Deciphering the spatial stickiness, co-occurrence, and co-change of industries at the scale of city

Lock Yue Chew,[1,2,*] Ning Ning Chung,[3] Huynh Hoai Nguyen,[4] Wen Xuan Sia,[1] and Alvin Chua[5]

[1]*School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore.*
[2]*Data Science & Artificial Intelligence Research Centre, Nanyang Technological University, Singapore.*
[3] *Centre for University Core, Singapore University of Social Sciences, Singapore.*
[4]*Institute of High Performance Computing, Agency for Science, Technology and Research (A*STAR), 1 Fusionopolis Way, Connexis North Tower #16-16, Singapore 138632, Republic of Singapore.*
[5]*Design & Planning Lab, Urban Redevelopment Authority, Singapore.*

(Dated: April 29, 2023)

This paper analyses the location of industries in two metropolitan cities to investigate how firms are spatially distributed, and identify the trends in firm location choice. By defining a "Region Index" metric, we show how industries have concentrated in or diffused from regions-of-interest over time. We apply this metric to the Central Area (CA) of Singapore and the London region to decipher the relative tendency of industries to locate at CA as well as London. To analyse spatial co-location of industries, we constructed a community network, which reveal subtle spatial relationships between firms. Extending this approach temporally to create a co-change network, we deduce the occurrences of industries in Singapore and London that move into or out of a spatial region synchronously. Based on this approach, we elucidate firm-types that not only co-change but also co-occur in space and stick to a region-of-interest. Overall, we observe certain common and contrasting patterns of industry movements in the context of Singapore and London through the analysis. "Other activities auxiliary to financial serivices" is found to have the consistent tendency of diffusing out of the CA and London throughout the period of study. Conversely, opposite movement patterns are observed where we see the proclivity of IT and digital service firms to locate in London but relocate out of the CA, while the media, financial services, and life insurance sub-sectors tend to diffuse out of London but concentrate in the CA of Singapore.

I. INTRODUCTION

The location preference of firms are affected by a number of factors. This includes access to skilled labor and specialized supplier, the cost of real estate, the level of regulation and tax, and the availability of connected businesses that give rise to economies of scale [1–7]. Because these factors are more optimal within certain geographical regions, there are tendencies for firms and industries to situate in these regions. A case in point is Silicon Valley which is a hub for technological firms. By being a nexus of innovation activities that include research and development efforts between firms and universities, Silicon Valley has become a site where tech companies spatially agglomerate to reap the gains of knowledge and information spillover. Another geographical region of choice is the urban centre [8–15]. Firms that set up businesses in this central area (CA) may enjoy grants and stimulus that incentivize their economic activities. The dynamism in this area together with the benefits of the aforementioned factors have served to enhance the economic productivity of firms located there [16–18].

While firms may concentrate in these economic zones, certain firms may diffuse out of it due to other economic reasons [19]. For example, a set of firms may be galvanized to move into another region-of-interest to gain the economic advantage that the government had instituted into these regions in their land-use planning. Another reason for firms to shift their place of operation results from a change in their business strategy [20]. In fact, firms are known to spatially co-occur to exploit the benefits accrued from the agglomeration economy, and when opportunities arise, they may change their location in sync with other firms that are commercially synergistic with them. Such spatial relocation is expected to enhance the efficiency of their business operations.

With the availability of data on the temporal location of firms, we can probe the location choice of firms and their migratory patterns empirically [21]. In this respect, we shall focus our study on the localization behavior of firms within a city or a metropolitan area of interest. For this purpose, we have built tools and methods to identify the tendency of firms to stick to a certain geographical region, to co-locate with each other, or to co-move into another

*Electronic address: lockyue@ntu.edu.sg

region of space. This involves the development of three indices which we termed the co-occurrence index, the co-change index, and the region index, to draw out these spatial morphological features and behaviors of firms. In addition, we have employed concepts and methods from statistical physics and the field of complexity science such as the approach of complex networks to decipher the behavioral relationship between firm types when they co-locate or co-move with respect to each other. Our studies encompass a data size of the order of 5 hundred thousand firms in Singapore, and 7 million firms in the UK.

II. DATA AND METHODOLOGY

A. Data

Our study uses the Standard Industrial Classification (SIC) to organize and classify the firms based on their business activities.

Data for Singapore businesses is provided by the Accounting and Corporate Regulatory Authority (ACRA) [26]. ACRA collects business information (e.g. business name, registered office address, entity type, etc) when firms register and/or make statutory filings with the authority. We analyze the number of firms in 55 planning areas in 2009, 2013 and 2019 with the use of the Singapore Standard Industrial Classification (SSIC) to classify their business activities.

