, which in turn generates a deadweight loss in the building size market[84] When implementing FAR regulations, policy makers could also take advantage of minimum FAR (as opposed to maximum FAR, that is widely used) to increase population density in the context of urban expansion.[85] Finally, the concept of dynamic FAR would be more relevant in a world where cities have a high population growth rate because the level of externality at a location can change remarkably over time.

Regulation on building height is another formulation of FAR that aids in restricting population and employment density in urban areas. For instance, FAR in cities such as New York; Hong Kong SAR, China; and Singapore ranges from 5 to 15 in CBD to 0.5 or below in suburbs and is dynamically increasing over the years as per changes in population and infrastructure network. At the other extreme is the city of Mumbai in India where population density is inefficiently regulated by imposing a maximum FAR of close to 1 since 1991, without much consideration to geographic locations or dynamic changes.[86] Studies on the efficiency of regulations on building size reveal that the welfare cost from building height restrictions could be non-trivial and one study estimated this cost to be about 2% of the household income in a monocentric city model. Welfare costs emerge from the additional commute costs incurred by the residents living at the edge of the city.[87]

Alain Bertaud suggests a host of policy options ranging from general land use regulations and planning including those that re-evaluate the status of buffer zone between formal and informal settlements, the ones that allow densification of high income area or the policy instruments that offer to develop low income housing based on clusters rather than on individual plots. Policies that allow and actively promote the concentration of employment and calls for investment to upgrade and increase the amenities in the central parts of cities would also go a long way in reducing the extent of employment dispersion throughout the city.

Although some cities such as, Brasilia offer subsidized low cost housing programs to ease congestion in the city center, they are often unmindful of the trade-offs that poor households face. Households prefer to locate closer to their work place while providers mandate higher infrastructure standards at distant

[82] Lot size zoning can control only the plot size, not the building size. Therefore, in an area that mainly consists of buildings that accommodate multiple households, lot size zoning is not as effective as FAR regulation to control population externalities.
[83] Wheaton (1998)
[84] Also see, Pines and Kono (2012)
[85] Sakashita (1995); Kono et al. (2010); Joshi and Kono (2009). Some cities such as Oregon City, City of Buffalo, and Colorado Springs have been practicing minimum FAR regulation in designated areas to prevent underdevelopment.
[86] See Bertaud (2011).
[87] Direct density regulations such as building height restrictions (Arnott and MacKinnon, 1977; Bertaud and Brueckner, 2005; Kono et al., 2012) city limits determined in the sphere of politics (Brueckner, 2000a) constitute land-use regulations which invoke distortions that may be associated with large welfare costs.

locations where land is cheap. Thus, the government ought to make subsidies "portable" and allow low income households to make the trade-off between land use standards, transport costs and location. Similarly, emphasis on minimum floor space (or FAR) and infrastructure and not on minimum land consumption should help correct the distortion in land consumption.[88]

Land use policies can be strengthened with complementary transport policies to reach spatially optimal outcomes. A reduction in the cost of commute brings the equilibrium allocation closer to the optimal one through, (i) a direct reduction in workers costs per mile commuted and (ii) an indirect effect via the concentration of business areas. Policies to bring workers closer to job, such as those on road construction and improving public transportation should also consider this latter indirect additional gain in their cost-benefit analysis. [89]

Esteban Rossi-Hansberg's model suggests that policies on parking lots construction, highway investments which reduce the costs of working at business centers, are much needed to align the city structure with optimal allocation of land use.[90] These transport related policies are in essence similar to a government subsidy that lowers the labor costs for firms and motivates them to hire more workers has a positive impact on wages and rents.[91] In response to increases in rents, workers move out of prime locations, thereby making equilibrium allocation more concentrated and hence an optimum one.

In contemporary cities in the USA, about three quarters of the jobs are dispersed outside the sub-centers and the CBD. Studies on developed countries conclude that agglomeration economies are possibly metropolitan in scale.[92] Thus, the continued productive edge of cities is largely dependent on the ready availability of public transit and private automobiles that allow workers to reach out better and more productive economic opportunities available throughout metropolitan area. Such a model of dispersed employment may be acceptable in developed countries where car ownership rates are high and public transit systems are efficiently placed. However, for cities in low income countries, dispersion in employment is actually counter-productive because of lack of a well-organized public transit system, extremely low car ownership rates and poor road infrastructure. Developing countries may want to focus on policies that promote concentration of economic activity at the center or sub-centers and build infrastructure connectivity to and from such centers. Policies promoting metropolitan-wide connectivity and those that permit speedier and longer commuting would to aid agglomeration of firms and relocation of workers within tolerable commute range.[93]

Once zoning regulations, UGBs and transport policies can mitigate the problem of dispersed employment, agglomeration among firms is likely going to raise the productivity of the city as a whole and perhaps even attract more productive economic opportunities. At some point, congestion externalities would kick in

[88] Bertaud (2009)
[89] Rossi-Hansberg (2004)
[90] Studies that evaluate the cost of congestion and the policies for reducing such costs include - Anas and Xu (1999); Wheaton (1998). Such policies may have similar effects.
[91] We find similar suggestions in Kyriakopoulou and Xepapadeas (2013) as well.
[92] See Burger et al. (2009); Gordon and McCann (2000); Johansson and Quigley (2004) and Angel and Blei (2015a)
[93] Angel and Blei (2015b)

and necessitate the formation of multiple sub-centers of employment. Sub-center formation, however, may not be automatic. At a low population level, equilibrium and optimum coincide, however, as the population expands, the 'excessive' population must move away to form a new sub-center. Given that the social gains from relocation exceed the private gains, the second center is not formed until later. Thus, there may be a role for government in assisting sub-center formation, through the provision of infrastructure, the regulation of developers, or subsidization of developer activity or firm location.[94,95]

At the other extreme to concentrated employment sub-centers is the case of edgeless cities documented by Robert Lang.[96] An edgeless city is characterized by the development of isolated office and business space separated by large distances between each other. Such business areas are neither pedestrian friendly, nor easily accessible by public transit, and do not lend themselves to mixed use. Policies for curbing sprawl is motivated by the desire to foster agglomeration economies, reduce the cost of providing public services and infrastructure, and reduce the dependence on the automobile and the externalities associated with its use (e.g. air pollution and congestion). Development taxes and congestion tolls seem to be effective in reducing the spatial size of the city. Several anti-sprawl policies such as UGB, taxes on residential development, property taxes and fuel taxes have been proposed in the literature.[97]

Once the urban land use pattern comprises of multiple centers of economic activity and begins to look similar to the congested cities of developed countries, there is a separate set of policy instruments for dealing with congestion externalities. These include - tax on urban auto travel,[98] congestion pricing in the form of cordon charges, area-wide pricing, and variable-rate highway tolling (such as that in Singapore, London and in several of U.S. metropolitan areas).[99] Congestion pricing is expected to reduce congestion, thus lessen the associated negative externalities such as like traffic delays, air pollution, and greenhouse gas emissions etc. and also offer revenue to help fund transport system improvements, including public

[94] The same issues have been raised on a larger scale in the debates on France's poles de croissance, Britain's New Towns Policy, and LDC policies to control growth of primate cities.

