Digital Cities: Real Estate Development Driven By Big Data

By

Herman Donner,

Kent Eriksson, and Michael Steep

Stanford School of Engineering

Disruptive Technology and Digital Cities Program

Screenshot

Digital Cities: Real Estate Development Driven by Big Data

Herman Donner, Kent Eriksson, and Michael Steep

Abstract

Urban environments are composed of urban population, urban infrastructure, city governance and commercial markets within cities. The rapid growth of emerging technologies for sensing and communicating data is being leveraged by commercial companies to create digital applications where machine learning applications analyze multiple kinds of data now available from instrumented infrastructure, public and private urban transactions and citizens' mobility to transform urban environments. This kind of transformation is our view of what enables a "digital city". Commercial markets are at the heart of this concept, with commercial applications of digital infrastructure rapidly developing, because data from multiple sources are more easily available and analyzed across multiple data layers drawn from different sectors and regions of the city. It is now possible to visualize multiple kinds of outcomes across an entire city and its markets, and to do "What if?" analysis using predictive analytics to generate new insights and financial models across a wide range of vertical urban services. The ability to visualize real time data and insights drawn from that data about the urban environment that surrounds real estate and identify its connection with real estate value provides an unprecedented potential for enhancing real estate development decisions, primarily through better forecasts for building utilization, more accurate assessment of the purchasing power of users of real estate, and by better risk assessment of real estate users. This article presents an analysis of the potential benefits of digital cities for real estate development decision making.

Digital Cities: Real Estate Development Driven by Big Data

1. Introduction

This article is a proposal for fundraising for a new and transformative way of analyzing and determining real estate value by application of digital infrastructure in cities. End user demand determines real estate values through rents, vacancies, and risk characteristics. It is therefore surprising that so little research has analyzed how big data analysis of user characteristics and preferences can be applied to real estate development. In other areas, the impact of information and communications technologies (ICT) has been transformative, giving rise to new models for business, organization, finance, employment, and service delivery. We analyze the urban and societal impact of new technologies, and their application to real estate development.

Physical things are now creating enormous amounts of big data through sensors across the urban environment, on buildings, roads, street lights, infrastructure, and various other places. Smartphones, tablets, applications for consumption, social media platforms, and network connected vehicles are all part of what is known as the internet of things (IoT) that allows for collection and application of knowledge on the environment and its inhabitants (Yovanof and Hazapis, 2009). Consequently, businesses are increasingly using big data analysis to predict market trends and gain understanding of customers and products (Hashem et al., 2016). This development has led to the argument by some, that personal data is becoming the "*oil of the 21st century*", emerging as a new asset class (World Economic Forum, 2011).

From a perspective of real estate development, identification of relationships between movement, commercial transactions, social activities, and real estate value is a potential game changer that will increase the value and quality of the urban environment. We propose that a city can be conceptualized as layers of various big data that makes it is possible to identify where, by whom, and when money is spent, and where people are moving, in real time. When connected to real estate values, this allows for prediction of various *What If* scenarios. More specifically, we analyze possibilities created by merging data on a platform for modelling and visualization of such scenarios, something that has never been possible before, and is part of a current project at the school of engineering at Stanford University. When finished, the platform will allow various stakeholders such as real estate developers, banks, investors, and city governments to identify user preferences and purchasing power, and analyze the risk characteristics of income streams that determine real estate value. We find that real estate developers will be able to increase the value of existing real estate and new development, that lenders and financial institutions can improve their risk management through identification of risk characteristics of real estate users, and that new ways of financing real estate are possible when commercial flows within a city can be identified.

A major point of this paper is to illustrate that commercial enterprises are driving digitalization of cities that allows for collection, storage, and analysis of big data. A new way of thinking of infrastructure and service provision is therefore necessary. Examples of commercial investment in urban applications are network connected autonomous vehicles that allow for consumption while driving that are developed by General Motors; ridesharing with Uber; accessing restaurant reviews and making reservations on Yelp; and using Google Maps. Research on digital infrastructure in cities have used concepts such as digital cities', 'intelligent cities', and 'smart cities' when referring to the use of communications technologies and the provision of services and information to populations[1] (Yin et al., 2015). In this article, we choose to exclusively use the term ‘digital city’, because to be digital is more applicable than to be smart, and since digital is more generic than smart it is also a broader based definition for inclusion of commercial market development and R&D. Within the context of urban development, we choose to define a ’digital city’ as: *A city that enables digital infrastructure, creates new insights on commercial markets, that attracts human capital, and delivers means to achieve economic growth, and a higher quality of life through measures taken by private and public stakeholders.*

The commercial applications that we analyze are based on new technologies, with little or no earlier research on their impact on the economy and their commercial applications. The paper that is most like this study is probably Glaeser et al. (2018), that analyzes how big data can improve urban systems, in addition to looking at how Google Street View images can be used to predict income and real estate prices. This article provides a unique addition to earlier literature as we analyze new possibilities provided by crossing of datasets that allows for real-time visualization of activities within a city. Thus, creating a truly digital city.

To date, most research find positive effects of urban infrastructure and ICT technologies on urban systems, with studies typically being done in economics. Findings can be broadly categorized so that the impact of digitalization on the economy stems from two main sources; (1) a general economic impact through new technologies that enable for new functions, improved communications, knowledge spillovers, attracting human capital, and provision of means for increased productivity in existing industries. (2) a direct impact through new technologies and data sources provided by digitalization, that allows for identification of use patterns and demand.

The article is structured so that it starts with a description of the technologies that enable for digitalization of cities and their potential applications. This is followed by a literature review. Lastly, we then identify expected effects of digital cities on various stakeholders such as real estate developers, financial institutions, and local government.

2. Technologies driving change

Collection of large and increasing amounts of data on the urban environment is made possible by things such as sensors, vehicles, smartphones and tablets, applications for consumption, and social media platforms. Storage and analysis of this data is made possible by high-speed mobile networks (such as long-term-evolution networks, LTE, and the advent of the fifth-generation of mobile networks, 5G) and cloud computing, which are essential parts of the internet of things. A good illustration of these technologies is the gps in network connected vehicles, that continually uploads its position to a server that calculates the best route to the set destination and sends this back to the vehicle (Han et al., 2015). Similarly, in addition to identifying location of people in a city,

[1] Clear definitions have yet to emerge, although a ‘smart city’ typically refers to both sustainable behavior and technologies, in contrast to a ‘digital city’ that indicates the use of communications technologies and strategies to supply information and e-services that improve quality and services to inhabitants (Dameri and Cocchia, 2013). Reviews of the definitions of ‘smart cities’ and ‘digital cities’ are provided by Dameri and Cocchia (2013), Cocchia (2014) and Yin et al (2015).

smartphones also allow for information of the mode of transportation (Bierlaire et al., 2013 and Carrel et al., 2015). Applications such as Zillow (real estate listings) and Yelp (restaurant reviews and reservations) provide knowledge on preferences of populations (Glaeser et al., 2018). Data on credit card transactions can tell us where money is spent, on what, and by whom. In summary, it possible to identify commercial and social activity within a city, such as income, consumption and movement patterns of the population.

The underlying trend and driver that makes this possible is Moores's Law, the observation that computing power doubles approximately every 18 months, and correspondingly, the cost per processor would go down by half during the same period. Although typically referring to computing power, Moore's law has come to be pervasive across all technologies used for big data analysis, such as storage technologies, algorithms, artificial intelligence, and advances in material science. This exponential development in performance of multiple technologies is amplified by their convergence, and the advent of cloud computing that offers new economics of scale for storage and processing. This rapid pace of development is illustrated by the fact that just five years ago it was too costly to store and process enormous amounts of data, while it is now possible to collect information of say a million vehicle rides, store the data at low or no cost, and have access to computing power that allows for the application of algorithms that identifies patterns of interest in real time, at a reasonable cost. So, while the data has been available, it is only recently that technologies have allowed for cost effective analysis of this type of data. We define this as a *Quantum Engine™* that drives digitalization, that consists of five components that drive business, on top of which applications are built;

(1) *Computing power*, that provides the brute power to analyze the data and enables for complex visualization. Besides exponential development of microprocessors, a significant enabler for new service applications is the advent of cloud computing, that allows for storage and processing of data without the need for massive physical it-infrastructure, in addition to also making data access and processing more efficient. Consequently, new applications that leverage these advantages have emerged, which in turn has led to a tremendous increase in the amount of data generated (Hashem et al., 2015).