On the other hand, the data for UK firms is provided by the Office for National Statistics (ONS) [27]. In the dataset, UK businesses are broken down by legal status, industry, region, employment, and turnover size bands. We analyze the number of firms in 12 regions (i.e. North East, North West, Yorkshire and The Humber, East Midlands, West Midlands, East, London, South East, South West, Wales, Scotland, Northern Ireland) of UK according to their SIC from 2015 to 2021. In this dataset, we do not have the exact location of each firm. Hence, analysis on co-occurrence and co-change for this dataset is more coarse-grained compared to the analysis on the Singapore dataset.

B. Co-occurrence Index

The co-occurrence index measures the tendency of co-location between firms of two different types. It is based on pairwise correlation between spatially distributed data points [22]. To compute the index, we first partition the data points by R two-dimensional spatial regions. We then obtain the probability of data points of type i within region r as P_{ir}. Consolidating this set of probabilities over all r yield a one-dimensional vector P_i with its r-th element given by P_{ir}. The co-occurrence index between firms of type i and j is then quantified by the correlation between P_i and P_j as follows:

$$Co_{ij} = \frac{\sum_{r=1}^{R}(P_{ir} - 1/R)(P_{jr} - 1/R)}{\sqrt{\sum_{r=1}^{R}(P_{ir} - 1/R)^2}\sqrt{\sum_{r=1}^{R}(P_{jr} - 1/R)^2}}, \tag{1}$$

where $-1 \leq Co_{ij} \leq 1$.

When $P_i = P_j$, firms of type i co-locate perfectly with firms of type j and we have $Co_{ij} = 1$. In this case, the probability of finding an i-type and j-type firm in each region r is exactly the same. Conversely, if the presence of an i-type firm implies the absence of a j-type firm and vice versa, with the difference being uniform across all regions r, we have $Co_{ij} = -1$. Thus, a positive Co_{ij} implies that if an i-type firm is present in a certain region, there is a positive chance that a j-type firm is also present in the same region. On the other hand, if Co_{ij} is negative, the presence of an i-type firm is likely to indicate the absence of a j-type firm in the same spatial region. When the presence of an i-type firm is neutral to the presence of a j-type firm in the specified region, we have $Co_{ij} = 0$.

C. Co-change Index

The manner of calculating co-change is similar to co-occurrence except that the evaluation is done in time instead of space. Again, we consider R spatial regions, with Q_{ir} being the change in the number of firms of type i in region r within a specified interval of time. Because the change in number can be an increase or a decrease, the value of Q_{ir} is either positive or negative correspondingly. If the number of firms remain the same, $Q_{ir} = 0$. By putting the set of Q_{ir} over all the different r together into a vector, we form Q_i whose r-th element is Q_{ir}. The co-change index between firms of type i and j is then quantified by the Pearson correlation between the two trade categories of interest

as follows:

$$Cc_{ij} = \frac{\sum_{r=1}^{R}(Q_{ir} - 1/R)(Q_{jr} - 1/R)}{\sqrt{\sum_{r=1}^{R}(Q_{ir} - 1/R)^2}\sqrt{\sum_{r=1}^{R}(Q_{jr} - 1/R)^2}} . \tag{2}$$

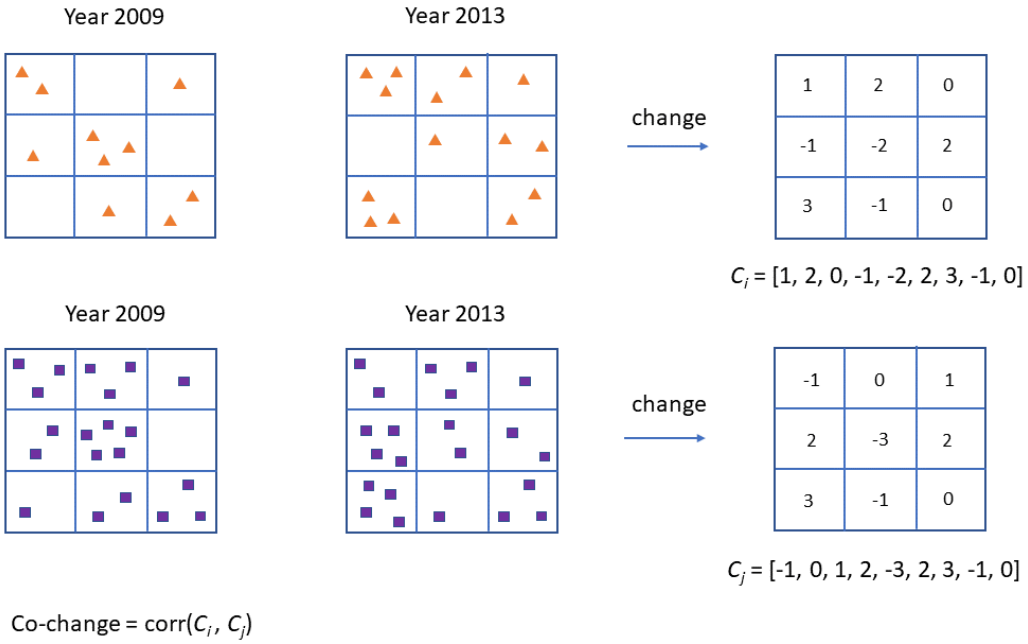

FIG. 1: Conceptual illustration of co-change. Here, the number shown in each cell in the rightmost grid gives the difference of the number of firms found in the corresponding cell in the grid from year 2009 to year 2013. C_i and C_j are the flatten vectors of the two-dimensional changes. $\mathrm{corr}(C_i, C_j)$ denotes the correlation between C_i and C_j.