[95] Developers and large firms (non-competitive) are sometimes critical in promoting subcenter formation. Edge cities", as they are popularly known, are subcenters created by large-scale land developers. See for instance, Henderson and Slade (1993), Henderson and Mitra (1996) and Fujita et al. (1997). Also see Anas et al. (1998) for a survey on subcenters from both the theoretical and empirical points of view. Takashaki (2014) finds that strategic interactions among firms could result in equilibrium with either agglomeration in the city center or decentralization where firms locate at either ends of the city. Higher commuting cost vis-à-vis shopping travel costs results in agglomeration, otherwise there is maximum dispersion in firm location decision. Another view is that employment centers emerge as a result of the decision making of local governments, including tax policy and land-use policy (Fujita, 1989; Sullivan, 1986; Zhang and Sasaki, 1997, 2000)

[96] Lang (2003)

[97] See for instance, Brueckner (2000b, 2001) who finds that development taxes and congestion tolls seem to be effective in reducing the spatial size of the city. Bento et al. (2006) find that, from an overall efficiency perspective, development taxes and UGB are equivalent instruments and the most effective anti-sprawl policies. By contrast, gasoline tax and the property tax are substantially less effective in combating sprawl. In an endogenous land use model, Anas and Rhee (2006) find that congestion tolls eliminate excess travel or excess sprawl by making the job and residence distribution of the city more compact while UGBs result in a substantial efficiency loss. Also see Anas and Rhee (2007) that makes a careful discussion on the reasons why UGB is not always a second-best policy option relative to congestion tolls.

[98] Although this may lead to excessive decentralization and perhaps over-investment in highways such as that in the US in 1950s, see, Wheaton (1978)

[99] Santos (2005); Gao (2012)

transit. Empirical work suggests that congestion tolls may catalyze land development around tolled roads, and perhaps also have a somewhat negative effect on the city center's economy.[100]

Alternatively, to reduce congestion, many innovative policies have been deployed. These policies range from adding carpool lanes, which is common in the United States, to rationing driving by allowing only certain license plate numbers to enter heavily congested areas on certain days or at certain times, which is common in Latin America. Even transit investment and ridership have shown signs of reversing long declining trends. Urban planners could encourage further concentration of jobs in sub-centers and then use public transit to link the sub-centers with population groups in need of better access to employment. Land-use and zoning policies might also be used in combination so as to induce a reduction in the use of the private automobiles and thereby reduce traffic congestion.[101]

The spatial structure of a city can be modified by a consistent action involving a close coordination between transport infrastructure investments, modification of land use, housing and planning regulations, taxation and distribution of subsidies. Nonetheless, spatial modification is a difficult task not only because of historical persistence but also due to the complexity of factors that determine urban spatial structure. Alain Bertaud argues that the consistency across the three tools is a hard find because they are often designed at different administrative levels of government and in most cases for very different objective, perhaps completely unrelated with city spatial structure.[102]

We conclude this section with three caveats. One, with the exception of zoning regulations, most policy prescriptions offered in this section do not directly affect urban form, but the effect on spatial organization comes through changes in density and land prices.[103] For instance, regulations on FAR does not directly determine whether land use in a given area would be allocated for business, residential or mixed purposes but it could potentially sway the density of a given area in the aimed direction. Specifically, stipulating a minimum FAR in a central region increases the supply of land and could potentially ease the transition from mixed land use area to a primarily business area. This effect may, however, depend on the several other factors such as the efficiency of transport networks, cost of communication, sectoral composition of the industry and the strength of agglomeration economies and so on.[104]

[100] See Gupta et al. (2006) for Austin; Santos and Shaffer (2004) for London

[101] See Cervero (1989) and Downs (1992) for a discussion on the policy for the development of subcenters in the US.

[102] Bertaud (2003)

[103] Zoning regulations can directly impact the land use pattern by designating areas for certain purposes. For instance, to reverse the chaotic growth of the city the government of the Federal District of Mexico issued a decree called Bando2 that established the main guidelines for urban development in the city for the next six years. Three zoning areas for housing development were created within the city: One, a re-densification zone that offered incentives to promote low cost housing and thereby achieve high density housing area; a second zone where restrictions were imposed on large housing developments that made high demands on water supply and infrastructure; and a third zone were zoning regulations were not changed (see Naranjo and Quintero, 2011).

[104] Similarly, UGBs, which are instituted with the objective of containing urban sprawl, typically drive land prices higher in the inner city area. UGBs are, therefore, likely to promote concentration of firms in the central regions while they hope to settle residences in the surrounding outer region. See Venkataraman (2014) for the impact of UGBs on land prices in Bangalore and Jun (2004) for an impact of UGBs in Portland on urban development patterns, transport choices and mobility.

Two, most available policies for affecting the pattern of land use are possibly only a second-best solution. For instance, zoning is agreeably not the first-best policy to influence land use because it does not completely internalize the externality. Similarly, it is argued that UGBs can easily yield undesirable outcomes, because they are not directly linked to the underlying market failures responsible for sprawl.[105] Regulations on maximum FAR are also considered to be at most a second-best policy because they only control the total land supply but that may not have the desired impact on actual per-capita consumption of space. Thus, before implementing the suggested policy, officials need to weigh the dead weight loss ensuing from such a change against the expected benefit it may have.

Three, our review of the literature points to several instances where the necessary policy-mix could be far more complex than assumed by economic models. We offer three such examples here:

(i) Strategies for coping with urban issues in a monocentric region should be distinct from those of polycentric settings because they face different issues on land use. Analysis from both theoretical as well as empirical findings, however, remains inconclusive regarding policy effects.[106]
(ii) The equilibrium urban form arising from, say, *imperfectly competitive markets,* such as that in intermediate inputs, à la Chamberlin (1933) style, could be far from optimum due to differences in demand conditions.[107]
(iii) Research in this field has only partially shed light on several questions pertaining to the economies of agglomeration and welfare. Firms agglomerate based on a variety of factors, including increasing returns and Smithian specialization, idiosyncratic matching, production externalities, and innovation. If such reasons motivate firms to agglomerate, the market cannot perhaps efficiently deal with all of them. Given our limited understanding of agglomeration economics at finer spatial scales, the policy implications of their incomplete internalization is not clear.

IV. Conclusions

This paper surveys the existing literature on the theory and evidence of jobs and firm location and land use changes within a city. The theoretical foundations of urban spatial organization suggest that both centrifugal and centripetal forces are at work. Agglomeration economies in the form of production externalities, urbanization economies or localization forces tend to enable clustering of firms while the cost of commute, congestion externalities, and the higher rates of decay of production externalities tend to disperse firm location. The equilibrium outcome is the result of a trade-off between these forces.

