Beijing City Lab

Long Y, Huang C, 2016, Does block size matter? The impact of urban design on economic vitality for Chinese cities. Beijing City Lab. Working paper #74

Does Block Size Matter? The Impact of Urban Design on Economic Vitality for Chinese Cities

Ying Long, School of Architecture, Tsinghua University, China *
CC Huang, Energy Innovation, USA

* Corresponding author, ylong@tsinghua.edu.cn

Abstract: The influence of urban design on economic vitality has been analyzed by a number of researchers and is also a key focus of many planning/design theories. However, most quantitative studies are based on just one city or a small set of cities, rather than a large number of cities that are representative of an entire country. With the increasing availability of new data, we aim to alleviate this gap by examining the impact of urban design upon economic vitality for the 286 largest cities in China by looking at a grid of geographical units that are 1km by 1km. We use these units and a set of new data to look at the impact of urban form indicators, such as intersection density (urban design), level of mixed-use, and access to amenities and transportation, on economic vitality represented by activities using social media data. Our results show that these urban design indicators has a significant and positive impact on levels of economic vitality for cities at every administrative level. The results contribute to a holistic understanding of how to improve economic vitality in cities across China at a detailed level, particularly at a time when China's economic growth will depend largely on growth of the service sector in urban areas. We think these results can help decision makers, developers, and planners/designers to improve economic vitality in cities across China.

Keywords: consumption vitality; intersection density; block size, big/open data; China

1. Introduction

The contribution of urban design principles to vibrant and prosperous cities, particularly its impact on urban vitality[1], has become common belief. Pioneering urbanists like Jacobs (1961) and Lefebvre (1962) were the among the first to specify the features of good urban design (e.g. small blocks) and how it cultivates community, sociocultural vibrancy, and healthy neighborhoods. Other influential urbanists such as Gehl (1971), Lynch (1981), Whyte (1980), Montgomery (1998) also established new ground on ways to understand urban vibrancy. Urbanists like Jacobs and Gehl, along with the New Urbanists, broadly promoted pedestrian-friendly, compact, walkability, and mixed-use neighborhoods (Katz et al., 1994). Since the advent of these urbanists, there has been extensive and qualitative discussion on how good urban design promote the formation of lively city and urban vitality. The variables that we quantitatively analyze in this paper – intersection density, level of mixed-use, access

[1] Urban vitality has been widely regarded as an indicator of a continuous gradient of diverse socio-economic performance (Lees, 2010; Holanda, 2011). It is this feature that makes this concept widely used in sociology, architecture, urban planning and urban design, environmental psychology, to represent various kinds of understandings of vitality.

to amenities and transportation – are aimed at using data to confirm the beliefs of these great urban thinkers.

This paper aims to quantitatively explore the impact of urban design by looking at the impact of variables such as intersection density, level of mixed-use, and access to amenities and transportation, on economic vitality. We are particularly interested in exploring the answers to the following two questions: (1) Do urban design principles positively contribute to vibrant cities in terms of economy? How much does a variable such as intersection density impact economic vitality compared to other factors? (2) Does this impact vary across cities of different administrative levels?

This paper aims to fill a gap in existing studies by using granular data to analyze 286 of the largest cities in China. Of all the countries in the world, China's urbanization rate is among the highest and its urban population is also among the largest. We use ordinary least squares (OLS) to look at the comparative impact of each urban design variable on economic vitality. We use social media comments, sign-ins, and housing price data as proxies for economic vitality. Our analysis includes urban form variables such as intersection density, level of mixed-use, access to amenities, and access to transportation. Our results show that intersection density has a significant and unique impact on economic vitality, even when controlling for variables such as point of interest density and population density. We also find that other urban form variables such as level of mixed-use, access to amenities, and access to transportation also have a significant effect on economic vitality.

This paper is organized as follows. Section 2 provides a review of the literature on this topic, looking at relevant literature while also demonstrating the emerging consensus around the results of our analysis. Section 3 describes our dataset in detail and Section 4 presents our regression results. Section 5 concludes with suggestions for future research.

2. Literature Review

The widespread use of new data sources, including mobile phone traces, public transportation smartcard records, social media data and geo-tagged data, has created the possibility to understand how urban design impacts economic vitality (Liu et al., 2015; Zhou and Long, 2016). These data sources allow researchers to have a more granular and human-scale understanding of how people experience cities with increased accuracy, consistency, and detail (Dunkel, 2015). Our literature review shows that while urban form indicators have been widely analyzed, there have only been a handful of studies that look at the impact of urban design on urban or economic vitality[2] through quantitative analysis.

We find several specific studies that connect urban design with urban (especially economic) vitality. One study (De Nadai et al., 2016), targets six cities in Italy, and aims to confirm four of Jane Jacob's principles for urban vitality using mobile phone data. Two of the principles overlap with our indicators (mixed-use and intersection density) and the study finds that both of these characteristics have a significant effect on economic vitality. Jane Jacobs's principles are also examined by Sung et al.

[2] We assume economic vitality is a dimension of urban vitality in this paper, which focuses more on economic vitality rather than the broad urban vitality.

(2015), using the conventional household travel survey, by examining their role on pedestrian activity and evidence that they have a positive impact are found in Seoul, Korea. Similar studies supported by urban GIS and emerging new data have been conducted in individual Chinese cities, like Beijing by Long and Zhou (2016) using mobile phone data, Wu (2016) using mobile phone positioning records, and Zheng et al. (2016) using catering establishment data, and Shenzhen by Yue et al. (2016) using mobile phone data. Most of the aforementioned publications in this paragraph are for a single city or a limited number of cities. For example, De Nadai et al. (2016) looks at six cities, rather than having an aggregate view of all the cities in a country. They are all addressing the role of urban form on population density represented social vitality, with Zheng et al. (2016) as an exception, rather economic vitality which is the focus of this paper.

