# Understanding Urban Sub-centers with Heterogeneity in Agglomeration Economies—Where do Emerging Commercial Establishments Locate?

**Abstract:**

This paper investigates the formation of employment sub-centers from a new perspective of heterogeneity in agglomeration economies. Using highly granular commercial and residential land-use data (2001-2011) in Chicago, we measure how the locations of jobs, population, quality-of-life amenities, and transportation networks shape specific and heterogenous sub-centers. First, the results suggest that the CBD as it was traditionally defined is no longer the primary source of agglomeration externalities for the new economic sectors; sub-centers with sector-specific positive agglomeration externalities have stronger correlations with new commercial establishments. Secondly, residents appear to give the highest weight to quality-of-life amenities in choosing where to live. Both trends imply dis-incentives for CBD agglomeration. These findings connect the heterogeneous production theories with land use planning and urban design, through new empirical insights into how urban sub-centers grow. Furthermore, we put forward a method for forecasting of future sub-center growth through measuring changes in the probability of commercial development, and discuss its practical implications for planning and design in Chicago.

**Keywords:**

Employment sub-centers; land-use; CBD; agglomeration; heterogeneity

## 1. Introduction

Studies on the urban structure of megacities record a trend away from “distance to the central business district (CBD)” city model. Among numerous analyses of sub-center development, recent evidence suggests a new perspective on urban polycentric evolution in terms of heterogeneity in agglomeration economies. These externalities from the agglomeration of economic entities are not necessarily positive due to differences in interaction patterns between sectors (Firgo and Mayerhofer, 2017; Wixe and Andersson, 2016), or they may even be negative due to competitiveness (Chung and Hewings, 2015). The opposing views on heterogeneity in agglomeration economies shed new light in explaining the departure from the CBD-based urban growth patterns by exploring locational choices of economic sectors with different production externality preferences.

The primary focus of this paper is to understand the employment sub-center development from the view of temporally changing agglomeration economies. The focus of attention is on commercial[1] and residential land-use development.[2] We aim

[1] It is noteworthy that in our study, commercial land-use is broadly defined as land-uses that includes land-use with codes 12XX (commercial), 13XX (institutional), 14XX (industrial), and 15XX (transportation, communication and utility) from the official land-use classifications from Chicago Metropolitan Agency for Planning (CMAP). It should be officially termed as “commercial, institutional, industrial, and TCU” land-uses and we abbreviate it as “commercial land-use” hereafter.

to address the following two research questions with empirical evidence: (1) do relatively new commercial land-use development show the same spatial agglomeration patterns as the current ones? and (2) how do residential land-uses react to commercial agglomerations in sub-centers?

This study contributes to the current literature on employment sub-center development within megacities (Boarnet *et al.*, 2017; Huang *et al.*, 2017; Zhong *et al.*, 2014) by providing a new perspective on production externalities and agglomeration patterns of commercial establishments. Agglomeration for enhanced production externalities does not solely favor CBD growth or increases in density. Rather, heterogeneity (temporal heterogeneity in our case) in commercial sectors can render some locations distant from existing CBD as the primary source for the generation of new or enhanced production externalities. We also show that residential trips apart from commuting have strong weights in deciding residential location choice when we evaluate performance and impacts of employment sub-centers.

The main practical contribution of this paper is that we examine the functional relationships between urban attractors and emerging commercial establishments as empirical evidence and reference for policy-makers and urban planners who deal with urban spatial structure planning, especially sub-center development. Traditional policy considerations on sub-center development and planning focus on issues such as jobs-housing balance (Hu *et al.*, 2018) and the ecological capacity of city centers (Czepkiewicz *et al.*, 2018). We provide a method to understand the business location preference of emerging economic sectors and their associated workers to locate in a growing city.

The structure of this paper is organized as follows; section 2 will review the current literature on spatial-related production externality while section 3 constructs our theory of how sectoral heterogeneity promotes employment sub-center development. In section 4, we propose several methods to examine and compare how newly allocated commercial and residential land-use in Chicago differs from existing patterns. Section 5 presents and discusses the results while section 6 offers some conclusions and future directions for research.

## 2. Literature review

### 2.1 Emergence of employment sub-centers

Joint forces from market and government policies stimulate the evolving urban structure towards growing employment sub-centers. From the perspective of public policy and urban planning, sub-centers are promoted as an efficient policy tool to decentralize population (Garcia-López and Muñiz, 2010; Huang *et al.*, 2017) and improve the standard of living with better environmental quality (Wang *et al.*, 2018).

[2] Another note is that land-use types in this paper are classified from satellite imagery. Thus we are agnostic about whether each land-use decision is made by landowners, government planning, or developers.

A decentralized CBD and compact employment sub-centers also reduce suburban workers' commuting time (Angel and Blei, 2016; Choi, 2018; Hu *et al.*, 2018; Zambon *et al.*, 2017; Zhao *et al.*, 2009). Richardson (1988) has posed that locally employed workers in sub-centers are better off than commuting workers. Huang *et al.* (2017) further identified a positive effect of employment centers on population distribution, where local residents are more likely to live close to sub-centers with the emerging maturity of a polycentric urban structure.

From the perspective of the spatial configuration of production, there has been a continuing focus on the emergence of employment sub-centers since Shukla and Waddell (1991), Waddell and Shukld (1993), Berry and Kim (1993), and Anas *et al.* (1998). A sub-center emerges through diminishing agglomeration diseconomies compared to the overcrowded CBD (Ahlfeldt and Wendland, 2013; Fujita and Ogawa, 1982; Fujita and Thisse, 2009) thus maximizing agglomeration externalities inside city-region against in contrast to increasing commuting costs (Anas *et al.*, 1998).

The spatial equilibrium choices of households and workers, with maximized utility benefits of agglomeration, were first discussed in the Alonso-Muth-Mills model in a monocentric city (Alonso, 1964; Mills, 1967; Muth, 1969), and then extended by Fujita and Ogawa (1982) to incorporate the impacts of additional urban sub-centers. Later modifications of urban structure and land-use models imply a departure from the determinant of agglomeration externalities by distance to CBD. Lucas and Rossi-Hansberg (2002), for example, pose a symmetric city structure model in which proximity to all other commercial employment determines positive production externalities of a firm. This model was further extended by Ahlfeldt *et al*. (2015) who incorporated discrete spatial units and modeled sub-center formation by considering asymmetries in locational fundamentals. Although these contributions introduced agglomeration and dispersion forces into the internal structure of cities, these theories of urban structure still imply a strong gravitational effect towards the CBD where the highest employment density is associated with the highest agglomeration economies.

### 2.2 Heterogeneity in agglomeration economies

The advantages of agglomeration within the CBD are widely corroborated through evidence for both developed countries (Ahlfeldt and McMillen, 2014; Glaeser, 2011; Ottaviano and Thisse, 2004; Zhang *et al*., 2018) and emerging economies (Yang *et al.*, 2016; Zhang *et al.*, 2017). However, the recent literature has challenged traditional CBD-based agglomeration with evidence that production externalities differ between commercial sectors. Neighboring commercial establishments may not necessarily create positive production externalities. The relationship can be competitive or complementary (Chung and Hewings, 2015), related or unrelated (Firgo and Mayerhofer, 2017).

Specifically, Henderson *et al*. (1995) revealed that rapidly evolving high-tech industries favored externalities from diversified nearby industries, while mature industries, in contrast, enjoyed the positive effects from specialization. The diversity of economic activity is found to yield higher economic and employment growth

(Glaeser *et al.*, 1992) and encourage births of new establishments (Rosenthal and Strange, 2003). A broader literature review on the types of agglomeration economies could be found in Duranton and Puga (2004) and Rosenthal and Strange (2004).

Heterogeneity in agglomeration of this kind contributes to a new understanding of spatial externalities beyond merely scale economies and the emergence of sub-centers with sectoral incentives for higher economic benefits. Nonetheless, most studies focusing on sub-center evolution have not incorporated the production heterogeneity perspective. Furthermore, the limited literature in this field tend to focus on the city-regional scale (Chung and Hewings, 2015; Lucas and Rossi-Hansberg, 2002) and on static agglomeration effects (see e.g., Melo *et al*., 2009; Rosenthal and Strange, 2004), while the investigation into a finer spatiotemporal granularity into urban commercial and residential land-use growth patterns is often hampered by insufficient supporting empirical evidence.

## 3. Theoretical framework and hypotheses

This paper aims to bridge the research gap by examining how the temporal heterogeneity between existing sectors and emerging commercial sectors leads to the development of multiple employment sub-centers that may influence residential location choices in the Chicago metropolitan region. We provide a method of analysis down to 30x30-meter (about patch-level) resolution for commercial and residential land-use development with two periods (before and after 2001).

The theoretical framework is shown in figure 1. The agglomeration of the existing CBD consists of commercial establishments that share complementary relationships (positive spatial externalities) with each other. However, when technology and economic structures change, new and emerging commercial establishments may not share positive spatially-related production externalities with existing firms. As a result, they do not have any incentive to pay the higher land-rent in the existing CBD, and these firms may choose to agglomerate in an employment sub-center. This formation of employment sub-centers has far-reaching effects on urban structure, economic development, and residential location choices, as will be discussed later in this paper.