The survey illustrates that in the presence of externalities, equilibrium urban form generated by profit maximizing firms and utility maximizing workers may not be the socially optimal spatial structure. This

[105] Brueckner (2000b)
[106] . For example, in a city or region with only congestion externalities, marginal cost pricing is acceptably the first-best policy because it reflects the gap between marginal social and marginal private costs of each trip while this may not necessarily be true in a polycentric framework. See Pines and Sadka, (1985); Wheaton, (1998); Kono and Joshi, (2012). Also see, Giuliano and Small (1991); Crane and Chatman (2004)
[107] Abdel-Rahman and Fujita (1990), Henderson (1974).

calls for government action in preparing the city to take the most efficient spatial form that facilitates agglomeration of firms. Such policy options include changes to FAR, Urban growth boundary, congestion taxes, zoning regulations also help in defining the land use in central and peripheral locations. Further, increasing citywide connectivity through investments in transit networks and/or highways that reduce the costs of working at business centers is also likely to align the equilibrium outcome with the optimal solution. However, given that most of these are second best policy options, their merits need to be carefully weighed against the objective of changing the urban spatial structure.

Policies on urban spatial organization could be useful in accelerating structural transformation in cities such as those in Africa. Although traditionally urbanization was considered to be synonymous with structural transformation, studies on Africa reveal a perverse pattern. For instance, rapid urbanization in Ghana and Côte d'Ivoire in the past few decades has not been accompanied by an industrial or service revolution. Research shows that boom in the export of cash crops such as cocoa and coffee in African countries is responsible for the rise of "consumption" cities where spillover and linkage effects are lower.[108] Thus, even though the problem of Africa's urbanization without structural transformation may seem unique, the solution to this problem may, however, partly be found in urban spatial organization policies.

Finally, as a way forward to future research, our survey reveals that most empirical studies have not worked in tandem with the theoretical predictions provided. The future course of research in this area should align with the forces that have been identified in the literature as affecting the pattern of land use, returns to scale in production, spatial inhomogeneity, linkages, product variety, spatial interaction, traffic congestion, automobile pollution, and so on. Further, empirical work has largely focused on the trends in firm location and land use for developed countries. Much less is known about such patterns of industrial evolution in developing countries. The World Bank is responding to the lack of such studies through the "Spatial Development in Africa" program, which aims to understand the pattern of urban development in African cities and investigate how growing densities in Africa need to be supported by proactive policies and by appropriate infrastructure— including appropriate transport modalities, public transport, and so on. These countries remain widely unexplored. Finally, policy implications in this field are largely based on theoretical models and numerical simulations. Thus, there is more scope, in principle, to offer concrete policy lessons from studies on industrial evolution within cities once the theory and empirics are better aligned.

References

[108] See Jedwab (2013, 2011), Gollin et al. (2013), Collier and Dercon (2009) and Fafchamps (2003). In particular, Gollin et al. (2013) provide a conceptual framework to explain Africa's urbanization without structural transformation. Their generalization is based on the fact that urbanization in Africa is close to 40%, which compares with that in Asia, yet there has been little evidence of a Green Revolution in Africa. The continent has not seen an industrial revolution either. Its manufacturing and service sectors are relatively small and unproductive. It is perhaps this lack of evidence on concomitant structural transformation in Africa that a number of studies explain its urbanization in terms of push factors including civil wars, poor rural infrastructure, and climate change (Fay and Opal, 2000; Collier et al. 2008; Henderson et al., 2014).

Abdel-Rahman, H. and M. Fujita. 1990. Product variety, Marshallian externalities, and city sizes. Journal of Regional Science 30(2):165–183.

Agarwal, A., G. Giuliano, and C.L. Redfearn. 2012. Strangers in our midst: the usefulness of exploring polycentricity Ann Reg Sci, 48: 433–450

Alonso, William. 1964. Location and Land Use. Cambridge: Harvard University Press.

Anas, A. and Pines, D. 2008. Anti-sprawl policies in a system of congested cities. Regional Science and Urban Economics, 38: 408-423.

Anas A. and Rhee, H. 2007. When are urban growth boundaries not second-best policies to congestion tolls? Journal of Urban Economics, 61: 263–286.

Anas, H. Rhee, 2006. Curbing excess sprawl with congestion tolls and urban boundaries, Regional Science and Urban Economics 36 (2006) 510–541

Anas, A., Xu, R. 1999. Congestion, land use, and job dispersion: a general equilibrium model. Journal of Urban Economics 45: 451–473.

Anas, A., Arnott, R., Small, K.A. 1998. Urban spatial structure. Journal of Economic Literature, 36: 1426–1464.

Anas, A. and I. Kim. 1996. General equilibrium models of polycentric urban land use with endogenous congestion and job agglomeration. Journal of Urban Economics 40: 232–256.

Anderson, A.N. and W.T. Bogart. 2001. The structure of sprawl—identifying and characterizing employment centers in polycentric metropolitan areas. American Journal of Economics and Sociology, 60: 147–169.

Angel, S. and A Blei. 2015a. Commuting and the spatial structure of American cities. Marron Institute of Urban Management, New York University, Working Paper #20

Angel, S. and A Blei. 2015b. Commuting and the productivity of American cities. Marron Institute of Urban Management, New York University, Working Paper #19

Arnott, R. J. 1979. Unpriced transport congestion. Journal of Economic Theory, 21(2): 294-316.

Arnott, R., O. Hochman and G. C. Rausserc. 2008. Pollution and land use: Optimum and decentralization. Journal of Urban Economics, 64: 390–407

Arnott, R.J., MacKinnon, J.G. 1977. Measuring the costs of height restrictions with a general equilibrium model. Regional Science and Urban Economics 7: 359-375.

Arzaghi M. and J.V. Henderson, 2008. Networking Off Madison Avenue, *Review of Economic Studies* 75: 1011-1038.

Asabere, P. and K. Owusu-Banahene. (1983). Population Density Function for Ghanaian (African) Cities. Journal of Urban Economics, 14: 370-379

Atack, J., & Margo, R. A. 1998. Location, Location, Location! The Price Gradient for Vacant Urban Land: New York, 1835 to 1900. Journal of Real Estate Finance & Economics, 16(2), 151-172.

Avendaño A. and H. Enríquez. 2012. Polycentric structure and the co-agglomeration of economic activities in Bogota: A nonparametric perspective. Universidad Sergio Arboleda, Working Paper number 010394

Banerjee, A., E. Duflo and N. Qian, 2012. On the Road: Transportation Infrastructure and Economic Development, NBER Working Paper 17897.

Barlet M., A. Briant and L. Crusson. 2013. Location patterns of service industries in France: A distance-based approach. Regional Science and Urban Economics 43: 338–351

Barke, M., 1986. Transport and Trade. Oliver & Boyd.

Baum-Snow, N. 2014. Urban transport expansions, employment decentralization, and the spatial scope of agglomeration economies. Unpublished manuscript, Brown University.

Baum-Snow, N. 2007. Did highways cause suburbanization? Quarterly Journal of Economics 122(2):775–805.

Baum-Snow, N., L. Brandt, V. Henderson, M. Turner and Q. Zhang, 2012. Roads, Railroads and Decentralization of Chinese Cities, unpublished manuscript.

Baum-Snow, N. and B. F. Lutz. 2011. School desegregation, school choice, and changes in residential location patterns by race. American Economic Review 101(7):3019–3046.