(2) *Virtually unlimited free storage,* that makes it financially feasible to store enormous amounts of data. Storage has been one of the key challenges of big data, as rapidly increasing amounts of information requires new storage technologies. Several new networks based storage technologies have emerged, that provides the scalability that is required by large and rapidly increasing amounts of data. Concentrated computing and storage resources through cloud computing is an essential part of this development, with cloud based storage enabling efficient management of heterogeneous big data (Chen et al., 2014).

(3) *Smart algorithms,* that identifies patterns of interest in the data and makes predictive analytics based on big data possible. The unstructured and heterogeneous nature of big data makes it difficult to extract meaningful information from much of the data being generated, with videos, pictures, tweets, and similar data requiring tools to be analyzed. Computationally efficient algorithms enable for analysis of unstructured and heterogeneous big data, basically, finding the needle in the haystack. Types of such algorithms are Information extraction (IE), that extracts structured information from unstructured documents, such as drug name, dosage, and use from drug prescriptions, and automated speech recognition (ASR) algorithms that match sounds to words so that speech content is transcribed and then indexed. An additional form of data mining is Video Analytics,

which can process both recorded and real-time video. Algorithms can collect demographic information about customers, such as age, gender and ethnicity from CCTV videos which can be used to identify the number of customers, their movement patterns, the time spent in a store, crowds, and queues (Gandomi and Haider, 2015).

(4) *Artificial Intelligence* (AI) and *Machine Intelligence* (MI), that allows for intelligence that can go through an enormous number of scenarios and find new patterns. AI is a broad term for technologies that enables for computers to think and evaluate based on collected data, therefore making machines do tasks in an intelligent manner. An example of AI is autonomous vehicles as it involves the use of intelligent automation, which refers to technologies that actively selects data and processes information, resulting in decisions or controls processes (Lee and See, 2004). Machine Intelligence refers to putting AI to a purpose or task, it is the is the practice of using algorithms to analyze data, learn, and then make decisions or predictions. The machine is therefore trained to perform a specific task using large amounts of data, in contrast to having been given specific instructions.

(5) *Advanced Material Sciences*, is an interdisciplinary field referring to the development of new materials with properties not found in nature, known as metamaterials. The scope of potential applications ranges from improving antenna performance, super lenses, roofing materials that decreases temperature in buildings, to laser technology used for fiber optics. Another type of material development is the advancement of nanomaterials, that makes it possible to build microprocessors that are smaller than 100nm.

2.1 Application of technologies and their use in real estate development

We believe that the type of information that digital cities provide can be broadly divided into four main categories[2];

(1) *Digitization of records,* such as crime, taxes, education, and medical records, allowing for policy evaluation and crossing with other data. One potential use of this type of data is that it provides police enforcement with access to daily records of crime at a fine geography (Braga and Bond, 2008). From a real estate point of view, digitalization of building permits would provide valuable information, as future development can be used both as an indicator of the overall trajectory of a neighborhood and provide insight towards future supply of real estate. A potential source of data that would constitute an extremely useful measure of economic activity is credit card data at a store-level over time. This type of data could be used for policy evaluation, such as estimating the commercial impact of car-free zones. Consequently, value is created when this data is crossed with exogenous events (Glaeser et al., 2018). Credit card data can provide important information on two levels, at the store level, as described above, or when collected at the level of individual credit card holders, so that spending habits and preferences of various demographics can be identified.

[2] Our classification is somewhat different from that by Glaeser et al. (2018) that divide big data sources on urban systems into three main categories; digital exhaust of internet services (say TripAdvisor or Yelp), open government records (such as on crime), and corporate data (such as data on gym memberships). Although similar, the most notable differences in categorization is that we categorize corporate and public records in the same category, and choose to focus on the importance of sensor data and movement (that provide valuable insight when crossing data sources), adding this as a separate category.

(2) *Application provided information on user preferences,* from platforms for consumption such as Amazon, real estate listings from Zillow, restaurant reviews and reservations on Yelp, or social media postings on Facebook. Access to this type of data allows for identification of customer preferences, say what kind of real estate a certain demographic is most likely to want, or what kind of products they buy online. When linked to geographies, this can provide useful information when determining real estate value.

(3) *Sensor information on the urban environment,* provided by sensors that can sense things such as the number of people at a location, noise, and pollution, all of which can provide valuable insight in relation to real estate value. Identification of movement, crowds, and gatherings will tell us what kind of activities are taking place in a neighborhood, such as if people are shopping, eating and drinking at outdoor restaurants, or walking in a certain direction. The impact on commerce and flow of people caused by various types of development and tenant mix can also be analyzed. Linking pollution and noise to real estate is an easy use of this kind of data that can provide valuable insight for real estate developers. In addition to sensors on buildings and infrastructure, smartphones enable for tracking of environments, such as noise and weather (Han et al., 2015).

(4) *Sensor Information on movement*, provided by network connected cars, public transport, and smartphones. Although some privacy issues do arise, sensors on smartphones enable for tracking of individuals within a city (Han et al., 2015). Knowledge about movement provides valuable information when crossing datasets, so that patterns of consumption and social activity can be broken down by demographics. Identification of movement patterns in a city, such as knowing that a certain demographic tends to work at location A, and moves to retail and restaurants in location B at a certain time of day, is useful in understanding where to locate new development, and analyze how development impacts transportation. This type of data can also be used to identify trends, as it will be possible to see when millennials are moving to 'cooler' areas, in real time.

A good example of how sensor data can be used to improve the urban system is when such information is used to improve public transport, as is done in Seoul, South Korea. Knowledge about the number of passengers by location and time allows for better allocation of resources and real-time provision of estimated travel times so that travelers can make more informed decisions on how to get around. Less delays and shorter commutes will positively impact the economy, as it makes it easier to reach locations of work and consumption, in addition to attracting households and firms.

From a commercial perspective, valuable insights are provided by the fact that demand for various products and services varies by age, with elderly requiring more health care and younger age groups tending to demand more restaurants and bars. Identification of the location and movement of demographics allows for prediction of the type of customer that will buy a certain service, and real time estimates of how real estate values change when say the number of 25 to 34 year olds increases with a certain number within a set geography. This allows various stakeholders of real estate to estimate risks associated with real estate values, make prognoses on impacts on values (such as when a neighborhood gentrifies), and to identify what kind of development that maximizes value of new, and already existing surrounding properties. The data could be used so that a certain type of new development (residential, office, or retail) is determined by local demand (such as information provided from exhaust from applications), or so that development is aimed at attracting demographics that are found to increase surrounding real estate values.

Another example that illustrates the application and use of new technologies and big data is the development of indoor vertical farming that is currently being developed[3]. This allows for 35 acres worth of produce being farmed on less than 10 000 square feet where no soil is required and light is controlled by computers, determining the need for water, and when to harvest (by robots). Converging this production technology with data layers on the urban environment, a potential use is that demand can be estimated at a very local geography. In summary, the amount of say tomatoes that needs to be grown, and how much that should be sent to stores in various neighborhoods. This allows for more local production and minimizes waste in both supply chain and stores. In general, production technology coupled with data on population characteristics (on say income, age, and consumption) allows for estimation of local demand and to set corresponding production and local stock with a high degree of precision.