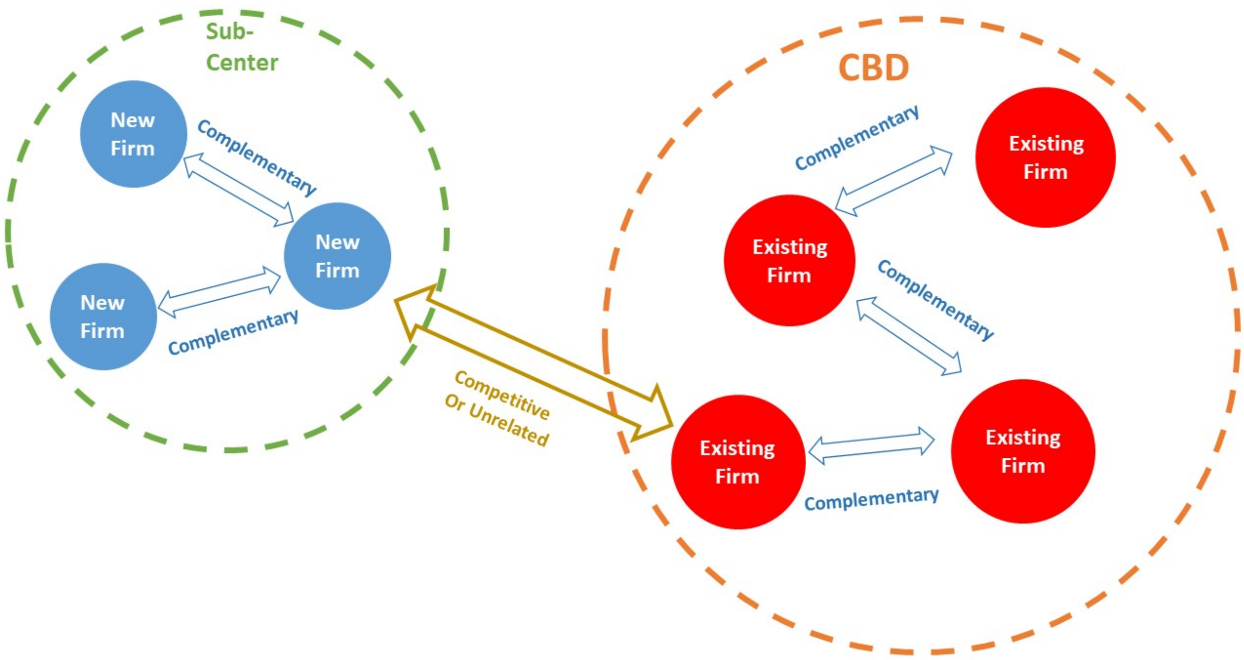



**Fig. 1.** Spatial externality heterogeneity that shapes employment sub-centers.

**Hypothesis 1: The heterogeneity of spatial externalities of commercial sectors promotes employment sub-center formation.**

Our first hypothesis is that due to the heterogeneity within commercial sectors, the CBD may not be the primary source of production externalities for certain economic sectors; as a result, there is a strong economic incentive for some firms to agglomerate in sub-centers with more firm-specific desirable production externalities. To examine this hypothesis, we propose a methodology that examines how new commercial land-use (post-2001) development in Chicago differs in comparison to existing commercial land-use development. The methodology focuses on much smaller spatial units of analysis (181 million cells) providing more flexibility in the analysis of intra-metropolitan structural dynamics. We measure the connectivity of all commercial land-use cells to existing population and job centers, quality-of-life amenities, and highways/major roads. Thereafter, we examine whether newly allocated commercial land-use cells show significant structural breaks from existing land-use cells. Our expected outcome for hypothesis 1 is that new commercial land-use allocated before and after 2001 would show a different pattern of relationships between connectivity to major urban attractors. This break in trend is motivated by the difference between the valuation of spatial externality sources between existing economic sectors and emerging economic sectors (post-2001).

**Hypothesis 2: Amenities are an equally (if not more) important factor in residential location choices in comparison to job trips.**

Our second hypothesis focuses on the residential land-use developmental patterns with regards to the newly formed commercial sub-centers. We suggest that amenities (such as restaurants, museums, and public spaces) serve as a more important factor in residential location choices, which need at least equal (if not more) emphasis when we discuss the influence of new employment sub-centers on the formation of new commuting catchments. In other words, we expect that neighborhood amenities that enhance the quality of life would have a strong influence on residential land-use allocation.

It is noteworthy that this study focuses on identifying land-use growth pattern after one time point and understanding how it reacts to urban agglomeration economies differently than previous patterns. It does not apply sub-center detection methods for multiple time periods such as McMillen (2001) and Giuliano and Small (1991). Instead, this study locates new employment sub-centers by finding new pattern of land-use growth in relation to agglomeration economy pattern, which is demonstrated in **Section 5.3**.

# 4. Study area, data, and methodology

## 4.1 Chicago Metropolitan Statistical Area: land-use and land-use change maps

The case study of this paper builds on McMillen and Lester (2003)'s study which shows that employment of all sectors in employment sub-center in the Chicago metropolitan region has grown from 6.8% to 25.9% from 1970 to 2000. We use 2001 land-cover data from the National Land Cover Database (NLCD) to identify residential, commercial, and undeveloped land-use types in Chicago. The developmental intensity classified by NLCD is overlaid on Google maps for manual data cleaning. The cleaned and classified 2011 NLCD is used for identifying existing land-use in Chicago enabling the identification of the difference between 2001 and 2011 land-use and thus urban commercial and residential land-use growth/change over this 10-year period. The existing (2001) land-use and land-use change (2001-2011) maps are shown in figure 2. For the land-use change map, already developed areas and no-growth zones (including forest preserve, flood zones, and parks) are not included in subsequent analysis.

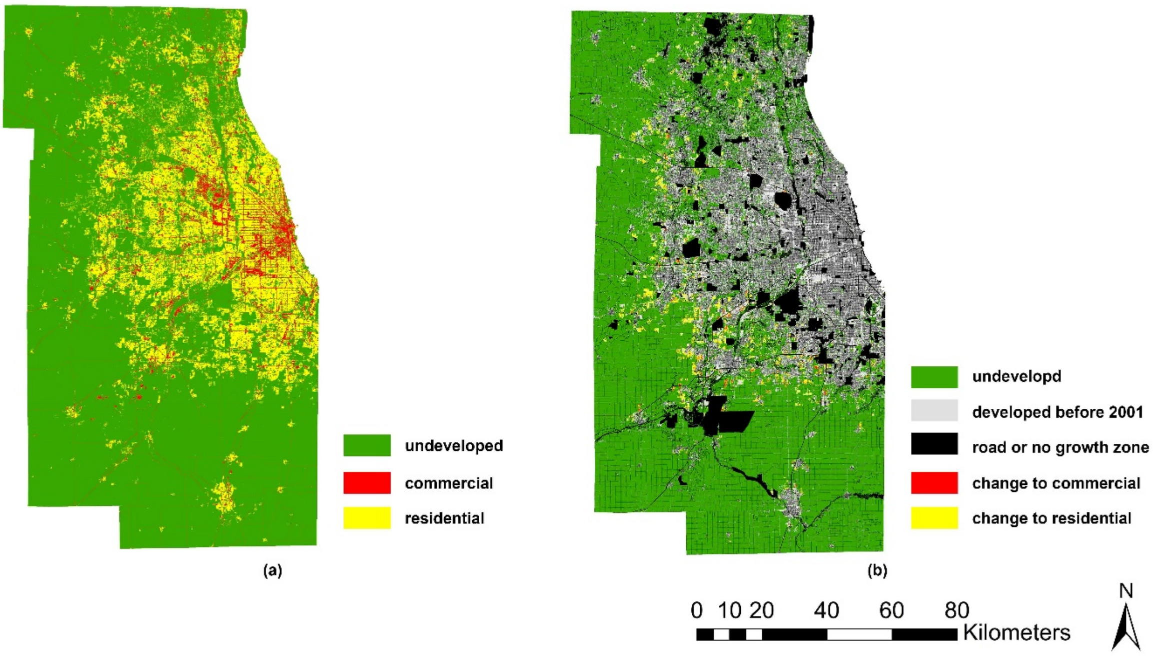



**Fig. 2.** (a) Existing land-use map of Chicago. (b) Land-use change map of Chicago from 2001 to 2011.

### 4.2 Mapping connectivity of urban attractors

To examine factors that influence commercial and residential land-use developmental processes, we map connectivity of each land-use cell in Chicago to all the four attractors (population, job, quality-of-life amenities, and transportation networks). The Chicago regional attractors we use in our analysis include: population centers, highways, major roads, and points of road network access—on-ramps, major intersections from the 2010 U.S. Census. Employment centers are obtained through D&B Hoover Industry Directory with total number of employees in each establishment. Using these data, we evaluate the connectivity to "the number of employees" with distance decay. This is very similar to Lucas and Rossi-Hansberg's (2002) methods in incorporating employment intensity into land-use studies. Quality-of-life amenities are represented by the points-of-interests (POIs) with commercial establishments with customer reviewers (including restaurants, shops, parks, hotels, and transport stations) using a Yelp API to obtain geotagged reviews for more than 10,000 restaurants, bars, stores, public and private services, hotels, and real estate purchases as of March 2016. As a result of the inability to access historical data from Yelp, the time point for Yelp data results are different than other dataset. It should not be a major problem as we do not conduct a longitudinal study. Rather, it focuses on the urban structure pre- and post-2001.

The most difficult computational challenge is to find the shortest route from 181 million cells to 10,000+ urban attractors to calculate connectivity. We optimize the computation with parallel computing techniques using a Stochastic Greedy Algorithm (SGA). The brief idea of SGA is illustrated in figure 3 and the pseudo code of the algorithm is attached in **Supplementary Materials Part 1**. First, the algorithm tries to find the local optimum (optimizing only one-step ahead at every movement step) from one land-use cell to one center using a greedy algorithm. In a greedy algorithm for path-finding, an agent tries to find the route to the destination by moving along the fastest road ahead on the correct direction for every time-step. However, the agent has only vision of 1 cell and cannot see the possibility of a detour onto a highway and use shorter total travel time for the whole route. In the improved version of SGA, we disseminate 1,000 agents from each of the population, employment, and quality-of-life attractors in the Chicago regional study area using high-performance paralleled computing resources. Each agent has been directed to move as far as it can travel in 1 cell increments at each time-step towards the destination. Direction is probabilistic to all adjacent cells with higher weights assigned on directions in a straight-line towards the destination or to roads with higher speed limits. It resembles human pathfinding in an environment without the aid of GPS devices. Each of the paralleled runs of the greedy algorithm is assigned a randomized decision rule (the blue dashed lines in figure 3). This provides the algorithm a chance to “jump” out of a local optimization to reach a globally optimal solution, as the algorithm considers both moving along a direct path to the destination as well as moving along the fastest route (such as via highways) based on probabilistic outcomes.