Behrens, K. and F. Robert-Nicoud. 2015. Agglomeration theory with heterogeneous agents. In the Handbook of Regional and Urban Economics, Volume 5, by G. Duranton, J. V. Henderson, and W.S. Strange, Elsevier.

Berliant, M. and H. Konishi. 2000. The endogenous foundations of a city: Population agglomeration and marketplaces in a location-specific production economy. Regional Science and Urban Economics 30: 289–324.

Berliant, M., Peng, S., Wang, P. 2002. Production externalities and urban configuration. Journal of Economic Theory 104: 275–303.

Berliant, Marcus & Wang, Ping, 2008. "Urban growth and subcenter formation: A trolley ride from the Staples Center to Disneyland and the Rose Bowl," Journal of Urban Economics, Elsevier, 63(2): 679-693

Bertaud, A. 2003. The Spatial Organization of Cities: Deliberate Outcome or Unforeseen Consequence?, World Development Report 2003: Dynamic Development in a Sustainable World, World Bank.

Bertaud, A. 2009. Spatial profile of three South African cities. Presented at Spatial Outcomes Workshop South African Reserve Bank Conference Centre Pretoria September 29 - 30, 2009

Bertaud, A. (2011). Mumbai FAR/FSI conundrum. Available at: http://alainbertaud.com/wp-content/uploads/2013/06/AB-Mumbai-FSI-Conundrun-Revised_June-2013_kk-ab1.pdf

Bertaud, A., Brueckner, J.K. 2005. Analyzing building-height restrictions: predicted impacts and welfare costs. Regional Science and Urban Economics 35: 109–125.

Bertaud, A. and S. Malpezzi. 2003. The spatial distribution of population in 48 world cities: Implications for economies in transition, Unpublished manuscript.

Bertaud, A. and S. Malpezzi. 2014. The Spatial Distribution of Population in 57 World Cities: The Role of Markets, Planning, and Topography, Unpublished manuscript.

Bento, A.M., Franco, S., Kaffine, D., 2006. The efficiency and distributional impacts of alternative anti-sprawl policies. Journal of Urban Economics 59: 121–141.

Bird, J., 1980, Seaports and seaport terminals (London, Hutchinson University Library)

Bleakley, H. and J. Lin. 2012. Portage and path dependence. Quarterly Journal of Economics 127(2):587–644.

Bogart, W.T. and W.C. Ferry. 1999. Employment centers in greater Cleveland: Evidence of evolution in a formerly monocentric city. Urban Studies, 36: 2099–2110.

Borukhov, E., Hochman, O., 1977. Optimum and market equilibrium in a model of a city without a predetermined center. Environment and Planning A 9: 849–856.

Boustan, L.P. 2010. Was postwar suburbanization “white flight”? Evidence from the black migration. Quarterly Journal of Economics 125(1):417–443.

Brinkman, J. C. 2013. Congestion, Agglomeration, and the Structure of Cities. Federal Reserve Bank of Philadelphia, Working Paper No. 13-25

Brueckner, J.K. 1979. A model of non-central production in a monocentric city, Journal of Urban Economics, 6: 444–463.

Brueckner, J.K. 2000a. Urban growth models with durable housing: an overview. In: Huriot, J.-M., Thisse, J.-F. (Eds.), Economics of Cities. Cambridge University Press, Cambridge: 263–289.

Brueckner, J.K. 2000b. Urban sprawl: Diagnosis and remedies, International Regional Science Review 23: 160–171.

Brueckner, J. K. 2001. Urban sprawl: Lessons from urban economics, in: W.G. Gale, J.R. Pack (Eds.), Brookings–Wharton Papers on Urban Affairs, Brookings Institution Press, Washington, DC: 65–89.

Brueckner, J. 2007. Urban growth boundaries: an effective second-best remedy for unpriced traffic congestion? Journal of Housing Economics, 16: 263-273.

Brueckner, Jan K. and Stuart S. Rosenthal. 2009. Gentrification and neighborhood housing cycles: will America's future downtowns be rich? Review of Economics and Statistics 91(4): 725–743.

Burger, M.J., F.G. Van Oort, K. Frenken, and B. van der Knaap. 2009. Networks and economic agglomerations. In Tijdschrift voor Economische en Sociale Geographie 11: 139N144.

Carlino, G. and Chatterjee, S., 2001. Aggregate Metropolitan Employment Growth and the Deconcentration of Metropolitan Employment, Journal of Monetary Economics, 48: 549-583

Carlino, G. and Chatterjee, S., 2002. Employment Deconcentration: A New Perspective on America's Postwar Urban Revolution, Journal of Regional Science, 42: 455-475.

Cervero, R. 1989. America's Suburban Centers: The Land Use-Transportation Link. Boston: Unwin Hyman

Cervero, R. and K.L. Wu. 1997. Polycentrism, commuting, and residential location in the San Francisco Bay Area, Environment and Planning A 29: 865–886.

Cervero, R. and K. Wu. 1998. subcentering and commuting: Evidence from the San Francisco Bay area, 1980–1990. Urban Studies, 35: 1059–1076.

Chatterjee, S. and B. Eyigungor. 2014. A Tractable Circular City Model with Endogenous Internal Structure. Federal Reserve Bank of Philadelphia. Paper presented at Conference on Urban and Regional Economics (CURE), Brown University.

Chamberlin, E. H. 1993. The theory of monopolistic competition. Cambridge: Harvard University press.

Chinitz, B. 1960. Freight and the Metropolis, Cambridge, Mass: Harvard University Press.

Cho, S. H., Poudyal, N., and Lambert, D. M. 2008. Estimating spatially varying effects of urban growth boundaries on land development and land value. Land Use Policy, 25(3): 320-329.

Ciccone A. and G. Peri. 2006. Identifying Human Capital Externalities: Theory with Applications. The Review of Economic Studies, 73(2):381–412

Collier, P., G. Conway, and A. Venables. 2008. Climate change and Africa. Oxford Review of Economic Policy: 24(2), 337-353

Collier, P. and S. Dercon. 2009. "African Agriculture in 50 Years: Smallholders in a Rapidly Changing World?" Paper prepared for the FAO Expert Meeting on How to Feed the World in 2050, FAO.

Combes, P. and L. Gobillon. 2015. The Empirics of Agglomeration Economies. In the Handbook of Regional and Urban Economics, Volume 5, by G. Duranton, J. V. Henderson, and W.S. Strange, Elsevier.

Combes, P. G. Duranton, and L. Gobillon. 2012. The costs of agglomeration:
Land prices in French cities. Processed, Wharton School, University of Pennsylvania.

Combes, P. G. Duranton, L. Gobillon, and S. Roux. 2011. ”Estimating Agglomeration Economies with History, Geology, and Worker Effects.” in The Economics of Agglomeration, edited by E. L. Glaeser, NBER and University of Chicago Press.

Craig, S.G and P. Ng. 2001. Using quantile smoothing splines to identify employment subcenters in a multicentric urban area. Journal of Urban Economics, 49: 100–120.

Crane, R. and D. Chatman. 2004. Travel and Sprawl: Evidence from US commuting, 1985 to 1997. In Richardson, H. W. and Bae, C-H C. (eds.) Urban Sprawl in Western Europe and the United States: 311-334.