In term of impact on the economy and the urban system, we believe that the impact of digitization can be thought of in term of two causal relationships;

(1) A general impact of new services and applications that increases economic output through improved communications, data sharing, and knowledge transfer. This improves productivity in existing industries, attracts human capital, and positively impacts firm location decisions. Examples are that Uber and Lyft provides means of transportation, and Zillow makes real estate markets function more efficiently, as market knowledge becomes more pervasive among individuals. This type of impact is general in nature and therefore difficult to quantify, and stems directly from improved communication, and the provision of new services.

(2) A direct impact through new technologies, data sources, and business applications. Exhaust of data from tax records, social media usage, and online purchasing patterns leads to a better understanding of preferences, use patterns, and demand.

3. Literature review

As field of research is new, there are few directly applicable earlier studies on how digitalization impacts urban environments. We therefore analyze two broad strands of literature;

Section 3.1 covers studies that relate to a general economic impact of digitalization, typically by looking at the impact of digital infrastructure and human capital, both of which are likely proxies for the overall level of digitalization. This category of research can be thought of as analyzing a more general impact on the economy and urban system, as new technologies facilitate knowledge spillovers and network externalities.

Section 3.2 covers the narrower field of big data sources, technologies and applications. The impact on the economy and urban system is through business applications of new technologies, and analysis of collected data.

3.1 Literature on digital infrastructure, agglomeration, and knowledge spillovers

[3] See for instance http://onepointone.us

Studies on the economic impact of improved communications

On the most general of levels, it is safe to assume that new means of data collection, sharing, and communication positively impacts the economy. As an illustration, the consulting firm McKinsey estimates that the internet accounted for 3.4% of GDP, and 21 percent of GDP growth in mature economies during 2005 to 2010 (Manyika and Roxburgh, 2011). Colecchia and Schreyer (2002) examines the impact of ICT technologies across nine OECD countries, finding that such technologies contributed to between 0.2 to 0.5 percentage-points in economic growth per year. A review of the literature on how ICT impacts productivity is provided by Cardona et al. (2013), finding that most studies show a positive and significant effect. The authors state that more research is needed on spillovers and externalities of ICT technologies.

Improved communications technologies positively benefits markets (Jensen, 2007; Aker, 2010; Aker and Mbiti, 2010). The effect is often tested in developing countries, with Jensen (2007) finding that mobile phone usage among fishers in India was associated with an increase in market efficiency by dramatically reducing in price dispersion and eliminating waste. Similarly, Aker (2010) found a positive impact on grain markets in Niger.

Services provides another a good way of illustrating the effects of ICT (although limited in the context of this paper), as they typically are consumed locally and require physical interaction. The internet provides a medium of exchange that has the potential to overcome the need to physically meet to transact. Freund and Weinhold (2002) estimates that the internet has had a significant impact on world trade in services, with a 10-percent increase in internet usage being associated with a 1.7 percentage-point increase in export growth and a 1.1-percentage-point increase in import growth. Analyzing trade in general, Freund and Weinhold (2004) find similar results. A 10 percentage-point increase in the growth of internet usage leads to a 0.2 percentage-point increase in exports. The authors find that the internet reduces market-specific fixed costs of trade. Although not affecting the relationship between distance and trade, the internet enhances local competition so that the overall effect of distance increases.

Studies on the economic impact of broadband access and usage

A large body of research that have analyzed the linkage between high-speed internet and local economic performance (Lehr et al., 2006; Gillett et al., 2006; Mack and Grubesic, 2009; Crandall et al., 2007; Koutroumpis, 2009; Holt and Jamison, 2009; Kandilov and Renkow, 2010; Mack et al., 2011; Qiang et al., 2009; Stenberg et al., 2009; Czernich et al., 2011; Kolko, 2012; Rohman and Bohlin, 2012; Kim and Orazem, 2012; Mack and Faggian, 2013; Whitacre et al., 2014; Mack and Rey 2014; Mack, 2014). Most of these studies have found that increased broadband capacity, or access, has a positive causal effect on productivity.

Precise estimates of how ICT technologies impacts the economy are hard to estimate as it is closely related to levels of human capital (Holt and Jamison 2009). The relationship might be such so that an already present highly-educated workforce benefits more from ICT technologies, as shown by several studies having found varying impact across industries (Lehr et al., 2006; Gillett et al., 2006; Crandall et al., 2007; Kolko, 2012). Alternatively, that the provision of such technologies might impact firm location and hiring, implying that ICT technologies causes a rise in human capital. Mack et al. (2011) finds that broadband is an important determinant of location for knowledge-intensive firms. Kim and

Orazem (2012) found that broadband will cause economic growth through an impact on productivity on current industries, in addition to positively impact firm entry and start-ups in rural areas. However, Mack and Grubesic (2009) did not find any relationship between broadband and firm location using data from Ohio. Similarly, Kandilov and Renkow (2010) found no impact on the number of businesses due to broadband access programs.

Mack (2014) highlights that the impact of broadband provision is highly heterogeneous and localized. It is found that broadband infrastructure can be essential to enable knowledge intensive firms to locate in places with lower costs and near a concentration of knowledge. However, broadband does not mitigate the negative externalities associated with remote locations. The author states that broadband should be viewed as a key component of economic development, but not the only component. Similarly, Mack and Rey (2014) find that broadband is of importance to knowledge intensive firms when analyzing 54 metropolitan areas across the U.S.

Studies on the economics of human capital and agglomeration effects

Adding to the relationship between digital infrastructure and human capital are agglomeration effects. A large body of research supports a positive relationship between city size and growth in employment and population (Glaeser et al., 1995; Simon, 1998; Simon and Nardinelli, 2002; Simon, 2004; Glaeser and Shapiro, 2003; Shapiro, 2006). Studies also support a positive (although varying in magnitude) relationship between city size and wages (Di Addario and Patacchini, 2008; Wheeler, 2006; Glaeser and Maré, 2001; Wheeler, 2001). A positive effect on wages is supported by a 33% wage premium in large metropolitan areas compared to non-urban locations in the U.S. (Glaeser and Maré, 2001). Several studies provide empirical results that support that the urban wage premia is not solely due to a concentration of workers with higher ability (Rosenthal and Strange, 2008; Glaeser and Maré, 2001; Wheeler, 2006; Glaeser and Gottlieb, 2009). This effect is typically attributed to increasing levels of transference of knowledge, so called 'knowledge spillovers': "*Even if cities are no better educated than the hinterland, urban density will increase interactions and intellectual spillovers.*" (Glaeser and Maré 2001). A deduction from this line of research is that given the benefits of physical connectedness, one would expect that technologies supporting communications and data sharing would bring similar economic benefits.

Agglomeration effects are amplified by human capital. Rauch (1993) makes the argument that the average level of human capital is a local public good. The author supports this argument by finding that cities with higher levels of human capital across the U.S. have higher wages and land values. The theoretical reasons for growth due to spillover effects of knowledge varies. Such effects can occur mainly across industries as suggested by Glaeser (1992), which implies that diversity of employment sectors is of importance (Jacobian diversification externalities). Alternatively, industry concentration might be of greater importance if knowledge spillovers occur mostly within industries (Marshallian specialization externalities). This is consistent with the findings by Henderson et al. (1995), that density increases growth in mature capital-intensive sectors.

Agglomeration economies might benefit households differently, with Wheeler (2001) finding that the wage premium increases with education while Adamson et al (2004) and Lee (2010) find the opposite, meaning that less skilled workers benefit more from agglomeration. Although increasing wages in general, Di Addario and Patacchini (2008) find that urbanization has no impact on individual returns on experience and decreases returns to education. Wheeler (2006) finds that larger cities increases wage growth, this effect is however not only attributed to an increased rate of acquiring skills. It is found that much of this wage growth is due to job changes, meaning that urbanization improves the search and matching process on the job market. Shapiro (2006) also analyzes the effects of human capital on rents and house prices, finding stronger effects than on wages. The author also runs a model that estimates that 60% of the effect of human capital on employment growth is due to increased productivity, with the remaining effect being a consequence of a relationship between concentrations of human capital and increased quality of life.