The optimum route (red dashed line in figure 3) is chosen from the algorithm for every land-use cell to each urban attractor and a gravity-like function calculates the aggregated connectivity to each attractor for each land-use cell. The implementation is realized through the open-source code on GitHub (Pan *et al.*, 2018) implemented on a parallel computing facility. Parallelization and randomization allow SGA to achieve a balance between computational performance and efficiency.

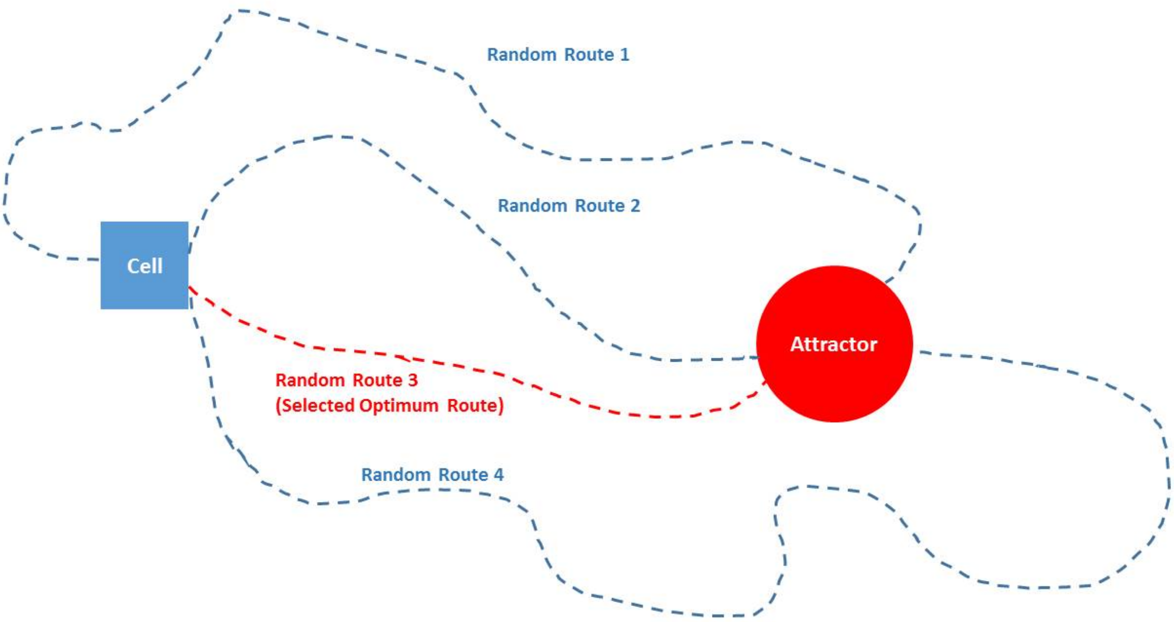



**Fig. 3. Illustration of SGA.**

One of the main groups of literature that try to improve from the early urban structure models (including Alonso, 1964; Fujita and Ogawa, 1982; Lucas and Rossi-Hansberg, 2002; Mills, 1967) is to use a grid-like urban road-network instead of measuring travel time solely by Euclidean distance (Dong and Ross, 2015; Tsekeris and Geroliminis, 2013). However, new evidence found by complex urban network studies (Batty, 2008) shows that actual urban networks have sharp difference to homogenized gridded networks. SGA method in our paper improves the ability to capture empirical realities with the ability to calculate shortest travel time from any land-use location to urban attractors through actual city networks. Also, the method is applicable to any urban network and land-use thus it is not just an *ad hoc* measure for our case study.

We also acknowledge two potential improvements for the SGA methods. First, public transportation plays an important role in commuting and can be included as alternative route with a different cost parameter to current SGA methods. Second, our SGA directly uses the number of population (or employment) in urban attractors as weights, while there are previously intense discussions to the weighting parameters of gravity-like models (Anderson and van Wincoop, 2003). Whether this turns out to be a limitation will be the focus of future work.

### 4.3 Comparison of current and future residential and commercial location choice patterns

The two hypotheses of this study address the potential break in the temporal externality pattern for both commercial establishments and residents in Chicago. To evaluate these hypotheses against assumptions of a dominant CBD-based urban structure or time-invariant urban structure and land-use models, we relate frequencies of commercial and residential land-use cells to each quantile value (50 in total) of the connectivity values of each attractor obtained in section 4.2. The functional forms of these relationships are not pre-selected. Instead, we use a “leave-one-out” cross-validation method to select one of the candidate functions (linear, bell, or cubic) that best depicts the relationship between land-use frequencies and connectivity to attractors.

The “leave-one-out” cross-validation method is used to allow higher order polynomial functional relationship between land-use frequencies and attractor values to be selected (Kohavi, 1995). Higher degree polynomials can explain the mechanism of land-use location selection when land tenants’ preferences of connectivity to attractors are not monotonic. For example, it is intuitive to think that residents do not want to live too close (due to noise) or too far away (due to longer commuting time) from a highway. In this case, the functional relationship that depicts frequencies of residential land-use to transportation network should be a bell curve that peaks in the middle of the attractor-value region.

In the “leave-one-out” cross validation process, we fit all three (linear and two higher polynomial forms) possible curves 50 times with the dataset, leaving one land-use frequency value out each time. We then try to predict the left-out data point using the fitted function and record the sum-of-squared errors of its prediction. We average the absolute mean errors for each functional candidate and choose the function with the smallest errors to be our final function. The merit of this method over traditional *R*-squared is that the cross-validation can compare models with different degrees-of-freedom or even with re-scaled data (Kohavi, 1995).

We conduct a similar process for land-use change frequencies in Chicago from 2001 to 2011. The fitted functional form results for land-use change are compared to the results for existing land-use to check our assumption: whether temporal break exists in locational preference between existing commercial land-use and newly allocated commercial land-use in regard to job, labor, quality-of-life amenities, and transportation network accessibility. Once again, the implementation is realized through the open-source code on GitHub (Pan *et al.*, 2018).

### 4.4 Future employment sub-centers

One of the practical contributions of this study is to understand likely locations of future employment centers. To explore this option, we apply the functional relationship between commercial land-use change frequencies and connectivity to various attractors to all lands in Chicago and find “hot-spots” for future employment sub-center development in undeveloped lands (the green areas in figure 2) in Chicago. The results are shown on a probability map with a 0-1 scale for these areas, and we will pick areas with higher probability of future sub-center development.

## 5. Results

### 5.1 Connectivity maps

Figure 4 shows the connectivity map of a) employment centers; b) population centers; c) quality-of-life amenities, and d) transportation networks. We observe that connectivity patterns to population and employment centers have a common structure — the highest connectivity occurs in the urban CBD and extends along major roads and highways. Employment centers generate similar patterns, although with less pronounced effects, while population centers show a more dispersed spatial distribution. Quality-of-life amenities are the most spatially dispersed and the major road structure is very prominent in the quality-of-life amenities’ connectivity map. The similarity shared between connectivity to population, employment centers and quality-of-life amenities is that the downtown of Chicago has a higher overall connectivity, and it is dispersed along road networks, especially highways. The main difference is that population connectivity is the most dispersed while quality-of-life connectivity shows a northward shift compared to employment connectivity. This shift of center is likely caused by a retail and tourist magnet north of the city center (the Michigan Avenue “Magnificent Mile”). Whether these externalities contribute positively or negatively (or with more complicated functional forms) to the location choices of residential and commercial land-use, and how their impacts are differentiated provides the main focus of further investigation in Section 5.2.