Cronon, W. 1991. Nature's Metropolis: Chicago and the Great West (W.W. Norton and Company, New York).Cullen, Julie Berry and Stephen D. Levitt. 1999. Crime, urban flight, and the consequences for cities. Review of Economics and Statistics 81(2):159–169.
Davis, J. and V. Henderson. 2008. The Agglomeration of Headquarters. Regional Science and Urban Economics, 38(5): 445–460.

De Silva D. and R. McComb. 2011 Geographic concentration and firm survival. MPRAPaper 32906. University Library of Munich, Germany

Deng, X., J. Huang, S. Rozelle and E. Uchida. 2008. Growth, population and industrialization, and urban

land expansion of China, Journal of Urban Economics, 63: 96–115.

Desmet, K. and M. Fafchamps. 2006. Employment Concentration across US counties. Regional Science and Urban Economics 36: 482–509

Desmet, K. and E. Rossi-Hansberg. 2009. Spatial Growth and Industry Age. Journal of Economic Theory, 144: 2477-2502

Desmet K. and E. Rossi-Hansberg. 2014. Spatial Development. American Economic Review, American Economic Association,104: 1211-43.

Desmet, K., E. Ghani, S. O'Connell and E. Rossi-Hansberg. 2015. The Spatial Development of India. Journal of Regional Science, 55: 10-30

Dixit, A., 1973. The optimum factory town. Journal of Economics and Management 4: 637–651.

Donaldson, D., 2014. Railroads of the Raj: Estimating the Impact of Transportation Infrastructure, American Economic review, forthcoming

Downs, A. 1992. Stuck in Traffic. Washington, DC: The Brookings Institution.

Duarte, C.M., C.A. NuÃ±ez, E.C. Mejia, C.P. Prieto and J.M. Tresserra. 2011. Polycentrism in the Spanish metropolitan system: an analysis for 7 metro areas. European Regional Science Association in its series ERSA conference papers with number ersa11p119.

Dumais, G., Ellison, G., Glaeser, E. 2002. Geographic concentration as a dynamic process. Review of Economics and Statistics 84 (2): 193–204.

Duranton, G. and D. Puga. 2001. Nursery cities: Urban diversity, process innovation, and the life cycle of products, Amer. Econ. Rev. 91: 1454–1477.

Duranton, G. and D. Puga. 2004. Microfoundations of urban agglomeration economies. In Vernon Henderson and JacquesFrançois Thisse (eds.) Handbook of Regional and Urban Economics, volume 4. Amsterdam: NorthHolland.

Duranton G. and D. Puga. 2005. From Sectoral to Functional Urban Specialization. Journal of Urban Economics 57: 343–70.

Duranton, G. and D. Puga. 2014. The Growth of Cities In: Handbook of Economic Growth, edition 1, vol. 2, chapter 5: 781-853 Elsevier.

Duranton, G. and D. Puga. 2015. Urban Land Use. In the Handbook of Regional and Urban Economics, Volume 5, by G. Duranton, J. V. Henderson, and W.S. Strange, Elsevier.

Duranton, G. and M. A. Turner. 2011. The fundamental law of road congestion: Evidence from US cities. American Economic Review, 101:, 2616-2652.

Ellison, G. and Glaeser, E. 1997. Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach. Journal of Political Economy, 105 (5): 889-927

Fafchamps, M. 2003. Engines of Growth and Africa's Economic Performance. In Emmanuel Nnadozie, African Economic Development, Academic Press, New York: 65- 98.

Fales, R. and L. N. Moses. 1972. Land Use Theory and the Spatial Structure of the Nineteenth Century City. Papers & Proceedings Regional Science Association, 28:49-80.

Fay, M., and C. Opal. 2000. Urbanization without growth : a not so uncommon phenomenon. World Bank Policy Research Working Paper 2412.

Fernandes, A. M. and G. Sharma. 2012. Together We Stand? Agglomeration in Indian Manufacturing, Policy Research Working Paper #6062, The World Bank.

Fujita, M. 1988. A monopolistic competition model of spatial agglomeration: A differentiated product approach, Regional Science and Urban Economics 18: 87- 124.

Fujita M. 1989. Urban economic theory: land use and city size. Cambridge University Press, Cambridge
Fujita, M., P. Krugman, T. Mori. 1999. On the evolution of hierarchical urban systems. European Economic Review, 43: 209–251.

Fujita, M. and T. Mori. 1997. Structural stability and evolution of urban systems. Regional Science and Urban Economics 27(4–5): 399–442.

Fujita, M. and H. Ogawa. 1982. "Multiple equilibria and structural transition of non-monocentric urban configurations," Regional Science and Urban Economics, Elsevier, vol. 12(2): 161-196

Fujita, M., J-F Thisse, and Y. Zenou. 1997. On the endogenous formation of secondary employment centers in a city. Journal of Urban Economics 41: 337–357.

Garcia-López, M.A. 2010a. The Accessibility City - When Transport Infrastructure Matters in Urban Spatial Structure (February 1, 2010). Document de Treball No. XREAP2010-1

Garcia-López, M.A. 2010b. "Population suburbanization in Barcelona, 1991-2005: Is its spatial structure changing?," Journal of Housing Economics, Elsevier, vol. 19(2): 119-132.

García-López, M.A., A Holl and E. Viladecans-Marsal. 2013. Suburbanization and the highways: when the Romans, the Bourbons and the first cars still shape Spanish cities, Institut d'Economia de Barcelona working paper 2013/05.

Ghani, E., A. Goswami, and W. Kerr. 2012. Is India's Manufacturing Sector Moving Away from Cities ?, NBER Working Paper #17992.

Ghani, E., A. Goswami, and W. Kerr. 2014. Highway to Success: The Impact of the Golden Quadrilateral Project for the Location and Performance of Indian Manufacturing. Economic Journal, *forthcoming*

Giuliano, G. and K.A. Small. 1991. Subcenters in the Los Angeles region, Regional Science and Urban Economics, 21: 163–182.

Giuliano, G., Redfearn, C., Agarwal, A., Li, C., & Zhaun, D. 2007. Employment concentrations in Los Angeles, 1980 – 2000; Environment and Planning, A39: 2935–2957

Glaeser, E. 1999. Learning in cities. Journal of Urban Economics 46:254–277.

Glaeser, E. and M. Kahn. 2001. Decentralized employment and the transformation of the American city. BrookingsWharton Papers on Urban Affairs :1–47.

Glaeser, E., Kahn, M.E., 2004. Sprawl and urban growth. In: Henderson, J.V., Thisse, J.-F. (Eds.), Handbook of Regional and Urban Economics, vol. 4. North Holland, Amsterdam, the Netherlands, pp. 2482–25

Glaeser E., H. Kallal, J. Scheinkman and A. Shleifer. 1992. Growth in cities. Journal of Political Economy, 100: 1126–1152

Glaeser, E.L., J.E. Kohlhase. 2004. Cities, regions and the decline of transport costs. Papers in Regional Science, 83: 197-228

Glaeser, E., Mare, D.C., 2001. Cities and skills. Journal of Labor Economics 19: 316–342.