There are also a number of studies that show there is a positive impact of indicators related to urban design, such as walkability, on property values and other economic indicators. While data such as property values can provide some insight into the role that urban design can influence a city's economy, social media data provides more granularity and more information on resident's direct *participation* in the economy. Li et al. (2015) and Loehr (2013) show that walkability can positively impact the property values of a neighborhood, while other researchers have shown that proximity to highways has a negative impact on property values (Tajima, 2003, Madison and Kovari, 2013).

While quantitative analysis that looks at the correlation between urban design and economic vitality is relatively new, studies that look at the relationship between the urban built environment and intensity of travel behavior is relatively mature. For example, there is research that show street design parameters such as block size have a positive impact on pedestrian volume (Hess et al, 1999; Vernez et al. 1997). In addition to the extensive studies on urban form and travel from Ewing and Cervero (2010), there are a number of studies that show that pedestrian activity increases with mixed-use, access to amenities, point of interest (POI) density, and intersection density (Sung et al. 2015, Sunga et al. 2015; Krizek 2003). This paper goes a step further to integrate pedestrian activity with participation in commercial or econoic activity, which is our view of "economic vitality."

3. Data

For this analysis, we have used a number of data sources, including social media data, geo-tagged data, property value data, and government-released economic data for control variables. In this section, we will review each dataset. Table 1 offers more information on the data we use in this paper.

Table 1. List of Variables and Data Sources Used in Analysis

Main Type	Name	Description	Data Source
Vitality	DIANPING	Dianping comments for each square kilometer grid in 2014 (#/ha)	http://www.dianping.com, and used by Long (2016)
	WEIBO	Weibo count of each square kilometer grid in 2014 (#/ha)	http://www.weibo.com
	HOUSING_PRIC	Average housing price for each square	http://www.Soufun.com

	E	kilometer grid in 2014 (CNY)	
Design	INTERSECTION	Number of intersections for each square kilometer grid in 2014 (#/ha)	Long (2016)
Density	POI_DENSITY	POI (point of interest) count of each square kilometer grid in 2014 (#/km^2)	Long (2016)
	POP_DENSITY	Population count of each grid derived from the township level population density of China in 2010 (#/km)	Long et al (2015)
Diversity	MIXED-USE	Mixed-use level determined per grid from data in 2014	Long (2016) using the method in Liu and Long (2016)
Amenities	AMENITIES	The total number of bus stops, education and research facilities, governmental facilities and convenient stores in each square kilometer grid in 2014 (#/km^2)	Derived from POIs in 2014, see Long (2016)
Access to Transit	ACCESSIBILITY	The average air distance to the closest city center, sub-center, green space, shopping center, hospital, and subway and HSR station in 2014 (km)	Derived from POIs in 2014 using Spatial Analyst of ArcGIS, see Long (2016)
City level control variables	GDP	GDP per capita of the city in 2014 (CNY)	MOHURD (2015)
	TERTIARY	Percentage of GDP that the tertiary sector accounted for in 2014	
	INCOME	Average income of each person in the city in 2014	
	CITY_LEVEL	The administrative level of the city in 2014 (5 for the metropolitan cities and 2 for the prefectural cities)	See details in Section 3.1

3.1. Geographic Data

In this paper, our data comes from all the major cities in China and we exclude data outside government-defined urban limits. The Chinese central government has specific boundaries and labels that demarcate and categorize Chinese cities and the data from the government's statistical yearbooks is based on this system. The government defines five main administrative levels of cities, including: 4 municipalities (MD, considered Tier 1), 15 sub-provincial cities (SPC, Tier 2), 17 other provincial cities (OPCC, Tier 3), 250 prefecture-level cities (PLC, Tier 4), and 367 county-level cities (CLC, Tier 5). As of 2014, there was a total of 653 Chinese cities from all these administrative levels[3] (Figure 1). Tier 1 cities are generally the largest and have the most mature economies. Beijing, Tianjin, Shanghai, and Chongqing are the four Tier 1 cities. Tier 2 cities are cities such as Chengdu and Wuhan, while Tier 3 cities include cities such as Hefei and Changsha. In general, Tier 1 and Tier 2 cities are larger and have more developed economies compared to Tier 3 and Tier 4 cities, with Tier 4 cities having the smallest economies. The administrative boundaries shown in Figure 1 were divided into 782,263 one square

[3] Due to changes in the urban system, several inconsistent cities have been excluded

kilometer units to represent the main units of analysis used in this paper.[4]