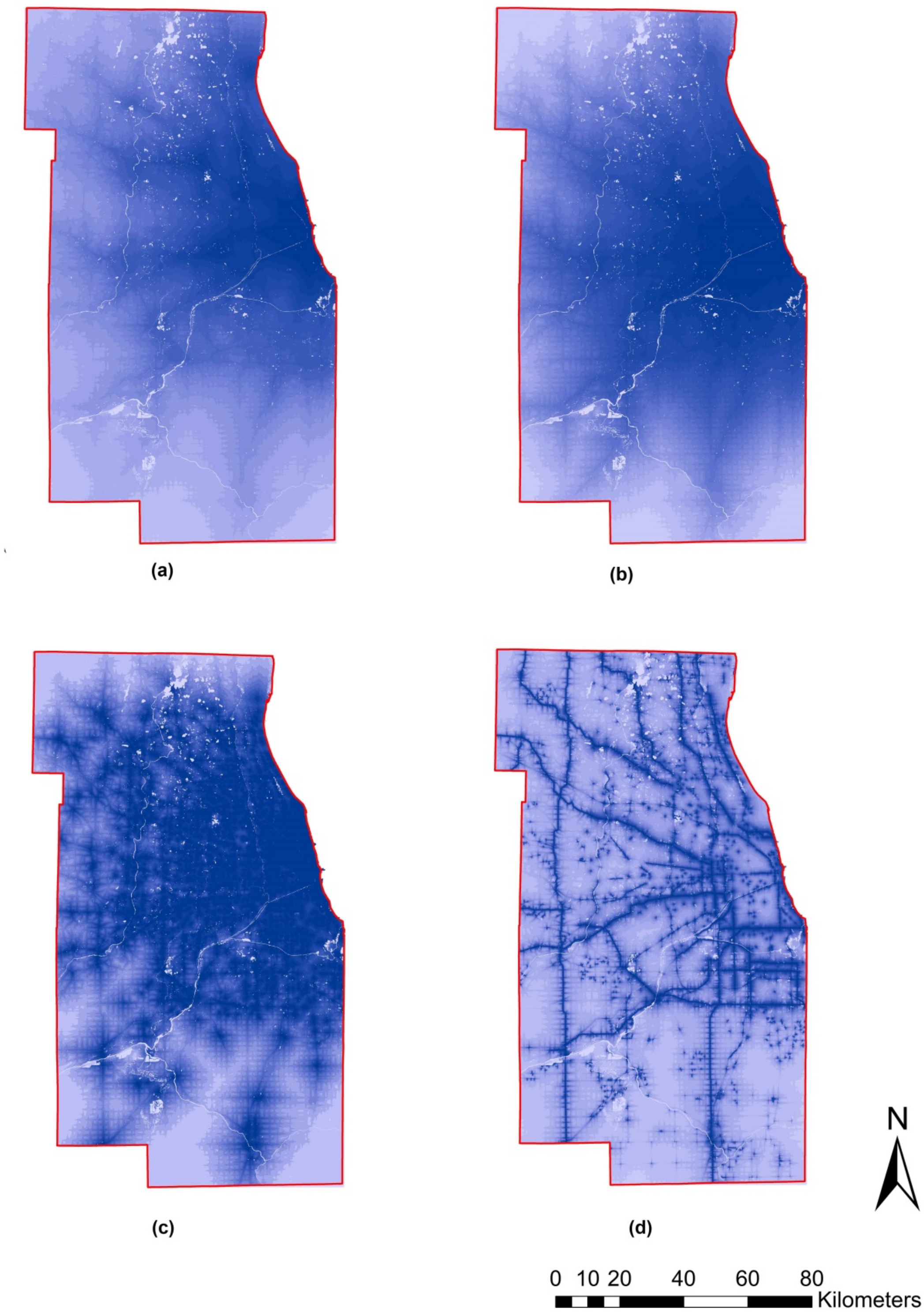


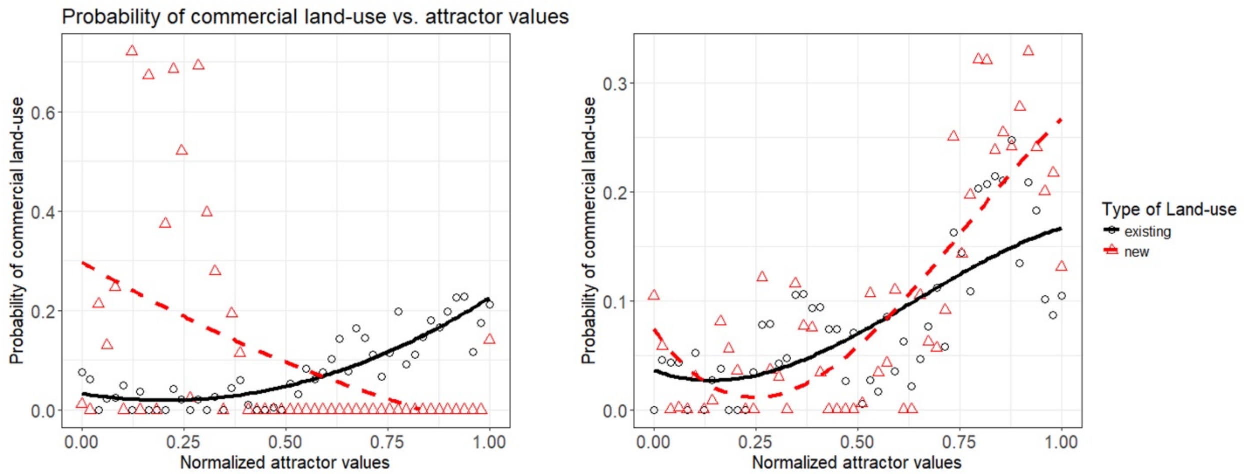

**Fig. 4. Connectivity maps of land-use cells to 4 urban attractors: a) employment; 2) population; 3) quality-of-life amenities; 4) transportation network. Darker color indicates higher connectivity**

### 5.2 Comparison of existing land-use and new growth

In this part, we try to find the best functional forms to fit the relationship between frequencies of commercial/residential land-use (change) and values of connectivity to the four types of urban attractors. The “leave-one-out” cross-validation results determine the selection of the best functional form. We record the averaged “leave-one-out” mean absolute error of each functional form in table 1.

**Table 1.** “Leave-one-out” cross-validation error of different.

| **Existing Land-use** | | | | | |
|---|---|---|---|---|---|
| | **Commercial Land-use** | | **Residential Land-use** | | |
| | **Employment** | **Transportation** | **Employment** | **Amenities** | **Transportation** |
| **line** | 0.02955 | 0.04139 | 0.05236 | 0.05638 | 0.08254 |
| **bell** | **0.01763** | 0.04263 | **0.01357** | 0.00708 | 0.09839 |
| **cubic** | 0.02047 | **0.03454** | 0.02779 | **0.00086** | **0.04134** |
| **Land-use Change** | | | | | |
| | **Commercial Land-use** | | **Residential Land-use** | | |
| | **Employment** | **Transportation** | **Employment** | **Amenities** | **Transportation** |
| **line** | 0.06567 | 0.01502 | 0.06120 | 0.10270 | 0.05540 |
| **bell** | **0.02110** | 0.01193 | **0.04380** | 0.10185 | 0.04389 |
| **cubic** | 0.06514 | **0.00606** | 0.06233 | **0.06306** | **0.03146** |




**(a) Employment Connectivity** **(b) Transportation Network Connectivity**

**Fig. 5.** Frequencies of commercial land-use (black solid line) and land-use change (red dashed line) vs. connectivity to attractors.

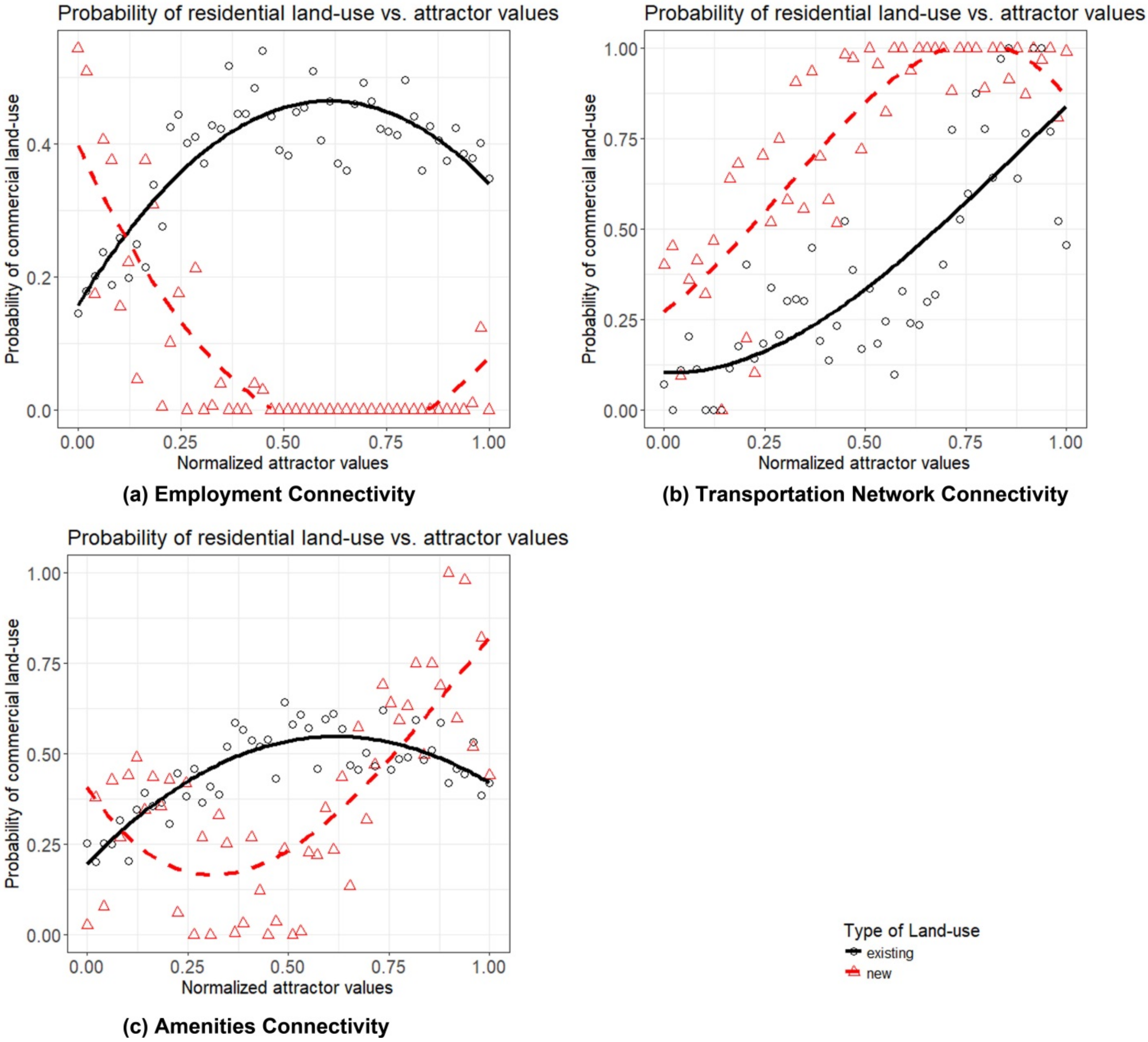



**Fig. 6.** Frequencies of residential land-use (black solid line) and land-use change (red dashed line) vs. connectivity to attractors.