The purpose of the co-change index is to quantify the statistical tendency of trade categories making similar changes in location choice. Its definition is analogous to that of the co-occurrence index, with the quantity in each spatial region now capturing the change from an initial year to a final year. This is illustrated in Fig. 1 where the initial year is 2009 and the final year is 2013. Note that the co-change index captures the correlated spatial changes of the two categories and its values span $-1 \leq Cc_{ij} \leq 1$. If $Cc_{ij} = 1$, the spatial changes of firm count of the two categories are exactly similar. If $Cc_{ij} = -1$, an increase in firm count in one trade is exactly match by a decrease of the same magnitude of firm count of the other trade. An index of $Cc_{ij} = 0$ then means the changes in firm count of the two trades are neutral of each other.

D. Region index

To evaluate how trades are sticky to a region of space, we define a region index to measure the stickiness of a trade to a certain region. Basically, the index measures the relative abundance of a trade in a specified region relative to that of all trades in the same region. In this paper, the region-of-interest is the central business district of Singapore, which is the central area of Singapore. Hence, we shall name this region index as the central-area index (i.e., CA-index) subsequently.

The idea of the region index originates from the D-index [23] which is defined as follows:

$$D = \sum_{r} p_{ir} \ln\left(\frac{p_{ir}}{p_{0r}}\right) . \tag{3}$$

Note that p_{ir} denotes the probability of a randomly sampled firm of type i locate in region r, while p_{0r} gives the probability of a randomly selected firm (of any type) locate in the same region r. Region index is defined by taking only the log-likelihood ratio part of the D-index, i.e.,

$$R - \text{index} = \ln\left(\frac{f_A}{f_0}\right) , \tag{4}$$

with f_A being the fraction of A-type firms locate in a specified region R, while f_0 is the fraction of all firms locate in the same region R.

R-index measures the affinity of an A-type firm situating in R relative to that of the baseline of all firms. For a R-index of zero, the affinity of A-type firms locating in R is the same as that of the baseline, and we say that A-type firms are neutral towards situating in R. However, when R-index is positive, it implies that A-type firms have a greater preference to locate in R relative to baseline, and A-type firms are attracted to R. On the other hand, if R-index is negative, we would observe a tendency of A-type firms to not locate in R relative to baseline, and it could imply that A-type firms are not attracted towards situating themselves in R.

III. RESULTS

A. Background

The aim of our study here is to decipher the preference or non-preference of industrial subsectors in concentrating at a particular spatial region, which we refer to respectively as being "sticky" or "nonsticky" to the space concerned. For our analysis, the central area (CA) of Singapore is taken as the spatial region of interest which leads us to define the CA-index. The timeframe of our analysis is from 2009 to 2019. In addition to exploring the stickiness to space of subsectors, we also investigate into the spatial co-location of sub-sectors through the co-occurrence network. After which, we deduce the occurrences of subsectors that change together in moving in or out of a spatial region, leading to a co-change network. We then investigate the possible presence of a set of subsectors that not only co-change but also co-occur in space in a region-of-interest.

B. CA-Index

We have computed the region index of all industrial sectors at the central area of Singapore, which we display in Fig. 2. Because the region-of-interest is at the central area (CA), we term the region index as the CA-index. Figure 2 shows that businesses in sections J, K, L and M (see Table I) of the CA of Singapore possesses a high CA-Index relative to other sections. This has led us to focus on these sections in our subsequent analysis since firms within these categories are *stickier* to CA. Note that Singapore's central area include 11 planning areas, i.e. Downtown Core, Marina East, Marina South, Museum, Newton, Orchard, Outram, River Valley, Rochor, Singapore river, and Straits View.

To understand the migratory behaviour of firms with respect to CA, we have evaluated the consistent increase or decrease of CA-index across the years. The results highlighted two distinct patterns: a concentration of firms into CA, and a diffusion of firms out of CA. Figure 3 shows a set of firms whose CA-index consistently increase from 2009-2019 in Singapore which illustrate their preference to concentrate in CA over the time frame of our study. Note that the nomenclature of the firms whose SSIC are indicated in Fig. 3 is given in Table II. In particular, we observe that firms in the "Banking, Finance, Life Insurance, and Real Estate" industry as well as those from the "Media (Motion Picture, Video, TV, Satcomm)" industry have a higher prevalence to consistently congregate in CA over the decade which suggest the attractiveness of CA to these industries. This is exemplified by firms from the Wholesale Banks which consistently site their operations in CA within a decade (from 2009 to 2019, see Figs. 4a and 4b). Another example is given by Carbon consultancy services, whose firms initially scatter about Singapore in the year 2009, but increasingly concentrate in the CA by 2019 (see Figs. 4c and 4d).