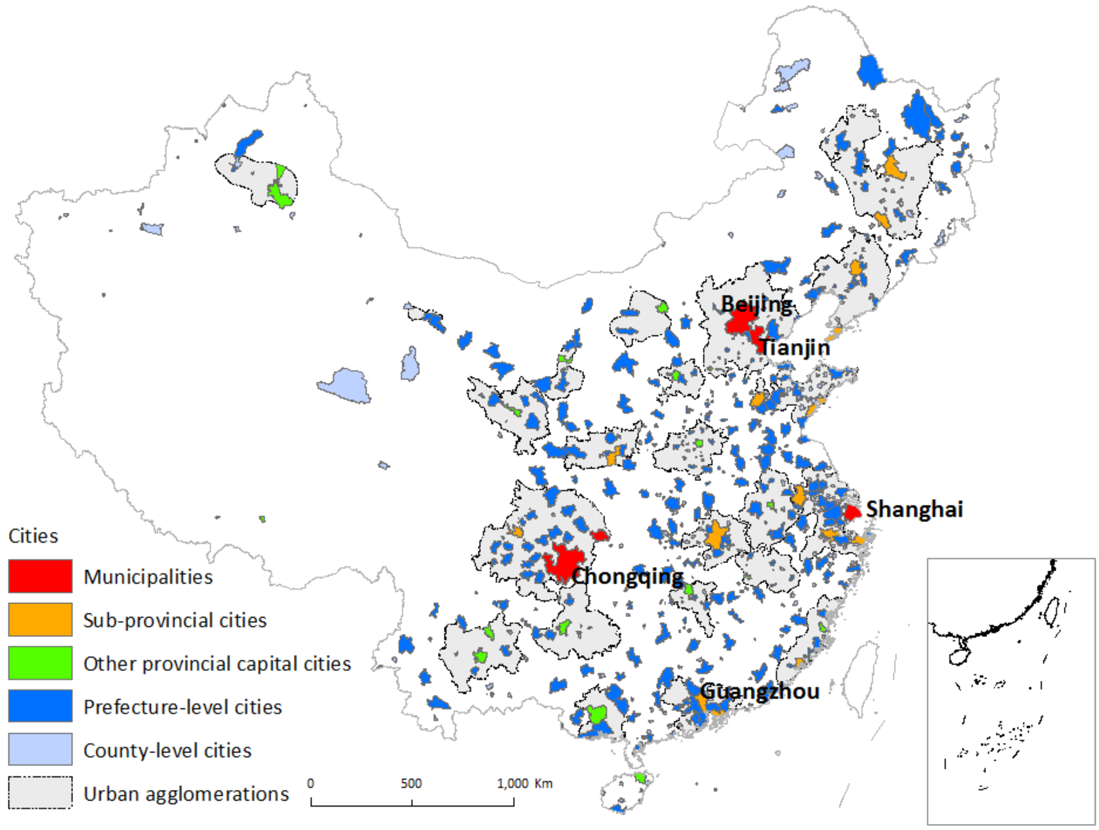

Figure 1. Chinese cities demarcated by administrative area (the polygon in color stands for the administrative area of each city)

Since not all of the one square kilometer areas located in the official administrative areas shown in Figure 1 are urbanized, we use data from Landsat TM images in the year 2010 by Liu et al. (2014) to identify the urbanized areas. In China, there were a total of 63,425 km^2 of urban areas in 2010, as indicated by the data. We categorize any area that has more than 0.5 km^2 of urban land areas as an urbanized area[5]. As a result, we determine that there are 43,646 geographic units of urbanized land in total within administrative areas of Chinese cities.

3.2 Economic Vitality Data

To evaluate the economic vitality of each area, we collect data from Dianping,[6] which is an aggregated social media tool used to rate restaurant and other service industry companies in China (Liu, 2014). For this analysis, the main data we use as a dependent variable is the quantity of comments on Dianping from the users per geographic unit and we interpret this as a proxy for economic vitality. There are major advantages to using Dianping. First, there is data on all of the Tier 1-4 cities that are used in our analysis. Second, there are no major competitors to Dianping, making it

[4] These grids are the same with the grids used in our open data project SinoGrids, see Zhou and Long (2015) for details.

[5] In this paper, we use the term urbanized grids to represent those grids within the administrative boundaries of Chinese cities, while they have half urban lands in each grid.

[6] Dianping can be understood to be the Chinese version of Yelp (www.dianping.com).

the most comprehensive and representative data source for both foot traffic in addition to actual engagement with commercial establishments within the geographic unit. In total, we were able to collect data from 13 million POIs from Dianping, and there are 1.9 m POIs that have at least one user comment. There are a total of 47 million comments for all POI's on Dianping. In this study, each establishment in Dianping is assigned to an urban geographic unit, and then each unit is linked with anywhere from a few to hundreds of establishments. We found that the Dianping data is not representative for county-level cities since the smallest cities do not always have enough Dianping users, hence we do not analyze cities that are smaller than Tier 4 cities. Moreover, we only include geographic units that contain at least one user comment, and we remove the top 1% of geographic units that contain the most comments to avoid outliers.

To complete the dataset, we aggregate the comment totals within each geographic unit as a representation of the economic vitality of that unit. After this process of improving data quality, of all of the geographic units we aggregated (as described in Section 3.1), there were 24,512 urban geographical units[7] in Tier 1, Tier 2, Tier 3, and Tier 4 cities that met the requirements outlined above. These 24,512 geographic units are the final units used in this analysis (Figure 2).