Referring to figure 5, we can see that a temporal break exists and is very obvious in the spatial externality patterns of employment connectivity measured by commercial land-use frequencies. Existing commercial land-uses are generally attracted by better connectivity to employment centers, but the relationship becomes negative when commercial land-use change after 2001 is examined (figure 5a). This is a major theoretical extension to Lucas and Rossi-Hansberg (2002)'s theory — better connectivity to employment is a source of positive production externalities for existing commercial land-use, but it turns negative for new commercial land-use distribution. On the other hand, both old and new commercial land-use frequencies appear to be positively correlated with connectivity to transportation networks (though not strictly monotonic, figure 5b), which suggests that proximity to major roads and highways is the time-invariant attractor for commercial land-use growth.

In terms of residential land-uses, we can also observe some results, shown in figure 6 that offer departures from prior urban structure theories. For example, existing residential land-use frequencies peak at about 60 percentiles of all value in regions of employment connectivity, a finding that conforms to the majority of findings in the existing literature that residents prefer places with shorter commuting time (figure 6a). However, new residential land-use growth frequencies drop to 0 in the range of the 50 to 80 percentile value region, which means finding places with better job accessibility has become less of a concern for new residential development. Figure 6b shows that transportation network connectivity appears to be the most important determinant of where old and new residential land-use location choices are made, though new development appears not want to be too close (frequency drops after the 90th percentile). Better connectivity to quality-of-life amenities (figure 6c) is also shown to be generally attractive for old and new residential lands, while the functional forms vary and suggest a change of preferences over the period 2001-2011.

Table 2 (old and new commercial) and table 3 (old and new residential) demonstrate the coefficients of higher order polynomials of each function and the inflection points of the curves shown in figures 5 and 6. These results highlight non-monotonicity of some relationships between connectivity and land-use frequencies. The results may indicate some heterogeneity among commercial and residential land-use types, or the results may also signal some important factors that are missing from this study.

**Table 2.** Coefficients and monotonic regions of commercial land-use models.

| | |
|---|---|
| **Existing Commercial/Employment** | |
| | **Coefficients** |
| **intercept** | 0.0326 |
| **x** | -0.1334 |
| **x^2** | 0.3261 |
| **Monotonic Support 1** | 0-20 percentile (decreasing) |
| **Monotonic Support 2** | 21-100 percentile (increasing) |
| **New Commercial/Employment** | |
| | **Coefficients** |
| **intercept** | 0.2968 |
| **x** | -0.4676 |
| **x^2** | 0.1349 |
| **Monotonic Support 1** | 0-100 percentile (decreasing) |
| **Existing Commercial/Transportation Network** | |
| | **Coefficients** |
| **intercept** | 0.0366 |
| **x** | -0.1431 |
| **x^2** | 0.5657 |
| **x^3** | -0.2926 |
| **Monotonic Support 1** | 0-14 percentile (decreasing) |
| **Monotonic Support 2** | 15-100 percentile (increasing) |
| **New Commercial/Transportation Network** | |
| | **Coefficients** |
| **intercept** | 0.0476 |
| **x** | -0.2044 |
| **x^2** | 0.4502 |

| | |
|---|---|
| **Monotonic Support 1** | 0-22 percentile (decreasing) |
| **Monotonic Support 2** | 23-100 percentile (increasing) |

**Table 3.** Coefficients and monotonic regions of residential land-use models.

| **Existing Residential/Employment** | |
|---|---|
| | **Coefficients** |
| **intercept** | 0.1580 |
| **x** | 1.0040 |
| **x^2** | -0.8216 |
| **Monotonic Support 1** | 0-61 percentile (increasing) |
| **Monotonic Support 2** | 62-100 percentile (decreasing) |
| **New Residential/Employment** | |
| | **Coefficients** |
| **intercept** | 0.3979 |
| **x** | -1.3043 |
| **x^2** | 0.9848 |
| **Monotonic Support 1** | 0-66 percentile (decreasing) |
| **Monotonic Support 2** | 67-100 percentile (increasing) |
| **Existing Residential/Transportation Network** | |
| | **Coefficients** |
| **intercept** | 0.1024 |
| **x** | -0.0378 |
| **x^2** | 1.2280 |
| **x^3** | -0.4539 |
| **Monotonic Support 1** | 0-1 percentile (decreasing) |
| **Monotonic Support 2** | 2-100 percentile (increasing) |
| **New Residential/Transportation Network** | |
| | **Coefficients** |
| **intercept** | 0.2726 |
| **x** | 0.8180 |
| **x^2** | 1.5629 |
| **x^3** | -1.7883 |
| **Monotonic Support 1** | 0-78 percentile (increasing) |
| **Monotonic Support 2** | 79-100 percentile (decreasing) |
| **Existing Residential/Amenities** | |
| | **Coefficients** |
| **intercept** | 0.0000 |
| **x** | 1.1564 |
| **x^2** | -0.9766 |
| **x^3** | 0.0440 |
| **Monotonic Support 1** | 0-62 percentile (increasing) |
| **Monotonic Support 2** | 63-100 percentile (decreasing) |
| **New Residential/Amenities** | |
| | **Coefficients** |
| **intercept** | 0.4058 |
| **x** | -1.6488 |
| **x^2** | 3.1469 |

| | |
|---|---|
| **x^3** | -1.0811 |
| **Monotonic Support 1** | 0-31 percentile (decreasing) |
| **Monotonic Support 2** | 32-100 percentile (increasing) |

There could be several reasons for non-monotonicity in the relationships between land-use (change) frequencies and connectivity to various attractors, and some reasons suggest limitations of our study. First, price (land rent) is not considered in this study. It can serve as a disincentive for development in some highly connected locations thus generating non-monotonic results. Secondly, other negative externalities apart from economic externalities can affect location decisions. For example, residents may not want to locate too close to highways due to noise and air pollution, and this is found in the frequency function of new residential land-use (the frequency drops over 63 percentiles in connectivity to transportation networks). Thirdly, heterogeneity within commercial and residential sectors can lead to a different locational preference related to connectivity levels. For example, Chen and Rosenthal (2008) find that elderly people do not prefer locations close to business centers, while younger people do.

### 5.3 Chicago's future

We provide a method to forecast the implications of agglomeration patterns for the formation of newer employment sub-centers. It can be assumed that land-use development planning and decisions predate allocation of new employment because places are first needed to accommodate workers. Thus, based on the existing connectivity to spatial externalities, we extrapolate the functional relationship between frequencies of commercial land-use and connectivity and calculate the potential of new commercial development for undeveloped lands in Chicago and obtain the resulting probability map of newer land-use development illustrated in figure 7. This can be viewed as a "one step ahead" commercial and industrial development for Chicago. This method of identifying new sub-center formation differs from previous methods (including McMillen, 2001 and Giuliano and Small, 1991). Instead of understanding sub-center using "clusters" of commercial or residential land-uses, this study identifies most likely new land-use development and sub-center formation to the agglomeration pattern. From the results in Section 5.2, it can be seen that frequencies of new commercial land-use growth occur at 10-30th percentile of employment connectivity and 75-90th of transportation connectivity. Figure 7 can be seen as identifying land locations that satisfy both criteria for newly or potentially growing employment sub-centers.

For the purpose of liking to previous literature using traditional sub-center identification method, we circle out in the sub-centers in Chicago as identified by McMillen (2001) methodology as a comparison. Our generally approach is that we focus on the patterns of land use growth in a continuum, rather than simply those nodes that reach a particular threshold. Emerging sub-centers may not reach a particular size at a given point in time, but that kind of emergence is important to highlight. We add value to the literature by identifying places with momentum to become sub-centers in the future, and they deserve attention by prospective promoters in a commercial context, and by planners in a regulatory and development strategy context. It can be observed that three “insignificant” sub-centers (Arlington Heights, Bolingfield, and Plainfield) by the criteria of McMillen’s (2001) method do appear to have the momentum to become significant sub-centers, while an already “significant” sub-center (the O’Hare Airport area) will continue its current trend of fast development.

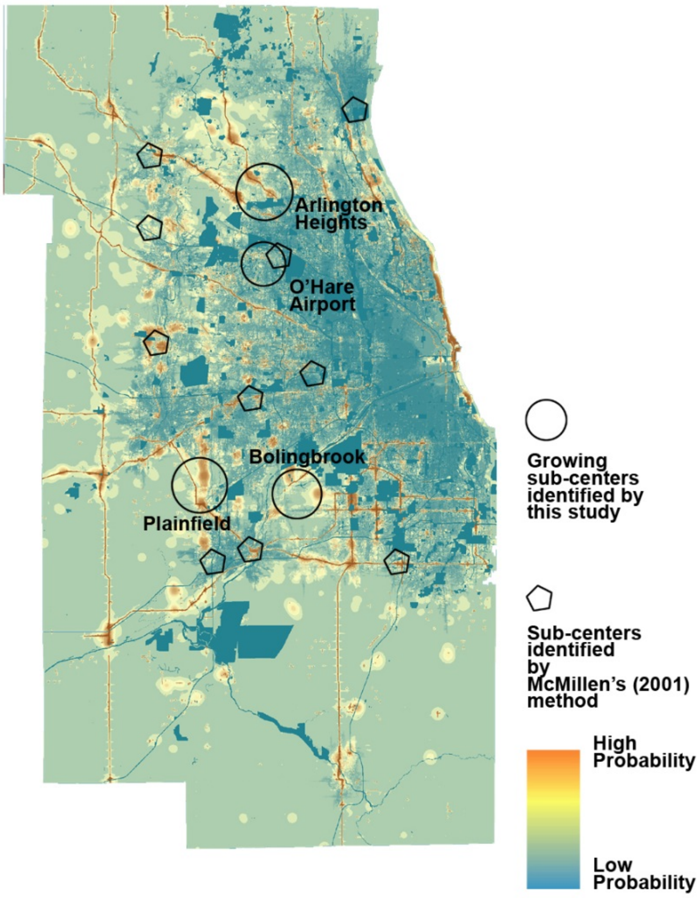



**Fig. 7. Future commercial developmental probabilities for Chicago with sub-centers identified by this study and McMillen’s (2011) study.**

We have identified four types of commercial new developmental patterns. First, redevelopment of vacant lands occurs on available lands with the best connectivity to the existing CBD. In traditional economic theories, these are places with the best source of spatial externalities. However, with the new patterns we have identified for new commercial land-use location choices, these types of development become very scarce as predicted. Secondly, we can see some major development as extensions to existing major employment sub-centers. On the northwest edge, development is predicted in Arlington Heights areas as extensions to major sub-centers around O’Hare International Airport. In the southwest urban edge, the Bolingbrook area is predicted to have strong growth as a continuation of dynamics associated with the burgeoning Plainfield area. The third and fourth developmental patterns we identified offer chances for emergence of new employment sub-centers. The third pattern is growth along highways. Some new linear growth along highways can result in growth of a large area around it with firms that have similar spatial externalities. The fourth and final type is more random with the “emergence” of high probabilistic employment sub-centers “out of nowhere.” In figure 7, some of those emerging locations can be observed in the southern Chicago Heights region; the non-monotonicity we find in section 5.2 provides the main reason for these outcomes.