The importance of physical connectivity between individuals is illustrated by the fact that knowledge spillovers that impacts wages and innovation are found to decline sharply with geographical distance (Rosenthal and Strange, 2008; Rosenthal and Strange, 2001; Andersson et al., 2009). A large body of research supports that space itself constitutes a barrier to the diffusion of knowledge (Krugman 1991; Feldman 1994; Jaffe et al 1993; Audretsch and Feldman 1996).

Although the causal relationships between agglomeration economies and human capital are difficult to disentangle[4], the importance of attracting educated households is illustrated by a positive relationship between wages, population and educational attainment. This trend is also amplifying as the share of adults with university degrees have increased more rapidly in US cities having started with a high level of educational attainment during the last 30 years (Berry and Glaeser 2005; Moretti 2004a). Consequently, Glaeser and Berry (2006) propose that cities adapt policies for attracting highly educated households. Similarly, Florida (2002) stresses the importance of 'soft infrastructure' such as crime-free environments, entertainment, knowledge networks, and volunteer organizations which in combination with high-tech and creative industries drive urban growth. The premise is that these industries are a growing part of the economy, forcing cities to attract creative people to succeed, or as stated in a review of Richard Florida's book 'The Rise of the Create Class' by Glaeser (2004): *"Florida's basic thesis is that the economy is transforming, and creativity is to the 21st century what the ability to push a plow was to the 18th century"*.

Studies on the economic impact of knowledge intensive industries and universities

Closely related to levels of human capital, and a possible proxy for digitalization is the presence of knowledge intensive industries or universities. Knowledge spillovers from universities are typically found to promote firm innovation (Stuart and Sorenson 2003; Hall et al 2003). Andersson et al (2005) found that patenting activity increases in larger and denser labor markets when examining Swedish data. Andersson et al (2009) analyzes the impact of decentralization of universities in Sweden, a policy change that provides a natural experiment. The authors find positive effects on productivity and innovation. Interestingly, aggregate productivity also increased because of decentralization policies. Effects are however highly localized, with spillovers rapidly decreasing with geography. The sharp attenuation of spillovers is like other studies of agglomeration, such as Rosenthal and Strange (2008).

[4] This issue primarily stems from human capital not being distributed randomly, as workers tend to locate based on preferences on wages, cost of living and amenities ((studies supporting that people with high human capital seek to live in places with a high quality of life are provided by Kahn (2000) and Cullen and Levitt (1999)), and firms correspondingly choosing location based on wages, cost of land and matching of costs with city characteristics (Moretti, 2004a). Several studies mention that more research is needed to disentangle the effect of human capital on wages, productivity and land prices from that of unobserved characteristics of workers and cities (Moretti, 2004a 2004b; Combes et al., 2008).

Andersson et al (2009) estimates that half of productivity gains were found to be within 5-8 kilometers from the community in which higher education investments were made. Audretsch et al (2005) states that spillover effects are likely to be heterogeneous and dependent on the context. Examining high tech start-up location choices in Germany, the authors find that such firms locate near universities to access knowledge spillovers. Acs et al (2002) found that both university research and industry R&D a positive effect on patent activity, with R&D activity having a substantially stronger positive impact. Florax (1992) did however not find that proximity to a university had a significant impact on the number of startup companies and their location.

Several studies confirm that technological knowledge is a principal input to R&D activity, as knowledge spillovers positively impacts firm productivity (Autant-Bernard et al 2013). A review of studies that examine spillovers of R&D activity is provided by Sena (2004).

Theory on human capital and house price growth

The traditional analysis of urban growth would assume that house price growth is a consequence of increased productivity. Consequently, digitalization would positively impact real estate values through a general economic impact. A study that offers an alternative framework to explain house price growth, that directly relates to human capital, is provided by Gyourko et al. (2013). The authors analyze the causes for the increased dispersion of long-run house price appreciation rates across U.S. metropolitan areas. As an example, San Francisco has experienced an average annualized rate of appreciation of 3.5% between 1950 and 2000, while the equivalent number for Buffalo is 0.5%. By 2000, house prices in the most expensive cities where four times the US average, compared to having been just twice as expensive in 1950. The authors show that given constant preferences for location among households, and housing supply not being perfectly elastic everywhere, changes in aggregate housing demand will impact cities differently so that local house price growth leads to a change in local population composition. Basically, aggregate demand will raise prices more in so called 'superstar' cities with excess demand and low supply elasticity, which in turn impacts the population composition when lower income potential residents are being crowded out. The authors state that as much two-thirds of the increase in house price dispersion, and almost all the increase in income dispersion, can be explained by the increase in high-earning households on the national level (Gyourko et al, 2013). Given that this effect, cities that can create excess demand are going to experience increasing real estate values. Policy towards attracting human capital, the provisions digital infrastructure and R&D will therefore directly impact real estate values.

In summary, the above-mentioned research tells us that an increased level of digital infrastructure is closely related to human capital, agglomeration economies and knowledge spillovers. In most cases, more developed digital infrastructure results in higher economic growth because of corresponding developments in human capital, knowledge spillover, and agglomeration effects.

3.2 Big data sources, technologies and applications

Studies on the business of big data the internet of things (IoT)

The volume and variety of data from mobile users and communications networks have been increasing exponentially since the early 2000s (Andrews et al., 2014; Musolesi 2014). Part of this development is what is known as the ‘internet of things’ (IoT), allowing for items such as vehicles and mobile devices to collect and exchange data (Atzori et al., 2010; Zanella and Vangelista, 2014). Although the benefits of universal computing (i.e. collection and application of knowledge about an environment and its inhabitants) was identified long ago, it is recent innovations in information and communications technologies (ICT) that has made large scale deployment of such systems possible (Yovanof and Hazapis 2009). The impact of big data on the economy is illustrated by it having led to the creation of entirely new business models (Hartman et al., 2016). In the retail sector, the ability to analyze buyer behavior in near-real time allows for adjustment of products, stocking, and prices (Hagen et al., 2013). A review of business models based on big data is provided by Hartman et al. (2016), that also proposes a framework for evaluation such business models. The authors state that despite the need to understand how to capture value from data, there has been surprisingly little research on the topic. However, it is known that commercialization of new technologies and innovation is typically done by start-up companies (Criscuolo et al., 2012).

Studies on 'digital' and 'smart' cities

Collection and exchange of data within the urban environment is closely related to what is often called a 'smart' city, that uses advanced communications technologies to provide services to its inhabitants (Zanella and Vangelista 2014; Hashem et al., 2016). Concepts such as 'digital cities', 'intelligent cities', and 'smart cities' have similar, although different meanings (Yin et al., 2015). Clear definitions have yet to emerge, although a 'smart city' typically refers to both sustainable behaviour and technologies, in contrast to a 'digital city' that indicates the use of communications technologies and strategies to supply information and e-services that improve quality and services to inhabitants (Dameri and Cocchia, 2013).

Zanella and Vangelista (2014) provides a review of the enabling technologies for smart cities, stating that they have reached a level of maturity that allows for realization of IoT solutions and services. Similarly, Hashem et al. (2016) analyzes communications technologies in the context of smart cities and how big data improves sustainability and quality of life. The authors also propose a business model for big data, as data analytics enables for prediction of market trends with increasing precision and decision automation. Additional examples of the use of big data is being able to strategically place advertising and gain understanding about customers and products, which can be used to identify opportunities and risks for a company. It is also possible analyze customer complaints to see revenue contributions of products, and framing of hypotheses based on big data, which can be tested by experiments (Hashem et al., 2016).