On the other hand, there are industrial sectors that have begun to site their operations away from the CA. Figure 5 displays such a set of industrial sectors whose CA-index consistently decrease from 2009 to 2019 in Singapore (see Table III on the nomenclature of the firms associated with the SSIC in Fig. 5). Interestingly, we observe that firms from the following group of industries show a consistent decline in their CA-index: "IT Development", "Media (Motion Picture, Video, TV) production", "Accounting and Auditing", "Auxiliary financial services", "Venture capital", and "Research and Development in Biotech, Life and Medical Science". While these industrial sectors continue to have a preference to locate in the CA of Singapore, their propensity to locate in CA is found to be waning. As the

Section	Sector
A	Agricultural and Fishing
B	Mining and Quarrying
C	Manufacturing
D	Electricity, Gas, Steam and Air-Conditioning Supply
E	Water Supply; Sewerage, Waste Management and Remediation Activity
F	Construction
G	Wholesale and Retail Trade
H	Transport and Storage
I	Accommodation and Food Service Activities
J	Information and Communications
K	Financial and Insurance Activities
L	Real Estate Activities
M	Professional, Scientific and Technical Activities
N	Administrative and Support Service Activities
O	Public Administration and Defence
P	Education
Q	Health and Social Services
R	Arts, Entertainment and Recreation
S	Other Service Activities
T	Activities of Households as Employers of Domestic Personnel
U	Activities of Extra-Territorial Organizations and Bodies

TABLE I: Description of the Section Categories referenced in this paper.

size of these diffusing industries expanded over time from 2009 to 2019, firms in these industries chose to locate in alternative locations outside of CA. This re-localization behavior is illustrated by firms of the "Other activities Auxiliary to Financial Service Activities NEC (eg Mortgage Advisory Firms)" and "Development of Software for Interactive Digital Media (except Games)" industrial sectors (see Fig. 6). For the latter, we observe a tendency of firms to concentrate in the Geylang and Bedok area as they migrate away from CA.

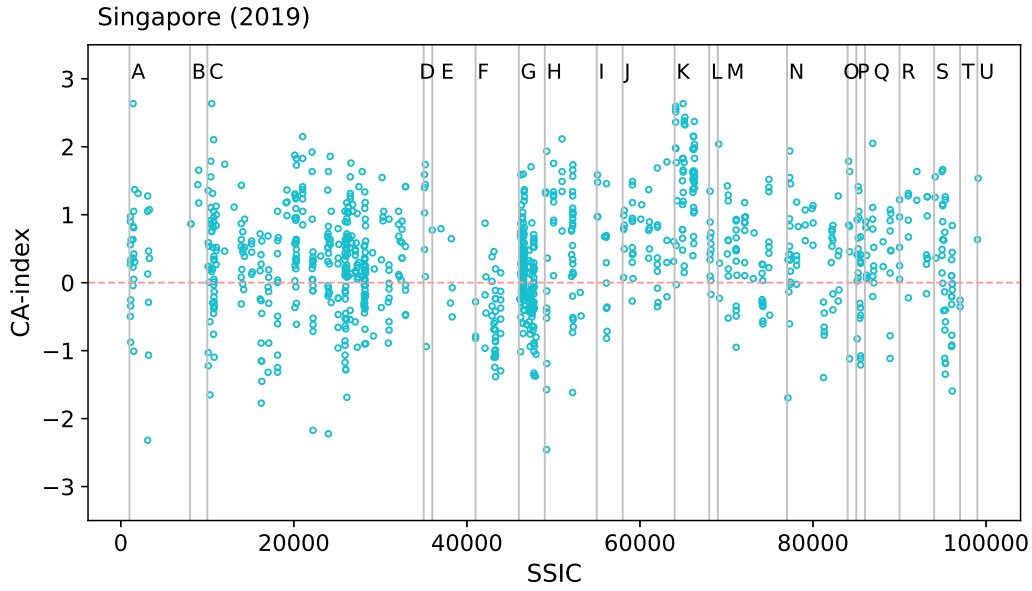

FIG. 2: Region Index of the Central Area of Singapore in 2019. Business activities are further grouped into a broad industry group structure consisting of 21 sections (i.e. A to U). Remarkably, most businesses in sections J, K, L and M are found to position themselves preferably in the country center with positive CA-index in Singapore.