Furthermore, the growing popularity of shared workspaces (e.g., WeWork) has its potential to accelerate the future growth of sub-centers as currently identified. With increasing trends in commuting reduction and community engagement, the decentralization trends of traditional CBD areas are expected to be enhanced, especially for promoting entrepreneurship in sparse regions (Fuzi, 2015). Although a substantial proportion of shared co-working hubs are currently located in the central city area (Wang and Loo, 2017), only a few of them favor the traditional CBD due to their pursuits of specific economic activities (e.g., university knowledge spillovers, local community engagement, etc.) and lower rent prices.

## 5.4 Discussion of results

### *5.4.1 Empirical results and policy implications*

The findings of our study have strong implications for the emergence of employment sub-centers in mega-cities. Our findings validate hypothesis 1 that the nature and influence of spatial externalities for firms may be evolving and changing in terms of the magnitude of their impacts and the influence of spatial location and connectivity factors. Thus, agglomeration in the existing CBD may not be the primary source of spatial production externalities for new firms. This extends the sub-center formation literature (see e.g., Ahlfeldt et al. (2015); Fujita and Thisse (2009) where a strong gravitational effect is still identified towards the CBD, and provides empirical evidence with finer spatial detail for examining diversification externalities (see e.g., Jacobs (2016). The temporal break in agglomeration patterns can be a strong rationale for employment sub-center growth and planning for mega-cities. The most important factor in this case is the structure of the transportation networks—connectivity to highways and major roads instead becomes the most important factors for new firm locations. With the current transformation in the delivery of retail goods and services - from traditional bricks and mortar locations to online purchases, the structure of the transportation networks will clearly play a more critical role in the delivery network.

Another finding that validates our hypothesis 2 is that residents appear to give the highest weight to quality-of-life amenities in choosing where to live, and quality-of-life amenities are spatially more disperse as shown in figure 4, further challenging the incentives provided by CBD agglomeration. This confirms the findings of literature that emphasizes the importance of quality-of-life amenities for residential location choices (see e.g., Chen and Rosenthal (2008) and Kuang (2017).

There are previous Chicago-centered studies (e.g. McMillen and McDonald (1998); Mcdonald and Mcmillen (1990); McDonald and McMillen (2000); McMillen and Lester (2003)) that point out re-centralization patterns (McMillen, 2003) or a decline in the numbers of significant employment sub-centers (McMillen and Lester, 2003). These studies apparently challenge some of our findings. Two explanations are possible. The first is the difference in temporal period of examination. McMillen (2003)'s study is mainly based on data available pre-millennium in contrast to our focus on pre-/post-millennium comparison. McMillen and Lester (2003)'s "decline of sub-centers post-2000" argument is a forecast study along the line of Northern Illinois Planning Commission's regional planning vision. Our spatial-explicit evidence post-2000 is indeed a fact check for the planning vision at 2000. Another possible explanation is the difference in approaches and data. McMillen (2003)'s study uses land values while McMillen and Lester (2003)'s study uses density as the major identifications for urban development. Our paper uses total land-use growth, which does not incorporate density and price. These three papers with different limitations serve nicely as complementary studies to each other.

In figure 8, we draw a density change in Chicago with regards to distance from CBD in the year 2002, 2007, and 2012 using data from the U.S. Economic Census as complementary evidence to our study. It can be argued that purely from an employment density perspective, Chicago has experienced a re-centralization in the millennium as density in the CBD continually increased while places far from the CBD (from 80-100km away from CBD) experiences employment density decreases in the same period. On the other hand, employment density in 25km to 50km distance from CBD experienced some fluctuations in the same period, likely caused by land-use development and change that occurred in some main sub-centers discussed in this paper, including O'Hare International Airport and Naperville.

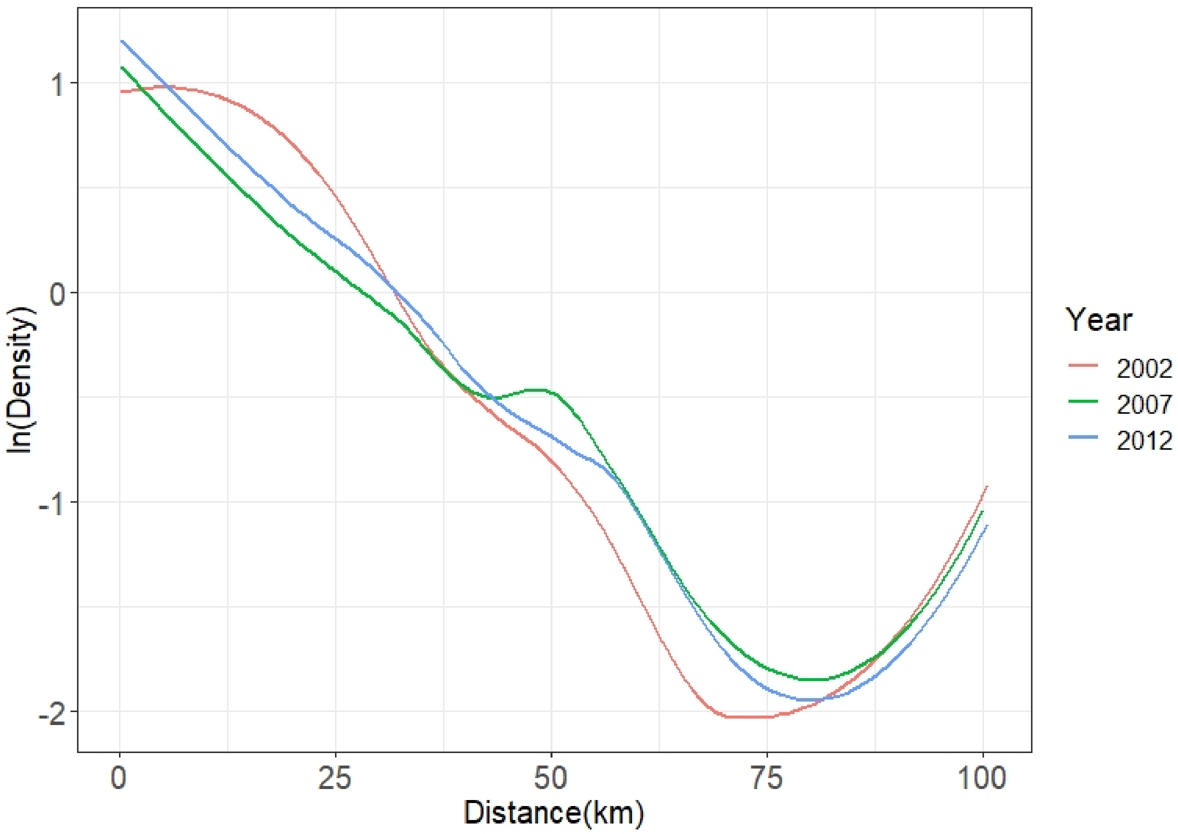



**Fig. 8.** The employment density change with regard to distance to CBD in Chicago from 2002-2012

As the decentralization of jobs from the CBD to suburbs has become a global trend (Angel and Blei, 2016; Gordon *et al.*, 1986), the findings and policy implications from Chicago are informative beyond the local context. First, we offer an alternative way to understand urban land-use planning and regulation for employment sub-centers in megacities beyond the current prevailing considerations of jobs-housing balance and ecological/resource capacity. Local planners should be cautious about the recent firm locational choices and market responses to sub-center policies, with a continuous monitoring of the changing agglomeration patterns in support of land-use planning and regulation.

Secondly, our findings show considerable heterogeneity in agglomeration economies accompanied by the sub-center growth, with one of the first attempts to use highly-detailed land-use data. While we still need other findings to explain the causes of the heterogeneity, the analysis here indicates in the future studies it is important to incorporate such land-use data particularly for cities which are known to be affected by the diversity and variability of their social makeups. As displayed in the divergent patterns between newly-developed and existing commercial land plots, decision-makers should also make the best use of the positive spatial-related production externalities generated by existing local industries. This should be firmly noted in future sub-center policy design especially for those rapidly-urbanizing megacities in contemporary East Asia, as their polycentric development are fueled by increasingly vast industrial park investments and industrial sector relocations (see e.g., Zhao *et al.*, 2011; Zheng *et al.*, 2017).

Thirdly, quality-of-life amenities should be highly valued in the polycentric development. Accessibility to local amenities and emergence of residential land-use development prove to have high functional relationship, the understanding of which will help to counter the negatives associated with longer commuting conditions and jobs-housing imbalance.