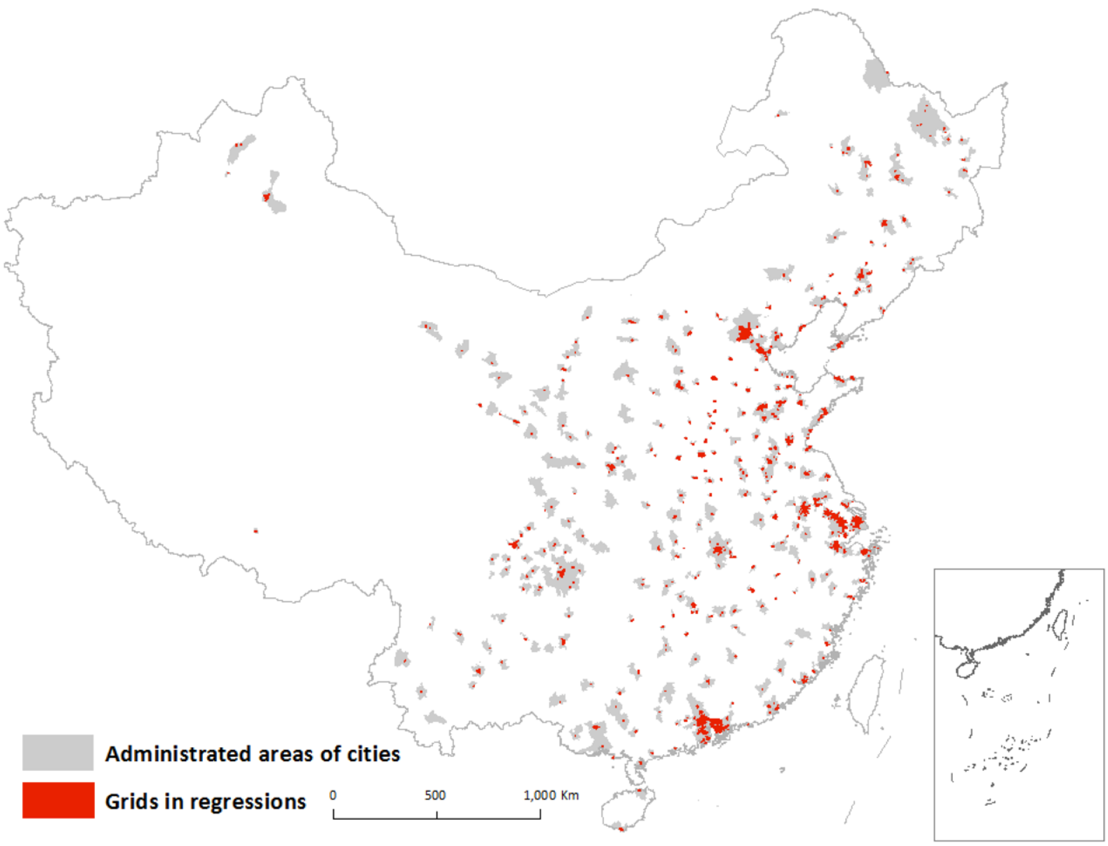

Figure 2. Geographic Units Used (urban grids)

3.3 Urban Data

According to three D's (density, diversity and design) proposed by Bernick and Cervero (1996) and

[7] We use the term "urban grids" here to differentiate the term "urbanized grids" used in Section 2. In the following context, we will use urban grids for representing our analysis samples.

five D's (density, diversity, design, distance to transit and destination accessibility) proposed by Belzer and Autler (2002), we also gather other data (mainly for measuring urban form) as predictive variables for economic vitality for all urban grids.

For the urban form variables, we use six distinct independent variables. This includes intersection density (as a proxy of urban design), POI and population density, access to amenities, and access to transit data for each city. Table 1 above shows detailed explanations and data sources for each data type. The first five types of variables are calculated for each geographic unit, similar to the Dianping data. This means that these variables are heterogeneous based on the geographic unit. For all of 24,512 urban geographic units used as samples in this paper, we calculate these five types of data per geographic unit.

4. Results

4.1 Explorative Data Analysis

For further insight on the variables in this analysis, this section will conduct explorative data analysis. Table 2 shows all the variables used in our main regression. The dependent variable DIANPING, representing total number of comments accrued per geographical unit ranges from 1 to 30,632 per km^2 and exhibits a long-tailed distribution, shows that there are more units with a very high number of comments than units with a very low number. To balance this, we take the natural log of DIANPING for the regressions. AMENITIES and ACCESSIBILITY are long-tailed distributed as well. In contrast, INTERSECTION, ranging from 0 to 462 intersections per square kilometer, exhibits a left-skewed normal distribution. MIXTURE is right-skewed and normally distributed.

Table 2. Descriptive Statistics

No.	Name	Min.	Max.	Mean	Std. Deviation
1	**DIANPING**	1	30,632	900	2,993
2	**INTERSECTION**	0	462	83	58
3	**POI_DENSITY**	0	2,914	200	254
4	**POP_DENSITY**	0	195,947	9,270	8,324
5	**MIXTURE**	0	1.94	1.38	0.57
6	**AMENITIES**	0	113.3	11.2	11.7
7	**ACCESSIBILITY**	0	336.4	18.0	28.2
8	**GDP**	0	249,040	65,170	27,928
9	**TERTIARY**	0	78.7	48.9	13.0
10	**INCOME**	0	71,923	43,052	12,165
11	**CITY_LEVEL**	2	5	3.0	1.2
	N=24,512				

Figure 3 shows the spatial distribution of DIANPING and INTERSECTION for several representative cities. Both variables exhibit a mono-centric spatial structure with the highest values concentrating at the center of a city. DIANPING and INTERSECTION in each city is similar in pattern, suggesting a strong correlation may exist between two variables. This is then examined in the following regression models.