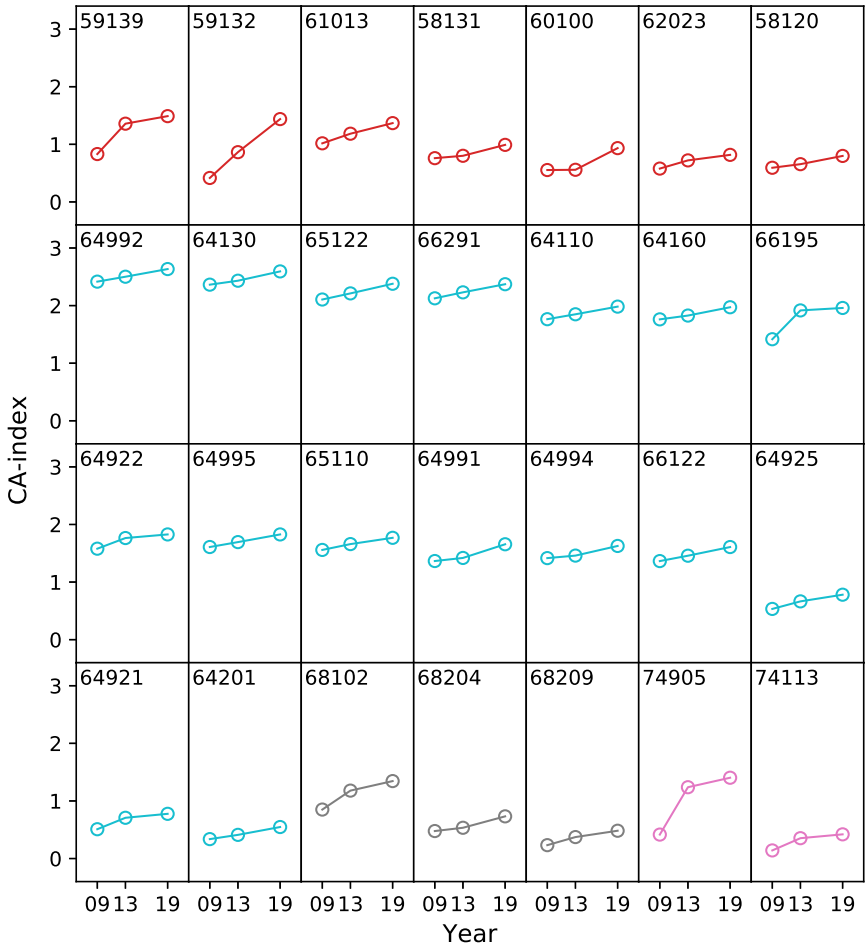

FIG. 3: CA-index for industries in Singapore in 2009, 2013 and 2019. The industries shown here are from sections J (red), K (cyan), L (grey) and M (pink) with CA-index increases consistently throughout the period under study. Only industries with significant increase in CA-index (≥ 0.2) are shown here.

SSIC	Industry	SSIC	Industry
59139	Motion picture, video and television programme distribution activities nec	64922	Credit card services
59132	Distribution of dramas, variety shows and documentaries (including television programme)	64995	Treasury centres
61013	Satellite telecommunications activities	65110	Life insurance
58131	Publishing of newspapers	64991	Factoring companies
60100	Radio programme production and broadcasting	64994	Remittance services
62023	Computer facilities management activities	66122	Foreign exchange brokers and dealers
58120	Publishing of directories and mailing lists	64925	Pawnbrokerage
64992	Representative offices of foreign banks	64921	Thrift and loan societies and credit associations
64130	Wholesale banks	64201	Bank/Financial holding companies
65122	Captive insurance	68102	Operating of serviced apartments
66291	Representative offices of foreign insurers	68204	Commercial and industrial real estate management
64110	Central bank	68209	Real estate activities on a fee or contract basis nec
64160	Finance companies (with deposit taking functions)	74905	Carbon consultancy services
66195	Transaction/Payment processing services	74113	Furniture design services

TABLE II: Description of SSIC shown in Fig. 3.