Glaeser, E., and G. Ponzetto. 2010. "Did the Death of Distance Hurt Detroit and Help New York?" In Agglomeration Economics, edited by Edward L. Glaeser, 302–37. Chicago: University of Chicago Press.

Gollin, D., R. Jedwab, and D. Vollrath. 2013. Urbanization with and without Structural Transformation. Unpublished manuscript, George Washington University, Department of Economics.

Gonzales de Olarte, Efraín; del Pozo Segura, Juan Manuel. Lima, una ciudad policéntrica. Un análisis a partir de la localización del empleo. Investigaciones Regionales, núm. 23, 2012, pp. 29-52. Asociación Española de Ciencia Regional. Madrid, España. http://www.redalyc.org/articulo.oa?id=28924472002

Gordon, I. and P. McCann. 2000. Industrial clusters: Complexes, agglomerations and/or social

networks? Urban Studies 37: 513-532.

Gordon, P., A. Kumar, and H.W. Richardson. 1989. The influence of metropolitan spatial structure on commuting time, Journal of Urban Economics, 26: 138–151.

Gordon, P. and H.W. Richardson. 1996. Beyond polycentricity: The dispersed metropolis, Los Angeles 1970–1990, Journal of the American Planning Association 62: 289–295.

Gupta, S., S. Kalmanje, and K. M. Kockelman. 2006. Road Pricing Simulations: Traffic, Simulations and Welfare Impacts for Austin, Texas. Transportation Planning and Technology 29(1): 1–23.

Hayuth, Y., 1981. Containerization and the load centre concept. Economic Geography 57: 160–176.

Hayuth, Y., 1988, Rationalization and deconcentration of the US container port system, The Professional Geographer, 40 (3): 279-288

Helpman, E., Pines, D., 1977. Land use and zoning in an urban economy: further results. American Economic Review 67: 982–986.

Henderson, J. Vernon. 1974. The sizes and types of cities. American Economic Review 64(4):640–656.

Henderson, J. Vernon. 2005. ”Urbanization and Growth.” in (eds) Philippe Aghion and Steven Durlauf, Handbook of Economic Growth, North Holland.

Henderson, J. Vernon. 2010. Cities and Development. Journal of Regional Science 50(1): 515-540.

Henderson, J.V. and A. Kuncoro, 1996. The Dynamics of Jabotabek Development, Bulletin of Indonesian Economic Studies.

Henderson J, Kuncoro A, Turner M. 1995. Industrial development in cities. Journal of Political Economy, 103: 1067–1085.

Henderson J.V. and Mitra A. 1996. The New Urban Landscape: Developers and Edge Cities, Regional Science and Urban Economics, 26: 613-643.

Henderson, V. and Y. Ono. 2008. Where do manufacturing firms locate their headquarters? Journal of Urban Economics, 63: 431-450

Henderson, J.V., Slade, E., 1993. Development games in nonmonocentric cities. Journal of Urban Economics 34: 207–229.

Henderson, J.V., A. Storeygard and U. Deichmann. (2014). Is climate change driving urbanization in Africa? Unpublished draft.

Holl, A. and E. Viladecans-Marsal, 2011. Infrastructure and Cities: The Impact of New Highways on Urban Growth, Universitat de Barcelona-Institut d'Economia de Barcelona Working Paper.

Hochman, O., Pines, D. 1982. Adjustment costs and the spatial pattern of a growing open city. Econometrica 6: 1371–1391.

Hotelling, H., 1929. Stability in competition. Economic Journal, 39: 41–57.

Hsu, W-T and H. Zhang. 2014. The Fundamental Law of Highway Congestion: Evidence from Japanese Expressways, Journal of Urban Economics, 81: 65-76.

Imai, Haruo. 1982. CBD hypothesis and economies of agglomeration. Journal of Economic Theory 28:275–299.

Jacobs, J. 1969. The Economy of Cities. Random House.

Jedwab, R. 2013. Urbanization without Structural Transformation: Evidence from Consumption Cities in Africa. Mimeo, George Washington University.

Jedwab, R. 2011. African Cities and the Structural Transformation: Evidence from Ghana and Ivory Coast. Mimeo, Paris School of Economics.

Johansson, B. and J.M. Quigley. 2004. Agglomeration and networks in spatial economics. Papers in Regional Science 83: 165-176.

Joshi , K.K. and T. Kono. 2009. Optimization of floor area ratio regulation in a growing city. Regional Science and Urban Economics, 39: 502-511.

Jun, M. (2004). The Effects of Portland's Urban Growth Boundary on Urban Development Patterns and Commuting. Urban Studies, 41(7): 1333–1348

Kim, S., 1995. Expansion of markets and the geographic distribution of economic activities: the tends in U.S. regional manufacturing structure, 1860–1987. Quarterly Journal of Economics 110: 881–908.

Kolko, J. 1999. Can I get some service here? Information technologies, service industries and the future of cities. Processed, Harvard University.

Kolko J. 2010. Urbanization, agglomeration, and coagglomeration of service industries. In: Glaeser E (ed) Agglomeration economies. National Bureau of Economic Research, Cambridge

Kono, T. and K.K. Joshi. 2012. A new interpretation on the optimal density regulations: Closed and open city. Journal of Housing Economics, 21(3): 223-234.

Kono, T., Kaneko, T. and Morisugi, H. 2010. Necessity of Minimum Floor Area Ratio Regulation: a Second-best Policy. Annals of Regional Science, 44: 523-539

Kono,T., Joshi, K.K., Kato, T., and Yokoi, T. 2012. Optimal regulation on building size and city boundary: An effective second-best remedy for traffic congestion externality. Regional Science and Urban Economics, 42: 619–630.

Koopmans, Tjalling C. 1957. *Three Essays on the State of Economic Science*. New York: McGrawHill.

Kuby, M., Reid, N., 1992, Technological change and the concentration of the U.S. General Cargo Port System: 1970-1988, Economic Geography, 68(3): 272-289

Kyriakopoulou, E. and A. Xepapadeas. 2013. Spatial Policies and Land Use Patterns Optimal and Market Allocations. University of Gothenberg, Working paper No. 566

Lang, R. E., 2003, Edgeless Cities: Exploring the Elusive Metropolis. Washington, DC: Brookings Institution Press.

Lucas, R. E. Jr. 2001. Externalities and Cities. Review of Economic Dynamics, 4: 245-274.

Lucas, R.E. Jr., E. Rossi-Hansberg. 2002. On the internal structure of cities. Econometrica 70: 1445–1476.

Marshall, A. 1890. Principles of Economics. Macmillian.

McCalla, R., 1999, From St. John's to Miami: containerisation at Eastern Seaboard ports, GeoJournal, 48: 21-28

McDonald, J.F. and P.J. Prather. 1994. Suburban employment centres: The case of Chicago, Urban Studies, 31: 201–218.

McMillen, D. P. 1996. One hundred fifty years of land values in Chicago: A nonparametric approach. Journal of Urban Economics 40(1):100–124.