Studies on big data collection and use

Although possible with current technologies, processing of large volumes of heterogeneous data that is multi-sourced and provided in real time poses a technological challenge, requiring scalable data storage infrastructure, and scalable performance (Zheng et al., 2016). The need for common data collection standards is stressed by many studies (Andrews et al., 2014; Bellavista et al., 2013). Literature on the field of big data and the challenges associated with data acquisition, storage and analysis is reviewed by Chong and Shi (2015), noting that new tools and methods will be required to efficiently extract and analyze information from large data. The question of what the fifth-generation of mobile networks (5G) will constitute is analyzed by Andrews et al (2014), noting that the explosion of mobile data will require network data capacities to increase by roughly 1000 times compared to 4G. Han et al. (2015) makes the case that 5G and improved mobile cloud technologies will provide new opportunities to take advantage of sensor data, which is central to the 'intelligent' revolution currently taking place.

A study that focuses on technologies for urban data collection is provided by Bellavista et al (2013). The authors find that so called wireless sensor networks (WSNs) that consist of small and inexpensive sensors that can take measurements, store and handle sensed data, and communicate to each other, can provide new opportunities for wide-scale monitoring of cities. An example that is relevant in the context of urban development are sensors that enable cities to sense and manage things such as water supply and transport networks (Kopetz, 2011).

Another technology that provides opportunities for urban monitoring are mobile ad hoc networks (MANETs) that enable connections between devices without predefined fixed infrastructures (Bellavista et al., 2013). Examples of such network are communication between vehicles (e.g. allowing for intelligent vehicle behavior during accidents) (Martinez et al., 2010), and mobile devices (e.g. allowing for communication without network access) (Conti and Giordano 2014).

Han et al. (2015) take smartphones as an example of a potential source of data, as they are already equipped with a rich set of sensors that can provide context information such as noise and weather. Smartphones also enable for tracking of individuals within a city and continuously upload locational information to central servers (Bierlaire et al., 2013 and Carrel et al., 2015). It is also possible to identify the mode of transportation through algorithms (often based on acceleration), such as walking, riding a bus, or train (Feng and Timmermans, 2013; Hemminki et al., 2013; Sankaran et al., 2014). Cloud computing allows for collection and analysis of this type of data without the constraints of the memory and computing resources of mobile devices (Han et al., 2015). An example of an application of this type of technology is a gps that continuously uploads the position of a vehicle and instantly returning the estimated best route (Han et al., 2015).

Thiagarajan et al. (2010) analyzed how smartphones could be used to track public transit usage in Chicago. Their simulations found that this kind of tracking could improve overall transit efficiency, with expected wait times being 2 minutes shorter if only 5% of transit users participated in the system that allowed for tracking. At a 20% level of penetration, the average wait time went from 9 to 3 minutes. A similar study by Zhou et al. (2012) analyzed prediction of bus arrival times based on tracking through smartphones. Knowing if a bus is going to be late based on information on other travelers, traffic, and accidents allow people to make more informed decisions on how to move within a city. The authors found that this type of system for prediction provides outstanding accuracy compared to current systems. Similarly, Gu et al. (2017) analyses measurement of interchange times in subway systems using smartphone sensors, allowing for identification of crowdedness of various stations, usage, and efficiency.

Tracking of individuals allows for more closely tailored products and pricing as illustrated of so called insurance telematics, where road vehicle information is collected and used for insurance purposes (Bruneteau, 2012). This type of data allows for individual risk assessments and premiums based on driving behavior. This type of monitoring can be done through fixed sensors, a semi-fixed device or smartphones (Händel et al., 2013). An example of how smartphones can be used is the *streetbump* app that detect potholes in roads through detection of shocks, this information is then uploaded to street repair services in Boston (Glaeser et al., 2018).

A review and analysis of how big data can improve urban systems is provided by Glaeser et al. (2018). A categorization of data sources and uses is provided, and examples of how big data can be gathered and analyzed. The authors provide examples of various uses of big data, with exhaust of user data from social media such as Facebook providing information on the 'pulse' of an area and structures of social networks; LinkedIn increasing knowledge on local labor markets; review platforms such as Yelp giving information on the quality of services; search engine data that sheds light on the needs and preferences of the population; and information on housing markets through services such as Zillow or Airbnb. The impact of digitalization of government records and corporate data is also reviewed. An example of the former being that more easily available crime data supports law enforcement, while the latter could be that gym memberships provide information on health or data on credit card transactions which allow for quantification of changes in spending. From a policy perspective, the authors give examples of value from the crossing of digital records with exogenous events, such as estimating the impact of a hotel tax through Airbnb listings and TripAdvisor reviews. Lastly, an example of how income and housing prices can be predicted with machine learning applied to Google Street View images is provided. Consequently, street view imaging provides a tool to better understand patterns of wealth, poverty, and segregation. If data is available over time, the effects of policy interventions can be estimated.

In the context of street view imaging, the findings of Naik et al. (2014) are also of interest, as the authors found that such data can be used measure the physical attributes of an area. Naik et al. (2015) applies an algorithm based on survey data to Google Street View images to create maps of perceived safety for 21 U.S. cities. This allows for analysis of the linkage between physical attributes of neighborhoods, with social attributes such as health and behavior. It is found that population density and share of college-educated adults are strong predictors of later improvements in the perceived safety of an area.

4. Discussion on the potential effects of digitalization on real estate development.

Based on a review of the literature on technologies, big data sources, and their applications, we find that opportunities for value creation through big data in the context of real estate can be divided into four main categories; (1) forecasting of user preferences and usage of real estate; (2) flexible utilization of real estate, (3) information on user purchasing power; and (4) identification of risks. These opportunities are elaborated upon below;

1. *Forecasting of user preferences and usage of real estate*. The capacity to gather and analyze data on user preferences will have a profound impact on real estate development and management, similarly to how retailers currently use such data to forecast market trends and analysis of products contribution to revenue (Hashem et al., 2016).

   As real estate is a long lived, heterogeneous asset that requires large capital outlays, forecasting abilities will be of even greater use. Being able to identify usage patterns and preferences before development provides opportunity for increasing real estate value through higher rents or prices, and lower vacancies. Similarly, identification of trends such as gentrification through real time data on movement and consumption through smartphones, credit card data, or online applications will provide valuable information to determine the location, and type of future development that will bring most value.

For existing real estate, this type of information can provide savings on costly tenant adjustments as the product can be more closely matched with customer preferences. Identification of behavior and socio-economic characteristics of people on a regional and local level will also provide insight into what kind of new urban development will bring most value to real estate users and the developers. For instance, analysis of multiple layers of data will enable better prediction of the proportions between residential, retail and office space in a high-rise building. Such an analysis might be based on identification of a certain type of demand through exhaust of Zillow searches on housing, restaurant reservations through Yelp, or identification of adjacent office usage through sensors. Movement data would tell us if the likely demographic for new office will want to have retail and restaurant locations nearby, or if new retail will increase office and residential values. In summary, the end goal of each analytical model is to deduce the composition of the real estate from the value it creates for end users, because that is the ultimate determinant of the value of the real estate.

This type of analysis can also provide valuable insight for other stakeholders, such as banks, when evaluating financing solutions for real estate. This as future cash-flows can be increased, and predicted more accurately. It will also be possible to predict future demand for housing by various demographics, and therefore also demand for financial services such as mortgages.

From an infrastructure perspective, this type of data analysis will also prove useful. As an example, knowing if future population growth will travel by car or public transport will support decisions on transport infrastructure. Knowledge on age and household sizes as well, so that usage of kindergartens and schools can be forecasted. Better abilities to track changes of activity such as movement, spending, and choice of transportation mode over time can provide valuable insight for policy evaluation.