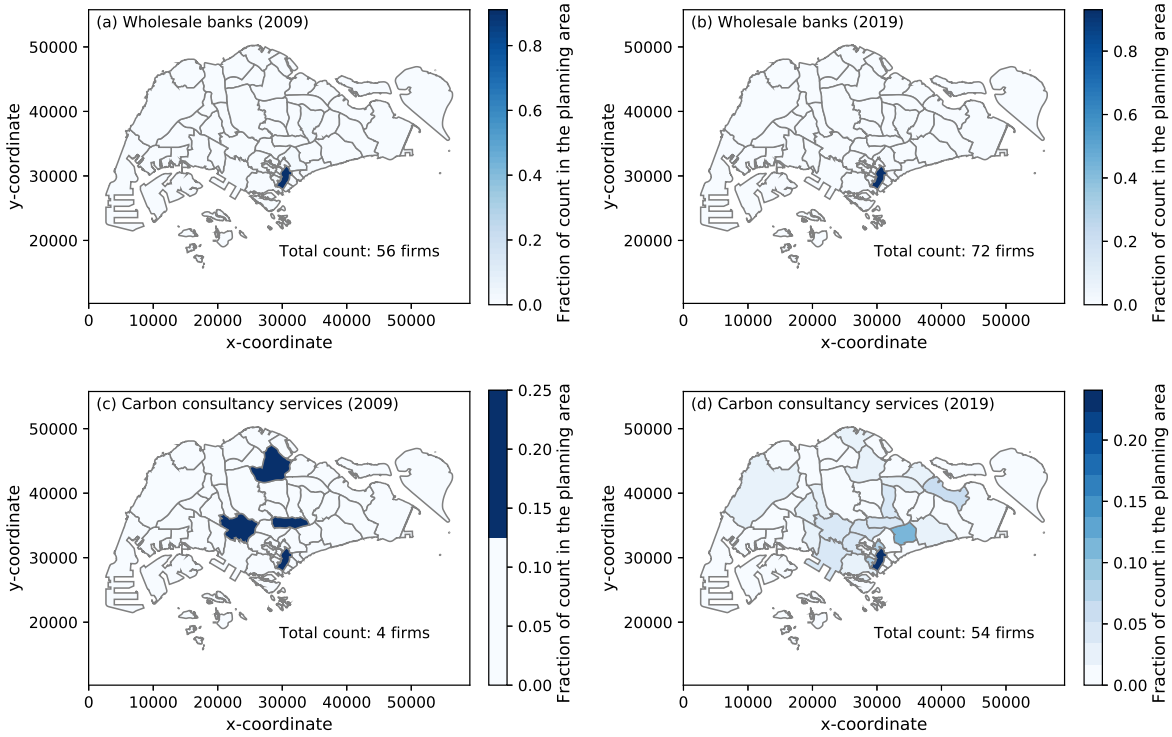

FIG. 4: Distribution of firms in the 55 planning areas of Singapore for wholesale banks ((a) and (b)) and carbon consultancy services ((c) and (d)). Wholesale banks exemplified a sector which possessed high preference to locate at the central area consistently throughout the period under study. On the other hand, carbon consultancy services was a sector which had a lower preference to locate at the central area in 2009 but became more sticky to the central area in 2019. Note that "Total count" in the figure indicates the number of firms in CA.

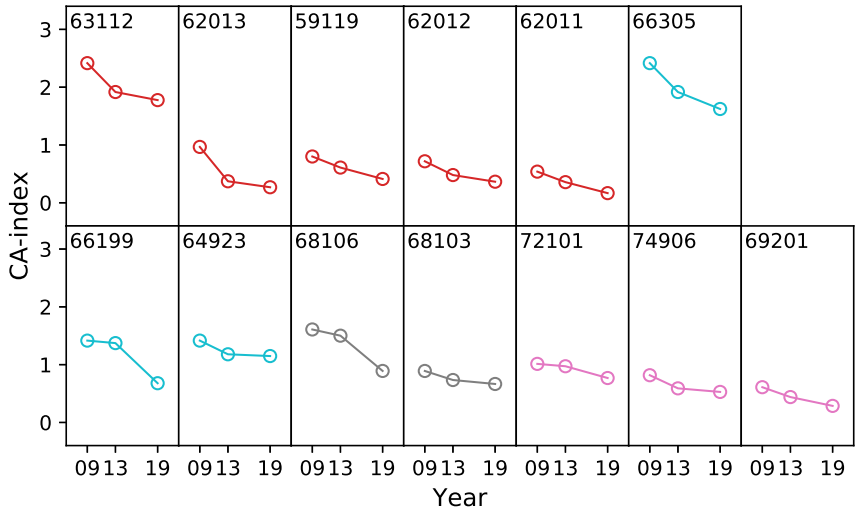

FIG. 5: CA-index for industries in Singapore in 2009, 2013 and 2019. The industries shown here are from sections J (red), K (cyan), L (grey) and M (pink) with CA-index decreases consistently throughout the period under study. Only industries with significant decrease in CA-index (≤ -0.2) are shown here.

SSIC	Industry	SSIC	Industry
63112	Data entry services	64923	Mutual benefits organisations
62013	Development of software for interactive digital media (except games)	68106	Management of self-owned strata titled property (i.e. Management Corporation Strata Title)
59119	Motion picture, video and television programme production activities nec	68103	Letting and operating of self-owned real estate except hotels, lodging and boarding houses
62012	Development of computer games	72101	Research and experimental development on biotechnology, life and medical science
62011	Development of e-commerce applications	74906	Agents for artistes, athletes, models and other performers
66305	Venture capital activities	69201	Accounting and auditing services (including taxation advisory services)
66199	Other activities auxiliary to financial service activities nec (eg mortgage advisory firms)		

TABLE III: Description of SSIC shown in Fig. 5.