McMillen, D.P. 2001a. Nonparametric employment subcenter identification, Journal of Urban Economics 50: 448– 473.

McMillen, D. P. 2001b. Polycentric urban structure: The case of Milwaukee. Economic Perspectives, 25(2): 15-27.

McMillen, D.P. 2003. Identifying subcenters using contiguity matrices. Urban Studies, 40: 57–69.

McMillen, D. P. and Lester, T. W. 2003: Evolving subcenters: employment and population densities in Chicago, 1970–2020. Journal of Housing Economics, 12: 60–81.

McMillen, D.P., Smith, S.C. 2003. The number of subcenters in large urban areas. Journal of Urban Economics 53: 321–338.

McMillen, D.P and J.F. McDonald. 1998. Suburban subcenters and employment density in metropolitan Chicago, Journal of Urban Economics, 43:. 157–180.

McMillen, D.P. and J.F. McDonald. 1997 A nonparametric analysis of employment density in a polycentric city, Journal of Regional Science 37: 591–612.

Mills, E. 1967. An aggregative model of resource allocation in a metropolitan area. *American Economic Review Papers and Proceedings* 57(2):197–210.

Mills, E. 1969. "The Value of Urban Land," The Quality of the Urban Environment (Harvey Perloff, Editor). Baltimore: Johns Hopkins Press for Resources for the Future, 1969: 231-253.

Mills, E. 1972. An Aggregative Model of Resource Allocation in a Metropolitan Area. The American Economic Review, 57(2): 197-210

Mills, E., and J.P. Tan. (1980). A Comparison of Urban Population Density Functions in Developed and Developing Countries, Urban Studies, 17: 313-321

Mirrlees, J. 1972. The optimum town. *Swedish Journal of Economics* 74(1): 114–135.

Moretti, E., 2004. Human capital externalities in cities. In: Henderson, J.V., Thisse, J.-F. (Eds.), Handbook of Urban and Regional Economics, vol. 4. Elsevier, Amsterdam, pp. 2242–2292.

Muller, E. 2001. Industrial suburbs and the growth of metropolitan Pittsburgh, 1870–1920. Journal of Historical Geography, 27(1): 58–73

Mu˜niz, I., Garcia-L´opez, M., Galindo, A. 2008. The effect of employment sub-centres on population density in Barcelona. Urban Studies 45: 627–649.

Muth, R. 1969. Cities and Housing. Chicago: University of Chicago Press.

Naranjo D. and O. Quintero. (2011). Effects of Urban Density Regulation on Land Prices: The Case of Bando2 in Mexico City. Lincoln Institute of Land Policy, working paper No. 49

Notteboom, T. 1997, Concentration and load centre development in the European container port system, Journal of Transport Geography, 5 (2): 99-115

Ogawa, H. and M. Fujita. 1980a. Equilibrium land use patterns in a nonmonocentric city. Journal of Regional Science 20: 455–475.

Ogawa. H and M. Fujita. 1980b. Multicentric urban configuration: equilibrium and optimum,

Working Papers in Regional Science and Transportation, 25 (University of Pennsylvania, Philadelphia, PA).

Ogawa. H. and M. Fujita. 1979. Nonmonocentric urban configurations in two-dimensional space, Working Paper, in Regional Science and Transportation, 18 (University of Pennsylvania. Philadelphia, PA).

Ota, M and M. Fujita. 1993. Communication technologies and spatial organization of multiunit firms in metropolitan areas. Regional Science and Urban Economics 23(6):695–729.

Overman, H. and A. Venables. 2005. Cities in the Developing World. CEP Discussion Paper dp0695

Ottaviano, G. and J. Thisse. 2004. Agglomeration and economic geography. In Vernon Henderson and JacquesFrançois Thisse (eds.) Handbook of Regional and Urban Economics, volume 4. Amsterdam: NorthHolland.

Paci R. and S. Usai. 1999. Externalities, knowledge spillovers and the spatial distribution of innovation. Geojournal 49: 381–390

Picard, P. M and T. Tabuchi. 2013. On Microfoundations of the City. Journal of Economic Theory, 148(6): 2561–2582

Pines, D. and T. Kono. 2012. FAR regulations and unpriced transport congestion. Regional Science and Urban Economics, 42: 931-937

Pines, D. and E. Sadka. 1985. Zoning, first-best, second-best and third-best criteria for allocating land for roads. Journal of Urban Economics, 17: 167-183.

Rauch, J., 1993. Productivity games from geographic concentration of human capital: Evidence from the cities. Journal of Urban Economics 108: 380– 400.

Redfearn, C. 2007a. The Topography of Metropolitan Employment: Identifying Centers of Employment in a Polycentric Urban Area. Journal of Urban Economics, 61: 519-541.

Redfearn, C. 2007b. Determinacy in Urban Form: Fixed Investment & Path Dependence in Urban Areas. School of Policy, Planning, and Development University of Southern California. Unpublished draft.

Redfearn, C. 2009. Persistence in urban form: The long-run durability of employment centers in metropolitan areas. Regional Scicence and Urban Economics, 39: 224–232

Roberts, M., U. Deichmann, B. Fingleton, and T. Shi. 2012. Evaluating China's Road to Prosperity: A New Economic Geography Approach. Regional science and urban economics, 42: 580-594

Rosenthal S. and W. Strange. 2001. The determinants of agglomeration. Journal of Urban Economics, 50: 191–229

Rosenthal S. and W. Strange. 2003. Geography, industrial organization, and agglomeration. The Review of Economics and Statistics 85(2):377–393

Rosenthal S. and W. Strange. 2004. Evidence on the nature and sources of agglomeration economies. In: Henderson J, Thisse J-F (eds) Handbook of regional and urban economics, vol 4. North Holland, New York, pp 2119–2171

Rossi-Hansberg, E. 2004. Optimal urban land use and zoning. Review of Economic Dynamics, 7: 69–106.

Rossi-Hansberg, E., P. Sarte, and R. Owens. 2009. Firm Fragmentation and Urban Patterns. International Economic Review, 50(1): 143-186

Rothenberg, A., 2013. Transport Infrastructure and Firm Location Choice in Equilibrium: Evidence from Indonesia's Highways, manuscript.

Small, K. and S. Song. 1994. Population and employment densities—Structure and change, Journal of Urban Economics, 36: 292–313.

Sakashita, N., 1995. An economic theory of urban growth control. Regional Science and Urban Economics 25: 427–434.

Santos, G. 2005. Urban Congestion Charging: A Comparison between London and Singapore. Transport Reviews 25 (5): 511–34.

Santos, G., and B. Shaffer. 2004. Preliminary Results of the London Congestion Charging Scheme. Public Works, Management and Policy, 9: 164–181.

Schafer, A. 2000. Regularities in travel demand: An international perspective. Journal of Transportation and Statistics, 3:1–31.

Schweizer, U., P. Varaiya and J. Hartwick. 1976. General equilibrium and location theory. Journal of Urban Economics, Elsevier, 3: 285-303

Simon, C. 2004. Industrial reallocation across US cities, 1977-1997. Journal of Urban Economics, 56:119-143.