2. *Flexible utilization of real estate*. Digital cities make it easier to change real estate, or to encourage changes in user utilization of the existing real estate. Digital models of real estate can be made by laser scanning during construction phases, thereby capturing all layers of construction. Once the building is occupied, utilization can be studied from a combination of sensors, phone, and gps data, and commercial transactions. Analysis of real estate usage can give reason for changing the real estate, or for changing user behavior. Detailed, rich and accurate information on buildings will make it less costly and time consuming to change the building to facilitate user needs. The tenant can be shown accurate digital representations of the real estate, and, together with digital representation of tenant user behavior, this can provide a good platform for a value adding discussions between the real estate operator and tenant, perhaps resulting in changed real estate, increased rents, and/or changed tenant utilization of the building.

   Rapid technological advances are made in sustainable urban systems, and these advances also make changes to real estate potentially very value adding. For instance, it is becoming more popular to have vertical gardens and vegetable farms in real estate. Making a digital representation of a building can make it easier to adapt to new urban real estate trends, effectively making the real estate usage modular.

3. *Identification of user purchaser power*. Knowledge about income and purchasing patterns of individuals on a local level can provide valuable insight for real estate professionals in determining rents and prices. Identification of customer purchasing power has many potential uses; (a) a tool for improved negotiating power, (b) allowing real estate developers to set price points more accurately and increase price differentiation after identification of local levels of purchasing power. Real estate managers and developers will be able to be better at accurately target different types of customers, e.g. at high-end or lower-end price-points.

4. *Identification of risks*. Digital city data can be used to assess the risk in building structure malfunction, for instance due to earthquake damage. Risk also emanate from users, because of factors that influence user demand. Digital data on user groups such as tenants and end customers, at retail locations, restaurants, offices, and residential properties provide information about individual risk characteristics that in turn also can be used to identify risks of commercial assets within a city. This type of information will be useful for many stakeholders, such as real estate investors, lenders, and city government. Of interest is data on commercial transactions in and around a real estate, because it can be combined with other layers of socio-economic data to better predict tenant and end customer risks. For instance, for a real estate owner, information on the retail tenant's customers may prove valuable for prediction of risks. This type of information can be used to evaluate real estate risk from the perspective of volatility in rental income, that consequently determines real estate value.

   This type of big data analysis can vastly improve pricing and risk management for financial institutions by better estimates of the cost of capital associated with financial products with real estate as the underlying asset. Identifying the relationship between movement of people, retail consumption, rents, and real estate value allows for a real estate asset's value elasticity in relation to retail spending, or the presence of a certain demographic to be estimated. This allows banks to more accurately price risk and allocate capital reserves, consequently improving market efficiency, pricing, profitability, and risk management.

   Credit risk can also be improved on the level of individual persons. Traditionally, credit risk is analyzed through models based on income and assets, such as mortgage default being modelled so that default will occur when a borrower has sufficiently negative equity to make strategic default worthwhile (Deng et al., 2000; Ambrose et al., 2001). Most households with negative equity do however not default (Foote et al., 2008). Being able to identify behavioral aspects of credit risk by crossing data on demographic characteristics, location, credit default, savings, and spending behavior, can provide valuable for management of risk by various stakeholders.

   Lastly, this new way of thinking of cities, with real time measurement and visualization of commercial flows within a city allows for new financial products. An example would be an asset backed security with retail transactions within a certain geography as the underlying asset. This provides a new way of financing new development, based on predictions of surrounding commercial flows that determine value of and risk of the financial asset.

The above categories of how digitization can increase value of real estate and provide tools for analysis of varying complexity. Most notably, collection and analysis of data can be done at one point in time, i.e. statically, or, continuously which provides a dynamic model. The former could take the form of an analysis of correlations within a database of converged data. In addition to providing insight through correlations, the latter will also provide valuable information by identifying changes in relationships. Being able to measure changes in economic activity, usage of infrastructure, or travel, on a local level can provide a setting to test of business plans for commercial actors (e.g. evaluation of marketing campaigns), and policy evaluation for city governments (e.g. testing impacts of changes in taxes, public transport or regulation).

In summary, this new way of thinking about digital cities will change how stakeholders in the urban environment do business and will lead to the questioning of prevailing theories and truths. The possibility to identify commercial and social activity of populations will be closely related to many of society's pressing issues. From a perspective of urban development, better forecasting and increased connectedness can provide new solutions that address questions such as housing shortages, crowdedness, and housing affordability.

References

Acs, Z., Anselin, L., Varga, A., 2002. Patents and innovation counts as measures of regional production of new knowledge. Research Policy 31 (7), 1069–1085.

Adamson, D.W., Clark, D.E., Adamson, M.D., 2004. Do urban agglomeration effects and household amenities have a skill bias. Journal of Regional Science 44, 201–223.

Aker, Jenny C., and Isaac M. Mbiti. "Mobile phones and economic development in Africa." The Journal of Economic Perspectives 24.3 (2010): 207-232.

Ambrose, Brent W., Charles A. Capone, and Yongheng Deng. "Optimal put exercise: an empirical examination of conditions for mortgage foreclosure." *The Journal of Real Estate Finance and Economics* 23.2 (2001): 213-234.

Andersson, Roland, John M. Quigley, and Mats Wilhelmsson. "Urbanization, productivity, and innovation: Evidence from investment in higher education." Journal of Urban Economics 66.1 (2009): 2-15.

Andersson, Roland, John M. Quigley, and Mats Wilhelmsson. "Agglomeration and the spatial distribution of creativity." Papers in Regional Science 84.3 (2005): 445-464. Andrews, Jeffrey G., et al. "What will 5G be?." IEEE Journal on selected areas in communications 32.6 (2014): 1065-1082.

Atzori, Luigi, Antonio Iera, and Giacomo Morabito. "The internet of things: A survey." Computer networks 54.15 (2010): 2787-2805.

Audretsch, D., Feldman, M., 1996. R&D spillovers and the geography of innovation and production. American Economic Review 86 (3), 630–640.

Audretsch, David B., Erik E. Lehmann, and Susanne Warning. "University spillovers and new firm location." Research policy 34.7 (2005): 1113-1122.

Autant-Bernard, Corinne, Muriel Fadairo, and Nadine Massard. "Knowledge diffusion and innovation policies within the European regions: Challenges based on recent empirical evidence." Research Policy 42.1 (2013): 196-210.

Berry, Christopher R., and Edward L. Glaeser. "The divergence of human capital levels across cities." Papers in regional science 84.3 (2005): 407-444.

Bellavista, Paolo, et al. "Convergence of MANET and WSN in IoT urban scenarios." IEEE Sensors Journal 13.10 (2013): 3558-3567.

Bruneteau, F., “Why insurance telematics matters – Overview of a future EUR 50 billion market,” in *Telematics Munich*, 29-30 October, Munich, Germany, 2012.

Bierlaire, M., Chen, J., Newman, J., 2013. A probabilistic map matching method for smartphone gps data. Transport. Res. Part C: Emerg. Technol. 26, 78–98.

Boccardi, Federico, et al. "Five disruptive technology directions for 5G." IEEE Communications Magazine 52.2 (2014): 74-80.

Carrel, A., Lau, P.S., Mishalani, R.G., Sengupta, R., Walker, J.L., 2015. Quantifying transit travel experiences from the users perspective with high-resolution smartphone and vehicle location data: methodologies, validation, and example analyses. Transport. Res. Part C: Emerg. Technol. 58, 224–239.

Caragliu, Andrea, Chiara Del Bo, and Peter Nijkamp. "Smart cities in Europe." Journal of urban technology 18.2 (2011): 65-82.

Cardona, Melisande, Tobias Kretschmer, and Thomas Strobel. "ICT and productivity: conclusions from the empirical literature." Information Economics and Policy 25.3 (2013): 109-125.

Chen, Min, Shiwen Mao, and Yunhao Liu. "Big data: A survey." *Mobile Networks and Applications* 19.2 (2014): 171-209.

Chong, Dazhi, and Hui Shi. "Big data analytics: a literature review." Journal of Management Analytics 2.3 (2015): 175-201.