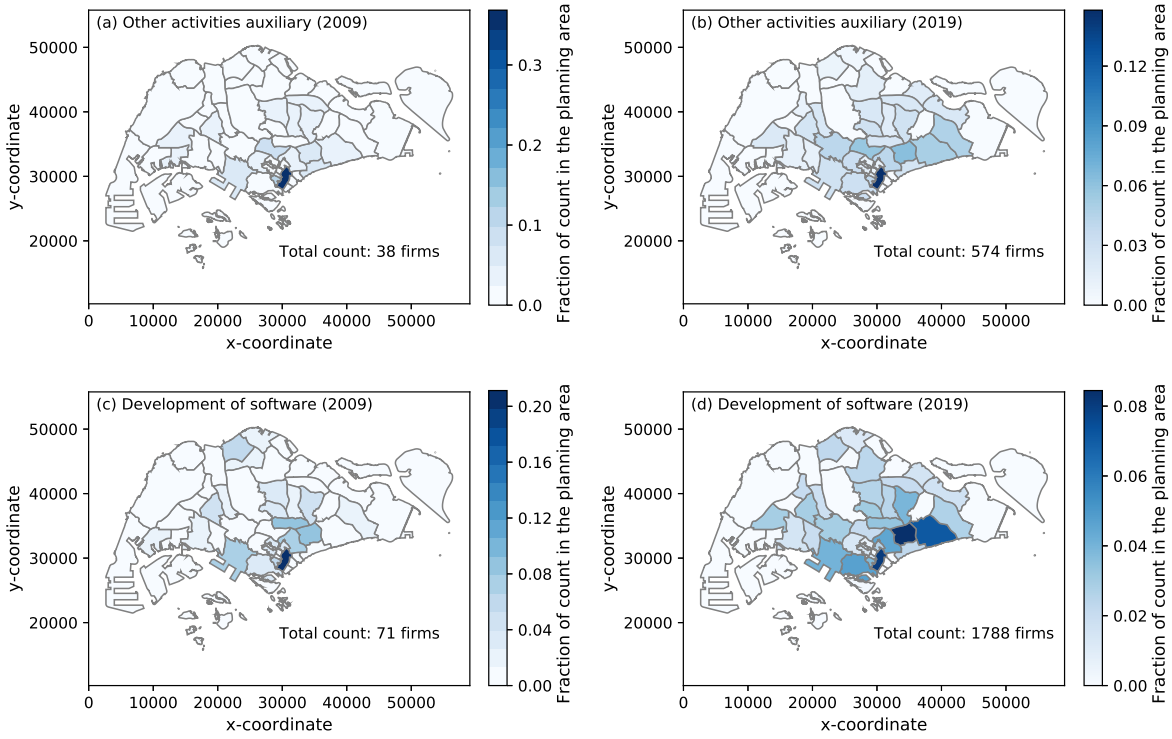

FIG. 6: Distribution of firms in the 55 planning areas of Singapore for other activities auxiliary to financial service activities nec (e.g. mortgage advisory firms) ((a) and (b)) and development of software for interactive digital media (except games) ((c) and (d)).

C. Spatial Co-occurrence

Certain trades tend to locate in proximity with each other and these trades can be uncovered through our measure of spatial co-occurrence. We compute this tendency of firms of two different sub-sectors co-locate together using our co-occurrence index. We perform such an evaluation over all pairs of industrial sub-sectors. By first representing each sub-sector as a node, we form a link between two nodes if the sub-sectors that correspond to them have a co-occurrence index greater or equal to a threshold of 0.85. (Note that in this paper the threshold is set at a level that gives about 50 network nodes such that visual analysis is tractable.) The outcome of this link formation process over the set of sub-sector nodes is a co-occurrence network which is illustrated in Fig. 7. This co-occurrence network displays the set

of trades that tends to co-occur together for the year 2009. Note that we have represented nodes in the same section in the same colour in the network with node size being proportional to the number of connections the node has.

We observe a separation of the co-occurrence network into two major sub-networks. A common feature that we uncovered in these sub-networks is the co-occurrence of sub-sectors in the same section. This is revealed through a community of sub-sectors from the same sections within the sub-network. For example, we noticed a small cluster of trades in the Transportation and Storage sector (section H) which co-occur together. Within this same sub-network, we found a large cluster of trades in the Financial and Insurance Activities sector (section K) that also co-locate with each other. In the other sub-network, we have detected spatial co-occurrence among a small cluster of sub-sectors in the Construction sector (section F). Interestingly, spatial co-occurrence can also happen across sections here, which is manifested among trades in the Wholesale and Retail Trade sector (section G), the Information and Communications sector (section J), and the Professional, Scientific and Technical Activities sector (section M).