*5.4.2 Endogeneity, casual relationships. longitudinal studies*

Similar to the earlier models from Alonso (1964), Fujita and Ogawa (1982), Muth (1969), Mills (1967), and Lucas and Rossi-Hansberg (2002), land-use is distributed simultaneously with connectivity to employment centers among other centers in this paper. By such analysis we describe and analyze the change in urban structure of Chicago pre- and post-2001 and explain its implications for employment sub-center formations. One major limitation to this approach is that the causality between land-use change and the establishment of major urban attractors (population, employment, and quality-of-life) cannot be identified. An important issue regarding the measurement of the marginal effect of connectivity to urban attractors in shaping urban land-use structure is the possible endogeneity between land-use growth and urban attractor evolution. This assumption is based on the fact that some urban attractors may be developed because new employment sub-centers are planned. If that is the case, the current model would fail to include the planned commercial development that yet have a correlation with both existing commercial growth or urban attractors, and it would incorrectly allocate the commercial land-use growth effect from planned growth to the connectivity to urban attractors.

To avoid this issue, Duranton and Turner (2012) develops an Instrumental Variables (IV) approach to study the effects of interstate highway stock on employment growth for the U.S. cities. Though econometric models and longitudinal studies are not the goal of this paper, we apply a simplified version of Duranton and Turner's (2012) approach by using the two-stage least squares (2SLS) estimator to ordinary least squares (OLS) estimates to test endogeneity issue. Hausman-Wu and Wald F-test as suggested by Chen and Haynes (2015) for studies on transport infrastructure and regional growth are taken to provide diagnostics for endogeneity issue. More precisely, we use economic places in Chicago that have less than median number of employment density in the 2002 American Economic Census as areas with high potential of commercial land-use development, and the key assumption of this model is that higher employment growth during 2002-2012 in these low employment density areas suggests higher probability of commercial land-use development. The choice of years 2002 and 2012 is because they are the years with available American Economic Census that best overlap with our study period. From figure 5(a) we can assume that land-use development in regions with lower employment density are least affected by the existing employment density, while employment density is intuitive to be highly correlated to employment attractor connectivity.

The OLS model is specified in Equation 1.

$$\Delta y_i = \beta_1 a_i + x'_i \gamma_0 + u_i \tag{1}$$

where $\Delta y_i$ is the employment growth at cell $i$ from 2002-2012; $x'_i$ is the matrix that includes the 4 urban attractors measures used in this study; $u_i$ is an independent Gaussian error term with mean 0. For comparison, a 2SLS-IV model is constructed as:

$$\hat{a}_i = \beta_2 z_i + x'_{i,t} \gamma_0 + v_{i,t} \tag{2}$$

and

$$\Delta y_i = \beta_1 \hat{a}_i + x'_{i,t} \gamma_0 + u_{i,t} \tag{3}$$

where $z_i$ is the employment density at location $i$ for 2002 coefficient $\beta_2$; $\hat{a}_i$ is the fitted value of employment connectivity treatment $a_{i,t}$ regressed by $z_{i,t}$ and all the other coefficients; $v_{i,t}$ is an independent Gaussian error term with mean 0. The remaining variables and specifications are similar to Equation 1.

We omit the report of regression coefficients as they are beyond our study goals while presenting the main diagnostic results. The Hausman-Wu test statistics (6,155 with df = 1;64,968) rejects the null hypothesis of exogeneity at the 95% confidence level while the null hypothesis of weak instrument is rejected by the Wald F-test statistics (2,552 with df = 1;64,967), which means that the choice of our instrument is warranted and a 2SLS approach is required. As the diagnostic results show, longitudinal econometric study with IV technique can be further conducted to understand the causal relationship behind the formation of our discovered urban structure. For example, Garcia-López *et al.* (2017) conducted a longitudinal study and provided a methodology for identifying the driving effects of transportation improvement with regards to sub-center formation in Paris from 1960 to 2010. Their approach uses a multi-year time series data while our approach uses a much finer spatial resolution (down to 30x30m parcel size), thus it serves as a complementary approach and supporting evidence for the findings in the Chicago study. With the advancement of remote sensing technology and emerging big data availability, a next step would be the analysis of longitudinal data at a fine spatial resolution to identify both micro-dynamics of urban structure as well as the causal relationships between land-use and connectivity to various attractive urban centers.

## 6. Conclusion

This paper examines the formation of employment sub-centers from a new perspective of heterogeneity in agglomeration economies. It captures the heterogeneous externalities in industrial sectors by using highly granular land-use data for the years 2001 and 2011, and it describes the changes in commercial and residential location choices patterns in relation to population, jobs, quality-of-life amenities, and transportation networks.

In terms of the agglomeration effects, it seems that the CBD as it was traditionally defined is no longer the primary source of agglomeration externalities for the new economic sectors. Comparatively, sub-centers with sector-specific positive agglomeration externalities appear to have stronger correlations with new commercial establishments. To our knowledge, this finding is the first to connect the heterogeneous production theories to land use planning and urban design with a finer geographical scale (30x30m patch level) and cross-temporal empirical evidence, providing new empirical insights into how urban sub-center grow. The more granular data on the one hand confirms some relationships we already know, e.g. proximity to major roads and highways is a time-invariant attractor for commercial land-use growth, and on the other hand has highlighted the heterogenous nature of agglomeration externalities, e.g. quality-of-life amenities in local neighborhoods have a much stronger influence on residential location choices than seen in more spatially aggregate studies.

We note that this study is only focused on the temporal heterogeneities in the business sectors. Although the conclusions highlight the importance of accounting for significant shifts over time regarding the effects, it is necessary in the next steps to investigate from where the sources of heterogeneities arise, e.g., from structural changes in business sectors, land use, built form, socioeconomic profiles, residents' preferences, or perhaps a combination of the factors. To pursue such further studies will demand significantly more data, but this may become viable as new data sources are likely to emerge through further availability of land use, sales transaction and other digital traces of urban activities. Another important note is that we are aware of possible endogeneities between commercial and residential developmental processes, although the data available so far is not sufficient for addressing such issues. Longitudinal econometric studies with IV or dynamic data panel techniques may need to be conducted to investigate causal links among the effects that we thus far uncover.

# References

Ahlfeldt, G.M., and McMillen, D.P. (2014). Land Values in Chicago, 1913–2010. Land Lines *26*, 15-21.

Ahlfeldt, G.M., Redding, S.J., Sturm, D.M., and Wolf, N. (2015). The Economics of Density: Evidence From the Berlin Wall. Econometrica *83*, 2127-2189.

Ahlfeldt, G.M., and Wendland, N. (2013). How polycentric is a monocentric city? Centers, spillovers and hysteresis. Journal of Economic Geography *13*, 53-83.

Alonso, W. (1964). Location and Land Use: Toward a General Theory of Land Rent (Cambridge: Harvard University Press).

Anas, A., Arnott, R., and Small, K.A. (1998). Urban Spatial Structure. Journal of Economic Literature *36*, 1426-1464.

Anderson, J.E., and van Wincoop, E. (2003). Gravity with Gravitas: A Solution to the Border Puzzle. American Economic Review *93*, 170-192.

Angel, S., and Blei, A.M. (2016). The spatial structure of American cities: The great majority of workplaces are no longer in CBDs, employment sub-centers, or live-work communities. Cities *51*, 21-35.

Batty, M. (2008). The Size, Scale, and Shape of Cities. Science *319*, 769-771.

Berry, B.J.L., and Kim, H.-M. (1993). Challenges to the Monocentric Model. Geographical Analysis *25*, 1-4.

Boarnet, M.G., Hong, A., and Santiago-Bartolomei, R. (2017). Urban spatial structure, employment subcenters, and freight travel. J Transp Geogr *60*, 267-276.

Chen, Y., and Rosenthal, S.S. (2008). Local amenities and life-cycle migration: Do people move for jobs or fun? Journal of Urban Economics *64*, 519-537.

Chen, Z., and Haynes, K.E. (2015). Impact of high speed rail on housing values: an observation from the Beijing–Shanghai line. Journal of Transport Geography *43*, 91-100.

Choi, K. (2018). The influence of the built environment on household vehicle travel by the urban typology in Calgary, Canada. Cities *75*, 101-110.

Chung, S., and Hewings, G.J.D. (2015). Competitive and Complementary Relationship between Regional Economies: A Study of the Great Lake States. Spatial Economic Analysis *10*, 205-229.

Czepkiewicz, M., Ottelin, J., Ala-Mantila, S., Heinonen, J., Hasanzadeh, K., and Kyttä, M. (2018). Urban structural and socioeconomic effects on local, national and international travel patterns and greenhouse gas emissions of young adults. J Transp Geogr *68*, 130-141.

Dong, X., and Ross, S.L. (2015). Accuracy and efficiency in simulating equilibrium land-use patterns for self-organizing cities. Journal of Economic Geography *15*, 707-722.

Duranton, G., and Puga, D. (2004). Micro-foundations of urban agglomeration economies. In Handbook of regional and urban economics (Elsevier), pp. 2063-2117.

Duranton, G., and Turner, M.A. (2012). Urban Growth and Transportation. The Review of Economic Studies *79*, 1407-1440.

Firgo, M., and Mayerhofer, P. (2017). (Un)related variety and employment growth at the sub-regional level. Papers in Regional Science.

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

Fujita, M., and Thisse, J.-F. (2009). New Economic Geography: An appraisal on the occasion of Paul Krugman's 2008 Nobel Prize in Economic Sciences. Regional Science and Urban Economics *39*, 109-119.

Fuzi, A. (2015). Co-working spaces for promoting entrepreneurship in sparse regions: the case of South Wales. Regional Studies, Regional Science *2*, 462-469.

Garcia-López, M.-À., and Muñiz, I. (2010). Employment Decentralisation: Polycentricity or Scatteration? The Case of Barcelona. Urban Studies *47*, 3035-3056.

Garcia-López, M.À., Hémet, C., and Viladecans-Marsal, E. (2017). Next train to the polycentric city: The effect of railroads on subcenter formation. Regional Science and Urban Economics *67*, 50-63.