Cocchia, Annalisa. "Smart and digital city: A systematic literature review." Smart city. Springer International Publishing, 2014. 13-43.

Colecchia, Alessandra, and Paul Schreyer. "ICT investment and economic growth in the 1990s: is the United States a unique case?: a comparative study of nine OECD countries." Review of Economic Dynamics 5.2 (2002): 408-442.

Conti, Marco, and Silvia Giordano. "Mobile ad hoc networking: milestones, challenges, and new research directions." IEEE Communications Magazine 52.1 (2014): 85-96.

Combes, Pierre-Philippe, Gilles Duranton, and Laurent Gobillon. "Spatial wage disparities: Sorting matters!." Journal of Urban Economics 63.2 (2008): 723-742.
Crandall, R., Lehr, W., & Litan, R. (2007). The effects of broadband deployment on output and employment: A cross-sectional analysis of US data. Issues in economic policy, 6(July), 1–34
Criscuolo, P., Nicolaou, N. and Salter, A. (2012) 'The elixir (or burden) of youth? Exploring differences in innovation between start-ups and established firms', Research Policy, Vol. 41 No. 2, pp. 319-333.
Cullen, Julie Berry, and Steven D. Levitt. "Crime, urban flight, and the consequences for cities." (1999): 159-169.
Czernich, Nina, et al. "Broadband infrastructure and economic growth." The Economic Journal 121.552 (2011): 505-532.
Dameri, R. P., and A. Cocchia. "Smart city and digital city: Twenty years of terminology evolution." X Conference of the Italian Chapter of AIS, ITAIS. 2013.
Deng, Yongheng, John M. Quigley, and Robert Order. "Mortgage terminations, heterogeneity and the exercise of mortgage options." *Econometrica* 68.2 (2000): 275-307.
Di Addario, Sabrina, and Eleonora Patacchini. "Wages and the city. Evidence from Italy." Labour Economics 15.5 (2008): 1040-1061.
Feldman, M., 1994. The Geography of Innovation. Kluwer Academic Publishers, Boston, MA. Feng, T., Timmermans, H.J., 2013. Transportation mode recognition using gps and accelerometer data. Transport. Res. Part C: Emerg. Technol. 37, 118–130.
Freund, Caroline, and Diana Weinhold. "The Internet and international trade in services." The American Economic Review 92.2 (2002): 236-240.
Freund, Caroline L., and Diana Weinhold. "The effect of the Internet on international trade." Journal of international economics 62.1 (2004): 171-189.
Florax, R., 1992. The University: A Regional Booster? Aldershot, Avebury. Florida, Richard. "The economic geography of talent." Annals of the Association of American geographers 92.4 (2002): 743-755.
Foote, Christopher L., Kristopher Gerardi, and Paul S. Willen. "Negative equity and foreclosure: Theory and evidence." *Journal of Urban Economics* 64.2 (2008): 234-245.
Gandomi, Amir, and Murtaza Haider. "Beyond the hype: Big data concepts, methods, and analytics." International Journal of Information Management 35.2 (2015): 137-144.
Gillett, S., Lehr, W., Osorio, C., & Sirbu M. (2006). Measuring the economic impact of broadband deployment. Economic Development Administration, Department of Commerce.
Glaeser, Edward L., et al. "Growth in cities." Journal of political economy 100.6 (1992): 1126-1152. Glaeser, Edward L., and David C. Mare. "Cities and skills." Journal of labor economics 19.2 (2001): 316-342.
Glaeser, Edward L. "Book Review of Richard Florida's" The Rise of the Creative Class"." (2004). Glaeser, Edward L., and Christopher R. Berry. "Why are smart places getting smarter." Taubman Center Policy Brief 2 (2006).
Glaeser, Edward L., and Joshua D. Gottlieb. "The wealth of cities: Agglomeration economies and spatial equilibrium in the United States." Journal of Economic Literature 47.4 (2009): 983-1028.
Glaeser, Edward L., et al. "Big data and big cities: The promises and limitations of improved measures of urban life." Economic Inquiry (2018).
Glaeser, Edward L., Jose A. Scheinkman, and Andrei Shleifer, "Economic Growth in a Cross-Section of Cities," Journal of Monetary Eco- nomics 36 (1995), 117–143.
Glaeser, Edward L., and Jesse M. Shapiro, "Urban Growth in the 1990s: Is City Living Back?" Journal of Regional Science 43:1 (2003a), 139–165.
Gyourko, Joseph, Christopher Mayer, and Todd Sinai. "Superstar cities." American Economic Journal: Economic Policy 5.4 (2013): 167-199.
Gu, W., Zhang, K., Zhou, Z., Jin, M., Zhou, Y., Liu, X., ... & Zhang, L. (2017). Measuring fine-grained metro interchange time via smartphones. *Transportation Research Part C: Emerging Technologies*, *81*, 153-171.
Hagen, C., Khan, K., Ciobo, M., Miller, J., Wall, D., Evans, H. and Yadava, A. (2013), Big Data and the Creative Destruction of Today's Business Models, ATKearney, Chicago
Hashem, Ibrahim Abaker Targio, et al. "The role of big data in smart city." International Journal of Information Management 36.5 (2016): 748-758.

Hall, B.A., Link, N., Scott, J.T., 2003. Universities as research part- ners. Review of Economics and Statistics 85, 485–491.
Han, Qilong, Shuang Liang, and Hongli Zhang. "Mobile cloud sensing, big data, and 5g networks make an intelligent and smart world." IEEE Network 29.2 (2015): 40-45.
Hanushek, E. (2002). "Publicly provided education". In: Auerbach, A., Feldstein, M. (Eds.), The Handbook of Public Economics, vol. 3. North-Holland, Amsterdam.
Hemminki, S., Nurmi, P., Tarkoma, S., 2013. Accelerometer-based transportation mode detection on smartphones. In: Proc. ACM SenSys, pp. 1–14.
Henderson, Vernon, Ari Kuncoro, and Matt Turner. "Industrial development in cities." Journal of political economy 103.5 (1995): 1067-1090.
Holt, Lynne, and Mark Jamison. "Broadband and contributions to economic growth: Lessons from the US experience." Telecommunications Policy 33.10 (2009): 575-581.
Horan, Thomas A. 'Digital places'. DISP 144 (2001): 12-19.
Hartmann, P. M., Zaki, M., Feldmann, N., & Neely, A. (2016). Capturing value from big data–a taxonomy of data-driven business models used by start-up firms. *International Journal of Operations & Production Management*, *36*(10), 1382-1406.
Händel, Peter, et al. "Smartphone-based measurement systems for road vehicle traffic monitoring and usage-based insurance." *IEEE Systems Journal* 8.4 (2014): 1238-1248.
Jaffe, A., Trajtenberg, M., Henderson, R., 1993. Geographic localization of knowledge spillovers an evidenced by patent citations. Quarterly Journal of Economics 434, 578–598.
Jensen, Robert. "The digital provide: Information (technology), market performance, and welfare in the South Indian fisheries sector." The quarterly journal of economics 122.3 (2007): 879-924.
Kahn, Matthew E., "Smog Reduction's Impact on California County Growth," Journal of Regional Science 40:3 (2000), 565–582.
Krugman, P., 1991. Increasing returns and economic geography. Journal of Political Economy 99 (3), 483–499.
Koutroumpis, Pantelis. "The economic impact of broadband on growth: A simultaneous approach." Telecommunications policy 33.9 (2009): 471-485.
Kandilov, Ivan T., and Mitch Renkow. "Infrastructure investment and rural economic development: An evaluation of USDA's broadband loan program." Growth and Change 41.2 (2010): 165-191.
Kolko, Jed. "Broadband and local growth." Journal of Urban Economics 71.1 (2012): 100-113. Kopetz H. Real-time Systems. New York: Springer, 2011. 307–323.
Kim, Younjun and Orazem, Peter F., "Broadband internet and firm entry: evidence from rural Iowa" (2012). Economics Working Papers (2002–2016). 79.
Kummerow, Max, and Joëlle Chan Lun. "Information and communication technology in the real estate industry: productivity, industry structure and market efficiency." Telecommunications Policy 29.2 (2005): 173-190.
Lange, Fabian, and Robert Topel. "The social value of education and human capital." Handbook of the Economics of Education 1 (2006): 459-509.
Lee, Sanghoon. "Ability sorting and consumer city." Journal of urban Economics 68.1 (2010): 20-33.
Lee, John D., and Katrina A. See. "Trust in automation: Designing for appropriate reliance." *Human factors* 46.1 (2004): 50-80.
Lehr, W., Gillett, S., Osorio, C., & Sirbu, M. (2006). Measuring broadband's economic impact. Broadband Properties, 12–24.
Manyika, James, and Charles Roxburgh. "The great transformer: The impact of the Internet on economic growth and prosperity." McKinsey Global Institute 1 (2011).
Mack, Elizabeth A., and Tony H. Grubesic. "Broadband provision and firm location in Ohio: An exploratory spatial analysis." Tijdschrift voor economische en sociale geografie 100.3 (2009): 298-315.
Mack, Elizabeth A., Luc Anselin, and Tony H. Grubesic. "The importance of broadband provision to knowledge intensive firm location." Regional Science Policy & Practice 3.1 (2011): 17-35.
Mack, Elizabeth, and Alessandra Faggian. "Productivity and broadband: The human factor." International Regional Science Review 36.3 (2013): 392-423.
Mack, Elizabeth A., and Sergio J. Rey. "An econometric approach for evaluating the linkages between broadband and knowledge intensive firms." Telecommunications Policy 38.1 (2014): 105-118.