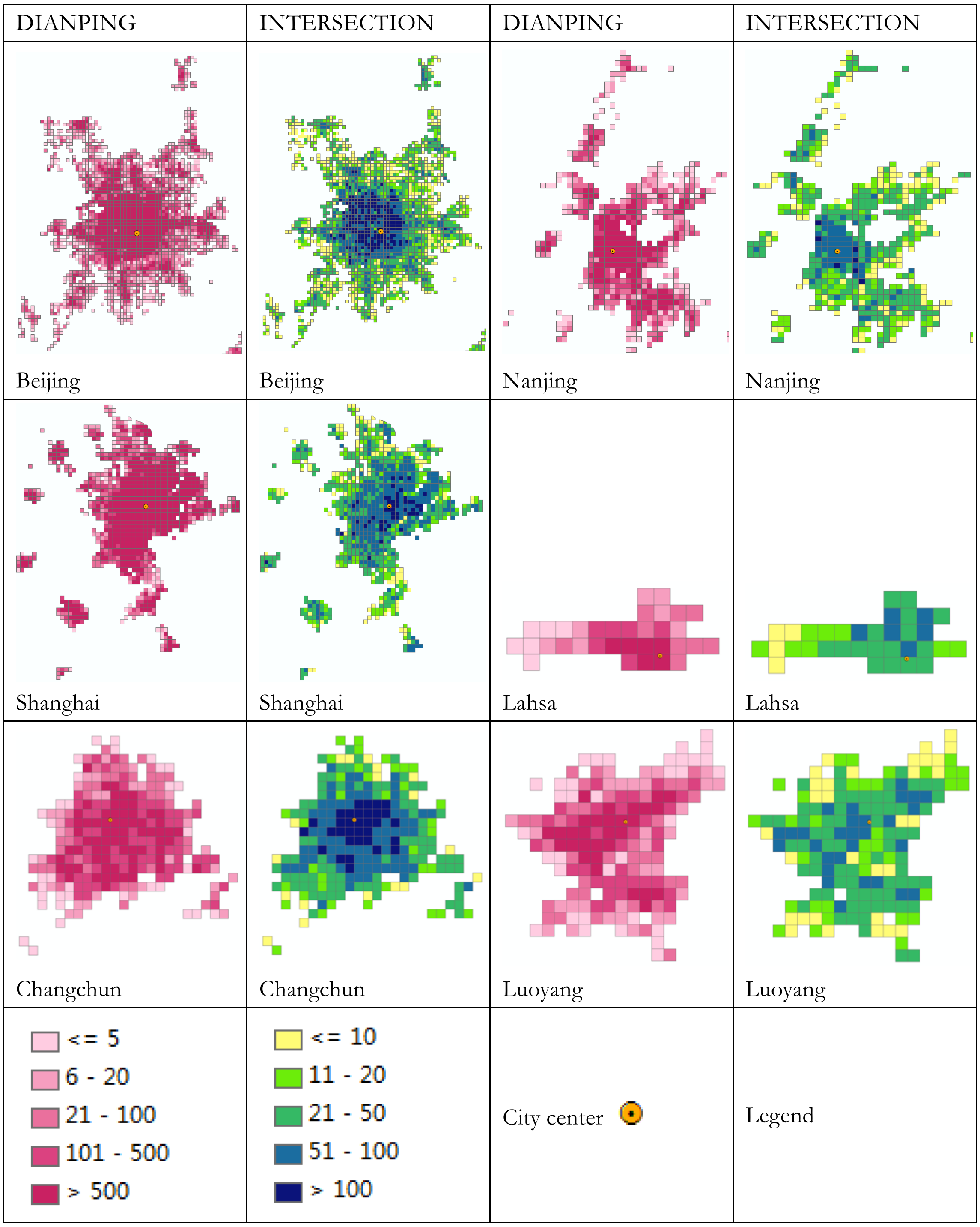

Figure 3. Showing economic vitality (DIANPING) and intersection density (INTERSECTION) for typical Chinese cities

4.2 Regression Results

To understand the effect of urban form variables on economic vitality, we use the regression equation in Equation (1). The natural log of the total comments of each grid Ln(DIANPING) is used as the dependent variable to represent economic vitality. We use Ordinary Least Squares (OLS) as the regression method to understand the effect of various urban form variables on economic vitality.

$$Ln(DIANPING) = ß0 + ß_1*INTERSECTION + ß_2*POI\_DENSITY + ß_3*POP\_DENSITY + ß_4*MIXTURE + ß_5* AMENITIES + ß_6*ACCESSIBILITY + ß_7*GDP + ß_8*TERTIARY + ß_9*INCOME + ß_{10}*CITY\_LEVEL \quad (1)$$

Before conducting regression analysis, we examine the possibility of multicollinearity (or collinearity), whose existence may lead a multiple regression model being overfitting or not reliable. Multicollinearity occurs when two or more of the variables in the model are highly correlated, which results in statistical bias and an unstable/unreliable model. One method to detect multicollinearity is to use the Pearson product-moment correlation coefficients, or Pearson's correlation. Generally speaking, if two variables have a Pearson's correlation above 0.8, there is a high degree of multicollinearity. Another test often used to evaluate the level of multicollinearity is looking at the Variance Inflation Factor (VIF). If any variable has a VIF value bigger than 7, it should also be dropped from the model. After conducting both tests on all the variables in Equation 1, we find that all variables pass correlation analysis and VIF examinations, showing that there is no multicollinearity.

We use several regression models to understand the effect of urban form on economic vitality. The regression results are listed in Table 3. Model 1 only uses one independent variable (INTERSECTION) to explain our dependent variable, Ln(DIANPING). The results of Model 1 indicate that INTERSECTION has a positive and significant effect on economic vitality. This single variable can explain about 40% of economic vitality (R-squared of 0.391), revealing that urban design does have a substantial effect on economic vitality. Next, Model 2 and Model 3 include other independent variables and control variables to develop a more refined understanding of the effect on economic vitality. Model 2 further introduces other urban form variables at the unit-level in addition to intersection density (the key variable for Model 1), and Model 3 includes city-level control variables in addition to these urban form variables. Both Model 2 and Model 3 again suggest that intersection density positively contributes to economic vitality and it is also the most important factor compared to all other variables used in the model.

For Model 1-3 use city-level fixed effects to control for any variability between cities. This means that our results show the average relative impact of intersection density within cities, we are not comparing cities like Beijing to Changsha. Rather, we are trying to understand, in a city like Beijing, how intersection density can explain the differences in economic vitality between different one square kilometer grids. Even after including POI_DENSITY and POP_DENSITY, intersection density still proves to be the most important factor that explains changes in economic vitality. This finding proves our first research question proposed in the introduction section.

In addition to intersection density, our regression models show that other urban form variables also have a significant effect on economic vitality. Access to amenities (AMENITIES) is also a very important influencing variable, along with access to transit. The coefficient on access to transit shows that for each additional kilometer that particular grid is further away from a public transit station, Ln(DIANPING) decreases by -0.11. As this variable is a continuous variable and the coefficient is the variation per kilometer, we can see that access to transit increases its effect on economic vitality the further away from transit a certain grid is. MIXTURE also has a significant effect on economic vitality.

In these regressions, there are also some inconsistent regression results. For instance, the effects of GDP and TERTIARY are negative, showing that a geographic unit in a city with a lower economic development levels and lower levels of industry tends to be associated with a greater economic vitality. The effects of these variables are much smaller than urban form variables. There are a few reasons this might be the case. First, the data available for GDP and TERTIARY are at the city-level, which makes them much less granular than all of the urban design or social media indicators, which are calculated per geographic unit. Second, a significant portion of GDP growth in China is due to the growth of state-owned enterprises, which is a very distinct sector from the commercial and service establishments that the social media data represents. The negative coefficient on GDP might be due to the fact that overall GDP growth and activity in the service sector are not tightly correlated. This has become a major barrier for the Chinese government in continuing economic growth, as the service sector has been historically difficult to grow in a market economy that is heavily influenced by the government. Finally, this might show problems with the quality of government economic data in China.