Giuliano, G., and Small, K. (1991). Subcenters in the Los Angeles region. Regional Science and Urban Economics *21*, 163-182.

Glaeser, E. (2011). Cities, productivity, and quality of life. Science *333*, 592-594.

Glaeser, E.L., Kallal, H.D., Scheinkman, J.A., and Shleifer, A. (1992). Growth in Cities. Journal of Political Economy *100*, 1126-1152.

Gordon, P., Richardson, H.W., and Wong, H.L. (1986). The distribution of population and employment in a polycentric city: the case of Los Angeles. Environment and Planning A *18*, 161-173.

Henderson, V., Kuncoro, A., and Turner, M. (1995). Industrial Development in Cities. Journal of Political Economy *103*, 1067-1090.

Hu, L., Sun, T., and Wang, L. (2018). Evolving urban spatial structure and commuting patterns: A case study of Beijing, China. Transportation Research Part D: Transport and Environment *59*, 11-22.

Huang, D., Liu, Z., Zhao, X., and Zhao, P. (2017). Emerging polycentric megacity in China: An examination of employment subcenters and their influence on population distribution in Beijing. Cities *69*, 36-45.

Jacobs, J. (2016). The economy of cities (Vintage).

Kohavi, R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. In Proceedings of the 14th international joint conference on Artificial intelligence - Volume 2 (Montreal, Quebec, Canada: Morgan Kaufmann Publishers Inc.), pp. 1137-1143.

Kuang, C. (2017). Does quality matter in local consumption amenities? An empirical investigation with Yelp. Journal of Urban Economics *100*, 1-18.

Lucas, R.E., and Rossi-Hansberg, E. (2002). On the Internal Structure of Cities. Econometrica *70*, 1445-1476.

Mcdonald, J.F., and Mcmillen, D.P. (1990). Employment Subcenters and Land Values in a Polycentric Urban Area: The Case of Chicago. Environment and Planning A *22*, 1561-1574.

McDonald, J.F., and McMillen, D.P. (2000). Employment Subcenters and Subsequent Real Estate Development in Suburban Chicago. Journal of Urban Economics *48*, 135-157.

McMillen, D.P. (2001). Nonparametric Employment Subcenter Identification. Journal of Urban Economics *50*, 448-473.

McMillen, D.P. (2003). The return of centralization to Chicago: using repeat sales to identify changes in house price distance gradients. Regional Science and Urban Economics *33*, 287-304.

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

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

Melo, P.C., Graham, D.J., and Noland, R.B. (2009). A meta-analysis of estimates of urban agglomeration economies. Regional Science and Urban Economics *39*, 332-342.

Mills, E.S. (1967). An aggregative model of resource allocation in a metropolitan area. The American Economic Review *57*, 197-210.

Muth, R. (1969). Cities and Housing: The Spatial Pattern of Urban Residential Land Use (Chicago: University of Chicago).

Ottaviano, G., and Thisse, J.-F. (2004). Agglomeration and economic geography. In Cities and Geography, pp. 2563-2608.

Pan, H., Deal, B., Chen, Y., and Hewings, G. (2018). A Reassessment of urban structure and land-use patterns: distance to CBD or network-based? — Evidence from Chicago. Regional Science and Urban Economics *70*, 215-228.

Richardson, H.W. (1988). Monocentric vs. policentric models: The future of urban economics in regional science. The Annals of Regional Science *22*, 1-12.

Rosenthal, S.S., and Strange, W.C. (2003). Geography, industrial organization, and agglomeration. Review of Economics and Statistics *85*, 377-393.

Rosenthal, S.S., and Strange, W.C. (2004). Evidence on the Nature and Sources of Agglomeration Economies. In Handbook of Regional and Urban Economics, J.V. Henderson, and J.-F. Thisse, eds. (Elsevier), pp. 2119-2171.

Shukla, V., and Waddell, P. (1991). Firm location and land use in discrete urban space: A study of the spatial structure of Dallas-Fort worth. Regional Science and Urban Economics *21*, 225-253.

Tsekeris, T., and Geroliminis, N. (2013). City size, network structure and traffic congestion. Journal of Urban Economics *76*, 1-14.

Waddell, P., and Shukld, V. (1993). MANUFACTURING LOCATION IN A POLYCENTRIC URBAN AREA: A STUDY IN THE COMPOSITION AND ATTRACTIVENESS OF EMPLOYMENT SUBCENTERS. Urban Geography *14*, 277-296.

Wang, B., and Loo, B.P.Y. (2017). Hubs of Internet Entrepreneurs: The Emergence of Co-working Offices in Shanghai, China. Journal of Urban Technology *24*, 67-84.

Wang, Z., Chen, J., Zheng, W., and Deng, X. (2018). Dynamics of land use efficiency with ecological intercorrelation in regional development. Landscape and Urban Planning *177*, 303-316.

Wixe, S., and Andersson, M. (2016). Which types of relatedness matter in regional growth? Industry, occupation and education. Regional Studies *51*, 523-536.

Yang, Z., Song, T., and Chahine, T. (2016). Spatial representations and policy implications of industrial co-agglomerations, a case study of Beijing. Habitat International *55*, 32-45.

Zambon, I., Serra, P., Grigoriadis, E., Carlucci, M., and Salvati, L. (2017). Emerging urban centrality: An entropy-based indicator of polycentric development and economic growth. Land Use Policy *68*, 365-371.

Zhang, T., Sun, B., and Li, W. (2017). The economic performance of urban structure: From the perspective of Polycentricity and Monocentricity. Cities *68*, 18-24.

Zhang, X., Guo, Q.-e., Cheung, D.M.-w., and Zhang, T. (2018). Evaluating the institutional performance of the Pearl River Delta integration policy through intercity cooperation network analysis. Cities.

Zhao, P., Lu, B., and de Roo, G. (2011). The impact of urban growth on commuting patterns in a restructuring city: Evidence from Beijing. Papers in Regional Science *90*, 735-754.

Zhao, P., Lü, B., and Woltjer, J. (2009). Conflicts in urban fringe in the transformation era: An examination of performance of the metropolitan growth management in Beijing. Habitat International *33*, 347-356.

Zheng, S., Sun, W., Wu, J., and Kahn, M.E. (2017). The birth of edge cities in China: Measuring the effects of industrial parks policy. Journal of Urban Economics *100*, 80-103.

Zhong, C., Arisona, S.M., Huang, X., Batty, M., and Schmitt, G. (2014). Detecting the dynamics of urban structure through spatial network analysis. International Journal of Geographical Information Science *28*, 2178-2199.

# Supplementary Materials

Supplementary Materials S1

**Figure A1** is a summary of SGA process for finding shortest path from one population center *k* to other cells. **Figure A2** is detailed pseudo code for SGA process.

```
Algorithm SGA{
    Initiate every cell on the map with infinity travel time;

    Repeat the following N times{
        Initiate a direction;

        Repeat the following T steps{
            Agent from current cell moves to an adjacent cell, with higher
            probability to the cell with lesser travel barrier and in accordance
            with the original direction #agent from population center k in the
            first step
        }
    }
    Update the least travel time from population center k for every cell on
the map
}
```

**Figure A1. Brief Pseudo-code of SGA**

*Algorithm SGA{*

*Initiate every cell on the map with infinity travel time;*

*Repeat the following 1,000 times{*

*Randomly draw a direction* ***d*** *from direction set* ***M*** *= {N, NE, E, SE, S, SW, W, NW} with equal chance; # N—North, E—East, S—South, W---West, NE—Northeast, etc;*

*Repeat the following 1,000 steps{*

*Define* ***P*** *as a probability vector with {*$\boldsymbol{p_N}$*,* $\boldsymbol{p_{NE}}$*,* $\boldsymbol{p_E}$*,* $\boldsymbol{p_{SE}}$*,* $\boldsymbol{p_S}$*,* $\boldsymbol{p_{SW}}$*,* $\boldsymbol{p_W}$*,* $\boldsymbol{p_{NW}}$*}, where every element of* ***P*** *is a continuous number in [0, 1];*

*Create set of neighboring direction vector of d as* ***E****={*$\boldsymbol{e_1}$*,* $\boldsymbol{e_2}$*,* $\boldsymbol{e_3}$*}; # for example, if* ***d*** *is NE, then* ***E****={N, E, NE}*

*Randomly draw 2 directions from* ***E*** *as vector* ***D****={*$\boldsymbol{d_1}$*,* $\boldsymbol{d_2}$*} with equal probability; #in this case, assume NE is the direction* ***d*** *and D={****NE****,* ***N****}*

*Assign the probability P for cell to move in direction of each of the elements of D as 0.35; for the 2 directions adjacent to E but not in E as 0.1; for the remaining 4 directions as 0.025; In our example,* ***P****={*$\boldsymbol{p_N}$*,* $\boldsymbol{p_{NE}}$*,* $\boldsymbol{p_E}$*,* $\boldsymbol{p_{SE}}$*,* $\boldsymbol{p_S}$*,* $\boldsymbol{p_{SW}}$*,* $\boldsymbol{p_W}$*,* $\boldsymbol{p_{NW}}$*}={0.35 ,0.35 , 0.025, 0.09, 0.025, 0.025, 0.025, 0.09};*

*Assign a* ***L*** *probability vector to each direction based on the travel barrier of the land-use type on each cell;*

*Calculate the final direction moving probability* ***Q****={*$\boldsymbol{q_i}$*:*$\mathrm{i}\in\mathrm{N}, q_i = p_i l_i / \sum_{j\in N} p_j l_j$*};*

*Agent from current cell moves to an adjacent cell with probability vector* ***Q*** *to {N, NE, E, SE, S, SW, W, NW};*

*}*

*}*

*Update the least travel time from population center k for every cell on the map*

*}*

**Figure A2. Detailed Pseudo-code of SGA**