Staley, S., J. Edgens and G. Mildner. 1999. A line in the land: Urban-growth boundaries, smart growth, and housing affordability. Reason Public Policy Institute

Steen, R.C., 1986. Nonubiquitous transportation and urban population density gradients. Journal of Urban Economics 20: 97–106.

Storeygard, A. 2012. Farther on down the road: transport costs, trade and urban growth in sub-Saharan Africa, manuscript.

Stull, W., 1974. Land use and zoning in an urban economy. American Economic Review 64: 337–347.

Sullivan A. 1986. A general equilibrium model with agglomerative economies and decentralized employment. Journal of Urban Economics, 20: 55–74

Taaffe, E., R. Morrill, and P. Gould. 1963. Transport expansion in underdeveloped countries: a comparative analysis. Geographical Review, 53: 503-529

Venables, A. 2010. Economic Geography and African Development. Journal of Regional Science 89(3): 469-483.

Venkataraman, M. (2014). Analysing Urban Growth Boundary Effects in the City of Bengaluru. Indian Institute of Management, Working Paper No : 464.

Walker, R. 2004. Industry Builds out The City: The Suburbanization of Manufacturing in the San Francisco Bay Area, 1850-1940. Published in The Manufactured Metropolis by Robert Lewis (ed.)
Temple University Press: 92-123

Wheaton, W. 1978. Price-Induced Distortions in Urban Highway Investment. Bell Journal of Economics, 9: 622-632.

Wheaton, W. 1998. Land use and density in cities with congestion. Journal of Urban Economics, 43: 258–272.

Wheaton, W. 2004. Commuting, congestion, and employment dispersal in cities with mixed land use. Journal of Urban Economics, 55: 417–438.

White, M. 1976. Firm suburbanization and urban subcenters. Journal of Urban Economics 3: 323–343.

Zhang, W. and K. Kockelman. 2014. Urban Sprawl, Job Decentralization, and Congestion: The Welfare Effects of Congestion Tolls and Urban Growth Boundaries. Proceedings of the 93rd Annual Meeting of the Transportation Research Board, 2014.

Zhang Y. and K. Sasaki. 1997. Effects of subcenter formation on urban spatial structure. Regional Science and Urban Economics, 27(3): 297–324

Zhang Y. and K. Sasaki. 2000. Spatial structure in an open city with a subcenter. Annals of Regional Science, 34(1): 37–53

Annex 1

Table A.1: Research on Policy Options for aligning equilibrium land use with optimal allocation

Name of authors	Context	Suggested Instruments	Main findings
Rossi-Hansberg (2004)	Non-monocentric (Endogenous land use)	government subsidy that lowers the labor costs, parking lots construction, highway investments	Positively impact rents, brings equilibrium allocation closer to optimum
Zhang and Kockelman (2014)	Non-monocentric (Endogenous land use)	Congestion toll policy	Leads to compact urban form and job decentralization, and serves as an effective strategy for correcting congestion externalities. First best policy
Arnott et al. (2008)	Non-monocentric (Endogenous land use)	Congestion toll policy	Spatially differentiated Pigouvian taxes per unit emission levied on industrial polluters does not support the optimum. Pigouvian taxes are optimal only under special circumstances.
Brinkman (2013)	Non-monocentric (Endogenous land use)	Congestion toll policy	In the presence of agglomeration economies, a naive optimal congestion pricing policy may lead to welfare loss
Pines and Sadka (1985); Wheaton (1998)	Monocentric	Congestion toll policy	Congestion tax is the first best policy option to encourage decentralization around the city core.
Anas and Kim (1996); Anas and Xu (1999); Wheaton (2004); Zhang and Kockelman (2014)	Polycentric	Congestion toll policy	welfare improving; delivers a more compact urban form and also corrects for the associated market failure
Zhang and Kockelman (2014)	Non-monocentric (Endogenous land use)	Urban growth boundary (UGB)	Alleviate congestion externalities and lower travel times, vehicle-miles traveled, and travel costs; but carry certain loss of social welfare owing to land rent escalation and UGBs' limitations on job decentralization.
Pines and Sadka (1985); Arnott (1979)	Monocentric	Urban growth boundary (UGB)	Effective second-best substitute for first-best tolling
Brueckner (2007); Kono et al., (2012)	Monocentric	Urban growth boundary (UGB)	welfare gain from an optimal UGB policy against unpriced congestion is only a tiny fraction of that gained under the toll regime, implying that the UGB policy is a poor substitute for the toll regime.
Pines and Sadka (1985); Wheaton (1998)	Monocentric	Lot size Zoning	Result similar to a first-best congestion toll
Arnott and MacKinnon (1977); Bertaud and Brueckner (2005)	Monocentric	Building height regulations	invokes distortions that may be associated with large welfare costs
Arnott and MacKinnon (1977); Bertaud and Brueckner (2005)	Monocentric	building size regulations	only maximumbuilding size regulation is considered, ignoring optimality of the regulation
Kono et al., (2012)	Monocentric	building size regulations (FAR)	welfare gain under the optimal FAR policy with a suitably chosen UGB makes up a significant fraction of that achieved under the toll regime. Even without a UGB, welfare gain under the optimal FAR policy alone is also substantial. As compared to optimal lot size zoning, FAR regulations can at most be second best policy option for controlling population density.
Kono and Joshi (2012)	Monocentric	Floor area ratio (FAR)	In a closed city, it is optimal to allow not only downward adjustment to the market density at boundary locations using maximum FAR regulation but also upward adjustment at central locations using minimum FAR regulation. The common practice of imposing maximum FAR regulation is effective in an open city.*
Kono et al., (2010); Joshi and Kono (2009); Sakashita (1995);	Monocentric	Floor area ratio (FAR)	In a closed city, FAR regulation is the second-best policy and that there is a need for minimum FAR regulations in addition to the currently used maximum FAR regulations
Bertaud and Brueckner (2005)	Monocentric	Floor area ratio (FAR)	welfare cost of FAR regulation increases with an increase in the commuting costs; however, only maximum FAR regulation is considered, ignoring optimality of the regulation.
Pines and Kono (2012)	Monocentric	Floor area ratio (FAR)	FAR optimal regulations render UGB redundant but it cannot always replace the fiscal instruments (e.g. spatially-variable excise subsidies, or taxes, on housing) for achieving the second-best utility.
Brueckner (2000b, 2001)	Monocentric	Anti-sprawl policies	Development taxes and congestion tolls seem to be effective in reducing the spatial size of the city
Bento et al. (2006)	Monocentric	Anti-sprawl policies	From an overall efficiency perspective, development taxes and UGB are equivalent instruments and the most effective anti-sprawl policies. By contrast, gasoline tax and the property tax are substantially less effective in combating sprawl.
Anas and Rhee (2006)	Non-monocentric (Endogenous land use)	Anti-sprawl policies	Congestion tolls eliminate the excess sprawl by making the job and residence distribution of the city more compact while UGBs result in substantial efficiency loss.

*Notes: * A closed city is defined as the one where utility is determined endogenously given the population size while in an open city, the size of the population is determined endogenously for a given utility level.*