In addition, the positive coefficient of CITY_LEVEL shows that the administrative level of a city can influence its economic vitality, which answers our second research question. We leave more in-depth analysis in Section 4.3.

Table 3. Regression results for all cities (N=24,512)

Variable	**Model 1**		**Model 2**		**Model 3**	
	Beta	Sig.	Beta	Sig.	Beta	Sig.
(Constant)				0.000		0.000
INTERSECTION	**0.625**	0.000	**0.327**	0.000	**0.266**	0.000
POI_DENSITY			0.011	0.210	**0.052**	0.000
POP_DENSITY			**0.116**	0.000	**0.110**	0.000
MIXTURE			**0.095**	0.000	**0.125**	0.000
AMENITIES			**0.286**	0.000	**0.290**	0.000
ACCESS TO TRANSIT			**-0.111**	0.000	**-0.054**	0.000
GDP					**-0.026**	0.000
TERTIARY					**-0.015**	0.016
INCOME					**0.149**	0.000
CITY_LEVEL					**0.108**	0.000
City fixed effect	Yes					
Dependent variable	Ln(DIANPING)					
N	24,512					
R^2	0.391		0.487		0.528	

Notes: All coefficients have been standardized for cross comparison, and coefficients in bold indicate being significant at the 0.05 level. The constant value is null as we only list the standardized coefficient in the table. Sig. indicates the significant level. The notes also apply to the following tables.

4.3 Regression Results for Different City Tiers

To further understand the impact of different urban form variables on economic vitality, we run the regression models for different municipal administrative levels. A CITY_LEVEL of 5 represents the highest tier cities (a metropolitan city that is directly governed by the state council) and a CITY_LEVEL of 2 represents the lowest tier cities (prefectural level in our samples). The regression results based on administrative level still show that intersection density and the other urban design indicators have a positive effect on economic vitality. Moreover, we find that the effect is larger for the highest tier cities. The regression model for Tier 1 cities can explain 62.4% of economic vitality. The results also show that access to amenities (AMENITIES), mixed-use, and access to transit all have a positive and significant effect on economic vitality for the different tiered cities.[8]

Table 4. Regression Results for Each City Tier

Variable	**Tier 1**		**Tier 2**		**Tier 3**		**Tier 4**	
	Beta	Sig.	Beta	Sig.	Beta	Sig.	Beta	Sig.
(Constant)		0.004		0.000		0.015		0.844
INTERSECTION	0.357	0.000	0.286	0.000	0.177	0.000	0.215	0.000
POI_DENSITY	0.066	0.000	0.038	0.020	0.140	0.000	0.080	0.000
POP_DENSITY	0.089	0.000	0.124	0.000	0.086	0.000	0.076	0.000
MIXED-USE	0.170	0.000	0.123	0.000	0.188	0.000	0.101	0.000
AMENITIES	0.148	0.000	0.317	0.000	0.329	0.000	0.322	0.000
ACCESSIBILITY	-0.187	0.000	-0.067	0.000	-0.024	0.107	-0.064	0.000
GDP	0.125	0.000	-0.162	0.000	0.155	0.000	0.045	0.000
TERTIARY	-0.149	0.000	-0.018	0.110	0.005	0.728	0.040	0.000
INCOME	0.108		0.088	0.000	-0.082	0.000	0.035	0.000
City fixed effect	Yes							
Dependent variable	Ln(DIANPING)							
N	4039		5467		2493		12513	
R^2	0.624		0.569		0.583		0.416	

4.4 Regression Results Using Other Dependent Variables

As there could be potential biases in solely looking at Dianping comments to understand economic vitality, we also look at other variables that can explain economic vitality. For these, we look at the

[8] We also take a more detailed look at the DIANPING data based on the administrative level of the city. We found that a cites with a higher administrative level were generally associated with a greater average value of DIANPING comments.

average residential housing price per one square kilometer unit and also Sina Weibo records. We use Dianping data for the main regressions is that this data is the largest and most comprehensive across all different types of cities, while housing price data and Weibo data is less comprehensive. We use a similar regression equation to Equation 1 above with no change in the independent variables (urban form variables) but we use housing price (HOUSING_PRICE) and Weibo records (WEIBO) as the dependent variables. Similar to the previous analysis, we exclude the lowest level administrative cities as the data quality is subpar for those cities.

Residential housing price represents the performance of real estate projects. We collected all online housing price information from www.Soufun.com in 2014 and have obtained about 400,000 housing price records. All records are geocoded according to the specific addresses provided on the website, and then this data is laid on top of the square kilometer grids to calculate the average housing price of each grid. We then use this average housing price of each square kilometer grid as the dependent variable in the new model. The regression results shown in Table 5 suggest that intersection density has positive and significant effect on housing price, another indicator that can represent economic vitality. This means that, the smaller the blocks in a particular area, the higher the average housing price in that block. Access to transit also has a positive and significant effect on housing price – the closer the area is to transit, the higher the housing price.

However, it is interesting to note that the impact of the other urban form variables on housing prices varies from their impact on Dianping comments. We find that level of mixed-use and access to amenities has a negative effect on housing price. This might be due to the fact that our housing price data is not as comprehensive as the Dianping comments data. It is also possible that level of mixed-use in some Chinese cities is still associated with older neighborhoods, which means that the negative effect is due to the fact that older neighborhoods have a higher level of mixed-use, and older apartments tend to be valued less.