Mack, Elizabeth A. "Businesses and the need for speed: The impact of broadband speed on business presence." Telematics and Informatics 31.4 (2014): 617-627.
Martinez, Francisco J., et al. "Assessing the impact of a realistic radio propagation model on VANET scenarios using real maps." Network Computing and Applications (NCA), 2010 9th IEEE International Symposium on. IEEE, 2010.
Moretti, Enrico. Human capital externalities in cities. In J. Vernon Henderson and Jacques-François Thisse, editors, Handbook of Urban and Regional Economics, chapter 8. North Holland-Elsevier, April 2003.
Moretti, Enrico. Estimating the social return to higher education: Evidence from longitudinal and repeated cross-sectional data. Journal of Econometrics, 121(1-2):175—212, July-August 2004. Musolesi, M. “Big Mobile Data Mining: Good or Evil?” IEEE Internet Com- puting, vol. 18, no.1, 2014, pp. 78–81.
Naik, N., S. D. Kominers, R. Raskar, E. L. Glaeser, and C. A. Hidalgo (2015). Do people shape cities, or do cities shape people? The co-evolution of physical, social, and economic change in five major US cities. NBER Working Paper No. 21620.
Naik, N., J. Philipoom, R. Raskar, and C. A. Hidalgo (2014). Streetscore – predicting the perceived safety of one million streetscapes. In IEEE Computer Vision and Pattern Recognition. Qiang, Christine Zhen-Wei, Carlo M. Rossotto, and Kaoru Kimura. "Economic impacts of broadband." Information and communications for development 2009: Extending reach and increasing impact 3 (2009): 35-50.
Rohman, Ibrahim Kholilul, and Erik Bohlin. "Does broadband speed really matter as a driver of economic growth? Investigating OECD countries." International Journal of Management and Network Economics 5 2.4 (2012): 336-356.
Rauch, J.E., 1993. Productivity gains from geographic concentration of human capital: evidence from the cities. Journal of Urban Economics 34 (3), 380–400.
Rosenthal, S.S., Strange, W.C., 2001. The determinants of agglomeration. Journal of Urban Economics 50, 191–229.
Rosenthal, Stuart S., and William C. Strange. "The attenuation of human capital spillovers." Journal of Urban Economics 64.2 (2008): 373-389.
Saxanion, A. (1996) Regional Advantage: Culture and Competition in Silicon Valley and Route 128, Cambridge, MA: Harvard University Press.
Sankaran, K., Zhu, M., Guo, X.F., Ananda, A.L., Chan, M.C., Peh, L.S., 2014. Using mobile phone barometer for low-power transportation context detection. In:Proc. ACM SenSys, pp. 191–205.
Sena, Vania. "Total factor productivity and the spillover hypothesis: Some new evidence." International Journal of Production Economics 92.1 (2004): 31-42.
Shapiro, Jesse M. "Smart cities: quality of life, productivity, and the growth effects of human capital." The review of economics and statistics 88.2 (2006): 324-335.
Stenberg, Peter, Mitchell Morehart, Stephen Vogel, John Cromartie, Vince Breneman, and Dennis Brown. 2009. “Broadband Internet’s Value for Rural America.” U.S. Department of Agriculture, Economic Research Service.
Stuart, T.E., Sorenson, O., 2003. The geography of opportunity: spatial heterogeneity in founding rates and the performance of biotechnology firms. Research Policy 25, 1139–1157.
Simon, Curtis J., “Industrial Reallocation across U.S. Cities, 1977–97,” Journal of Urban Economics 56:1 (2004), 119–143.
Simon, Curtis J., “Human Capital and Metropolitan Employment Growth,” Journal of Urban Economics 43 (March 1998), 223–243.
Simon, Curtis J., and Clark Nardinelli, “Human Capital and the Rise of American Cities, 1900–1990,” Regional Science and Urban Economics 32 (2002), 59–96.
Tabuchi, Takatoshi, and Atsushi Yoshida. "Separating urban agglomeration economies in consumption and production." Journal of Urban Economics 48.1 (2000): 70-84.
Thiagarajan, Arvind, James Biagioni, Tomas Gerlich, and Jakob Eriksson. 2010. Cooperative transit tracking using smart-phones. In Proceedings of the 8th ACM Conference on Embedded Networked *Sensor Systems* (SenSys '10). ACM, New York, NY, USA, 85-98.
Van Ark, B., & Inklaar, R. (2005). Catching up or getting stuck? Europe’s troubles to exploit ICT’s productivity potential (September)

Yin, Chuan, Tao, et al. "A literature survey on smart cities." Science China Information Sciences 58.10 (2015): 1-18.
Yovanof, Gregory S., and George N. Hazapis. "An architectural framework and enabling wireless technologies for digital cities & intelligent urban environments." Wireless personal communications, 49:3 (2009): 445-463.
Zanella, Andrea, et al. "Internet of things for smart cities." IEEE Internet of Things journal 1.1 (2014): 22-32.
Zheng, Kan, et al. "Big data-driven optimization for mobile networks toward 5G." IEEE network 30.1 (2016): 44-51.
Zhou, Pengfei, Yuanqing Zheng, and Mo Li. "How long to wait?: predicting bus arrival time with mobile phone based participatory sensing." *Proceedings of the 10th international conference on Mobile systems, applications, and services*. ACM, 2012.
Waldrop, M. Mitchell. "The chips are down for Moore's law." *Nature,* 530.7589 (2016): 144-147.
Wellman, B. (1999) The Networked Community, in Networks in the Global Village: Life in Contemporary Communities (Boulder, CO: Westview Press
Wheeler, Christopher H. "Search, sorting, and urban agglomeration." Journal of Labor Economics 19.4 (2001): 879-899.
Wheeler, Christopher H. "Cities and the growth of wages among young workers: Evidence from the NLSY." Journal of Urban Economics 60.2 (2006): 162-184.
Whitacre, Brian, Roberto Gallardo, and Sharon Strover. "Broadband's contribution to economic growth in rural areas: Moving towards a causal relationship." Telecommunications Policy 38.11 (2014): 1011-1023.
World Economic Forum. (2011) 'Personal Data: The Emergence of a New Asset Class '