As another indicator of economic vitality, we use data from Sina Weibo, a social media platform that has check-in data, showing where users tag their location on the platform. Sina Weibo, the Chinese version of Twitter combined with Foursquare, is the leading online microblogging platform in China, which allows users to publish, share and discuss short postings on their website (http://www.weibo.com). [9] We collected 30.3 million Weibo records that had geographical information (this represents about 1% of all Weibo[10] posts, as about 1% of Weibos are geotagged) for July and October 2014 for the entirety of China. Like the housing price data, this Weibo data is laid on top of the geographic units so we have a count of how many Weibo records there are per unit. Model 2 of Table 5 below shows the regression model with number of Weibo posts as the dependent variable. The model can explain about 60% of online activities, and INTERSECTION, MIXTURE, and AMENITIES all show a positive and significant effect on economic vitality as represented by Weibo. Again, like the Dianping and housing price regressions, we can see that intersection density is the most important factor impacting economic vitality.

[9] Weibo is used by over 30% of Internet users in China, with a market penetration similar to Twitter. It was launched by SINA Corporation on 14 August 2009, and at the end of the fourth quarter of 2014, the number of monthly active users has reached 176 million, 80% of which come from mobile terminals.

[10] The same way that a “tweet” is a short message a user would send on Twitter, a “Weibo” is a message that a user sends on Weibo.

Table 5. Regression results for housing price and Weibo records

Variable	HOUSING PRICE		WEIBO	
	Beta	Sig.	Beta	Sig.
(Constant)		0.000		0.000
INTERSECTION	**0.159**	0.000	**0.298**	0.000
POI_DENSITY	**-0.026**	0.009	**-0.027**	0.000
POP_DENSITY	**0.107**	0.000	**0.105**	0.000
MIXTURE	**-0.026**	0.000	**0.286**	0.000
AMENITIES	-0.018	0.076	**0.252**	0.000
ACCESSIBILITY	**-0.070**	0.000	**-0.046**	0.000
GDP	**-0.052**	0.000	**0.018**	0.000
TERTIARY	**-0.211**	0.000	**0.054**	0.000
INCOME	**0.473**	0.000	**-0.118**	0.000
CITY_LEVEL	**0.053**	0.000	**0.085**	0.000
City fixed effect	Yes			
Dependent variable	Ln(HousingPrice)		Ln(Weibo)	
N	17,997		31,823	
R^2	0.536		0.593	

Note: Not all our urbanized grids in the top four levels of cities have housing price information and Weibo records. Therefore, the samples used in the regressions for the both models vary with each other, and with our previous models for Dianping as well.

Similar to how the additional analysis we performed on the regressions using Dianping, we also looked at the regressions using housing price and Weibo records and divided the data by administrative level (CITY_LEVEL). In these regressions, intersection density continued to have a significant and positive impact on both housing price and Weibo records for every city level. This finding is consistent with our previous regression results using Dianping comments as the dependent variable.

5 Discussion

How urban design variables influences economic vitality continues to be analyzed by various stakeholders such as decision makers, developers, planners & designers as well as academic researchers. Home buyers or business owners should also be interested in this question as this may impact how how their investments on homes, offices, or businesses perform. This paper aimed to disentangle this important question by examining the relationship between urban design indiucators and economic vitality while controlling for other variables. This analysis was made possible by the emerging stream of new data sources – social media data, economic data, and other forms of Internet data. We look at 24,512 one square kilometer grids in 286 of the largest cities in China, in contrast to other existing studies that just look at a small set of cities or a single city. For all our regression models, we have found a positive and significant effect of intersection density on economic vitality. And for most of our models, we have also found a positive and significant effect of other urban design indicators such as mixed-use, access to amenities, and access to transit. We find that this effect holds for each administrative city level in China, showing that cities of all sizes must consider urban design principles to improve urban development. These findings on the important role of urban design is also shown by our two other proxy indicators for economic vitality – housing price and Weibo records.

We believe that this analysis has a few practical applications. First, the positive effect of intersection density on economic vitality should encourage real estate developers to propose urban planning schemes with smaller blocks. Second, our findings should also be considered by local governments and developers when creating long-term urban development plans. Finally, the database that we created for this analysis will be released in the public domain. We hope that the shared dataset can be used for other urban studies as well.

While we believe that the results of this paper are relatively robust and an important step forward in understanding how to study economic vitality in relation to urban design, we do think there are areas that this analysis could be improved. First, if possible, researchers should obtain data to create a more objective and comprehensive variable to signify economic vitality. Credit card transaction data would be one example of this sort of data. Even more granular data might allow researchers to break urban areas into even smaller geographical units, in contrast to our one square kilometer units, which would provide more accurate and pinpointed insights. Second, other indicators for measuring urban design at a finer level as proposed by Ewing et al. (2013), e.g., imageability, enclosure, human scale, transparency and complexity, could also be included in the regression models to gain more in-depth information on the effect of urban design on economic vitality. Last, in addition to the OLS based regression models applied in this paper, researchers could also try spatial regression models or geographical weighted regression (GWR) models to gain a better understanding.

In conclusion, this study has systematically examined the impact of urban form on economic vitality for the largest cities in China. Our findings contribute to the literature in using data to confirm the importance of urban design principles to further economic vitality. As China will continue to add 300 million people to cities by 2030, the has important implications for Chinese cities that are still developing greenfields and going through re-development of existing built-up areas. Empirical evidence continues to confirm the importance of designing cities that have small blocks, high levels of mixed-use, and provide access to transit and amenities. We hope that researchers in other countries will continue to use data to make the increasingly compelling case that cities must be built for the human-scale.

Acknowledgment: This paper was financially supported by Energy Innovation. We are grateful to Shiyun (Lily) Qiu for her contribution in completing the literature review. The data used in this analysis is available on Beijing City Lab's website (www.BeijingCityLab.com).[11]

[11] Available once this paper is accepted for publication.

References

1. Bernick, M., & Cervero, R. (1996). Transit Villages in the 21st Century. McGraw-Hill.
2. Belzer, D., & Autler, G. (2002). Transit oriented development: Moving from rhetoric to reality (pp. 06-15). Washington, DC: Brookings Institution Center on Urban and Metropolitan Policy.
3. De, Nadai, M., Staiano, J., Larcher, R., Sebe, N., Quercia, D., Lepri, B. The death and life of great Italian cities: A mobile phone data perspective. Proceedings of the 25th International Conference on World Wide Web. International World Wide Web Conferences Steering Committee, 2016, 413-423.
4. Dunkel, A. Visualizing the perceived environment using crowdsourced photo geodata. Landscape and Urban Planning. 2015, 142: 173-186.
5. Ewing, R., & Cervero, R. Travel and the built environment. Journal of the American Planning Association. 2010, 76(3): 265–294.
6. Ewing R, Clemente O, Neckerman K M, et al. Measuring urban design: Metrics for livable places[M]. Washington, DC: Island Press, 2013.
7. Gehl, J. (1971). Life between Buildings: Using Public Space. New York: Van Nostrand Reinhold.
8. Hess, P., Moudon, A., Snyder, M., Stanilov, K. Site design and pedestrian travel. Transportation Research Record. 1999, 1674: 9-19.
9. Holanda, F. (2011). Exceptional Space (FRBH, Brazil)
10. Jacobs, J. (1961). The Life and Death of Great American Cities, New York: Random House.
11. Katz, P., Scully, V. J., Bressi, T. W. (1994). The New Urbanism: Toward an Architecture of Community. New York: McGraw-Hill.
12. Krizek, K J. Residential relocation and changes in urban travel: Does neighborhood-scale urban form matter? Journal of the American Planning Association. 2003, 69(3): 265-281.
13. Jacobs J, 1961 The Life and Death of Great American Cities (Random House, New York)
14. Lees, L. Planning urban vitality?. Environment and Planning A. 2012, 42: 2302-2308.
15. Lefebvre, H. (1962). 'Notes on the new town', Introduction to Modernity. London: Verlso, 116-126.
16. Li, W., Joh, K., Lee, C., Kim, J., Park, H., Woo, A. Assessing benefits of neighborhood walkability to single-family property values a spatial hedonic study in Austin, Texas. Journal of Planning Education and Research. 2015, 35(4): 471-488.
17. Liu, J., Kuang, W., Zhang, Z., Xu, X., Qin, Y., Ning, J., Zhou, W., Zhang, S., Li, R., Yan, C., Wu, S., Shi, X., Jiang, N., Yu, D., Pan, X., & Chi, W. Spatiotemporal characteristics, patterns and causes of land use changes in China since the late 1980s. Journal of Geographical Science. 2014, 69(1): 3-14.
18. Liu, X. Visualizing urban gastronomy in China. Environment and Planning A. 2014, 46(5): 1012-1013.
19. Liu, X., & Long, Y. Automated identification and characterization of parcels with OpenStreetMap and points of interest. Environment and Planning B: Planning & Design. 2016, 43(2): 341–360.
20. Loehr, S. Mixed-use, mixed impact: Re-examining the relationship between non-residential land uses and residential property values. Columbia University Academic Commons, 2013, http://hdl.handle.net/10022/AC:P:20736.
21. Long, Y., & Zhou, Y. Quantitative evaluation on street vibrancy and its impact factors: A case

study of Chengdu. New Architecture. 2016, (1): 52-57. (In Chinese with English abstract)
22. Lynch, K. (1981). Good City Form. MA: The MIT Press.
23. Madison, C., & Kovari, J. Impact of green infrastructure on property values within the Milwaukee metropolitan sewerage district planning area: Case studies. 2013, The University of Wisconsin-Milwaukee Center for Economic Development.
24. Montgomery, J. Making a city: urbanity, vitality and urban design. Journal of Urban Design. 1998, 3, 93-116.
25. Moudon, A., Hess, P., Snyder, M., Stanilov, K. Effects of site design on pedestrian travel in mixed-use, medium-density environments. Transportation Research Record. 1997, 1578:48-55.
26. MOHURD. (2015) Chinese City Construction Statistics Yearbook 2014. Ministry of Housing and Urban-rural Development of the People's Republic of China, Beijing: China Planning Press.
27. Sung, H., Lee, S., Cheon, S., Operationalizing Jane Jacobs's urban design theory: Empirical verification from the great city of Seoul, Korea. Journal of Planning Education and Research. 2015, 35(2): 117-130.
28. Sunga, H., Gob, D., Choi, C., Cheon, S., Park, S. Effects of street-level physical environment and zoning on walking activity in Seoul, Korea. Land Use Policy. 2015, 49: 152-160.
29. Tajima, K. New estimates of the demand for urban green space: implications for valuing the environmental benefits of Boston's big dig project. Journal of Urban Affairs. 2003, 25(5): 641-655.
30. Yue, Y., Zhuang, Y., Yeh, A.G.O, Xie, J., Ma, C., & Li, Q. Measurements of POI-based mixed use and their relationships with neighbourhood vibrancy, International Journal of Geographical Information Science, 2016, in press, DOI: 10.1080/13658816.2016.1220561
31. Whyte, W. H. (1980). The Social Life of Small Urban Spaces. Washington: The Conservation Foundation
32. Wu, W., Wang, J., Li, C., Wang, M. The Geography of City Liveliness and Land Use Configurations: Evidence from Location-Based Big Data in Beijing. SERC DISCUSSION. 2016, PAPER, available from: http://www.spatialeconomics.ac.uk/textonly/SERC/publications/download/sercdp0201.pdf
33. Zheng, S., Hu, X., Wang, J., Wang, R. Subways near the subway: Rail transit and neighborhood catering businesses in Beijing. Transport Policy. 2016, in press, 1-12.
34. Zhou, Y., & Long, Y. SinoGrids: A practice for open urban data in China. Cartography and Geographic Information Science. 2016, 43: 1-14.