We next construct the co-occurrence network for the year 2019 under the same threshold condition. Here, we see an expansion of the network with a larger number of nodes (see Fig. 8). There is still two major sub-networks separated from each other as observed in 2009 but is now larger. And we again observe spatial co-occurrence of sub-sectors in the same section as before. The main difference is the spatial co-occurrence across sections which now possesses a greater variety of sub-sectors, with addition from Wholesale and Retail Trade (section G), Information and Communications (section J), Finance and Insurance Activities (section K), Real Estate Activities (section L), Administrative and Support Service Activities (section N), and Other Service Activities (section S). Finally, the triage of sub-sectors: "Restaurants", "Cafes and Coffee Houses" and "Food Court, Coffee Shops", whose co-location is expected due to the nature of their business, continues to exist in the sub-network of 2019 as they have in 2009.

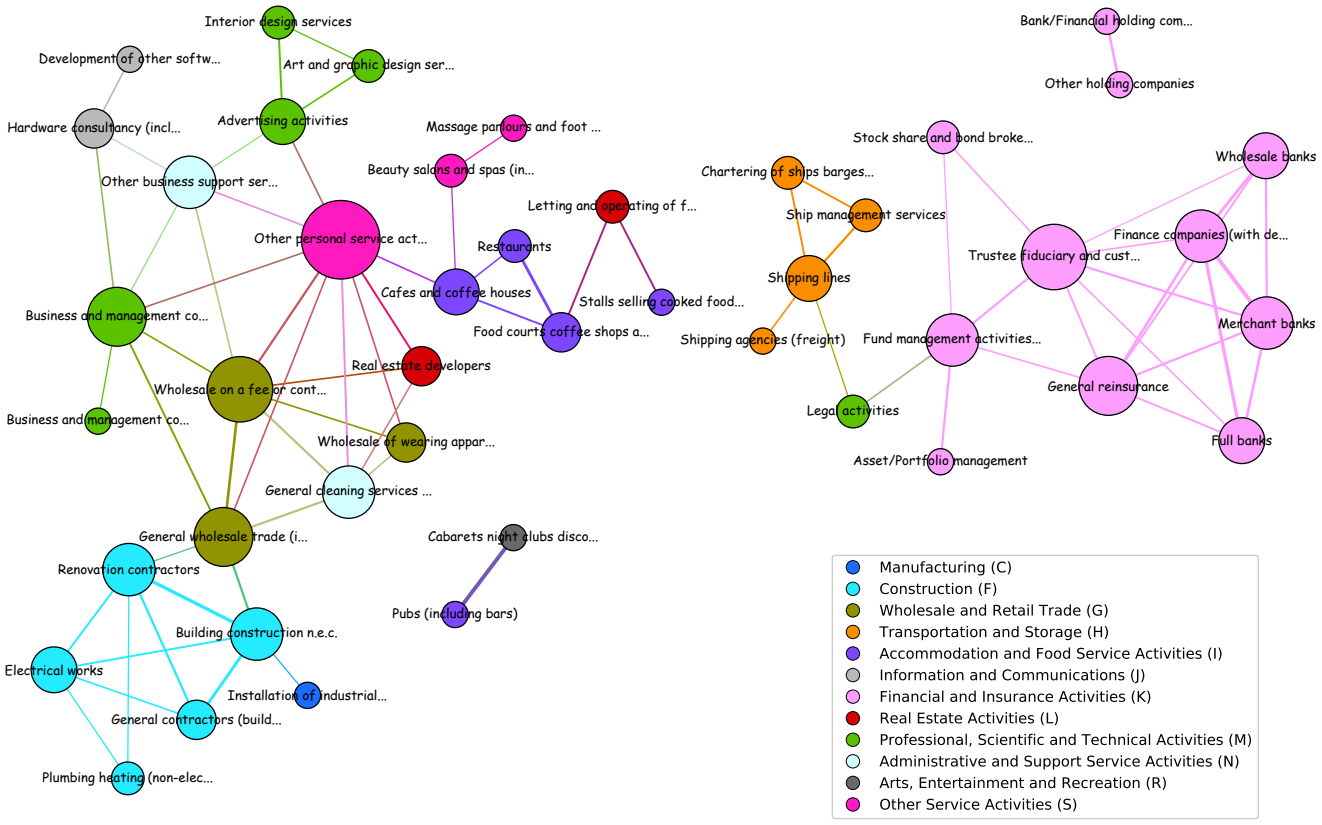

FIG. 7: Co-occurrence network of Singapore's industries in 2009. Only industries with more than 10 firms are included. Here, a threshold of $Co_{ij} \geq 0.85$ is used to determine whether two nodes are linked. Sub-sectors in the same section are shown as nodes of the same color. Note that the legend gives the name of the section with its category in bracket. Node size is proportional to node degree.

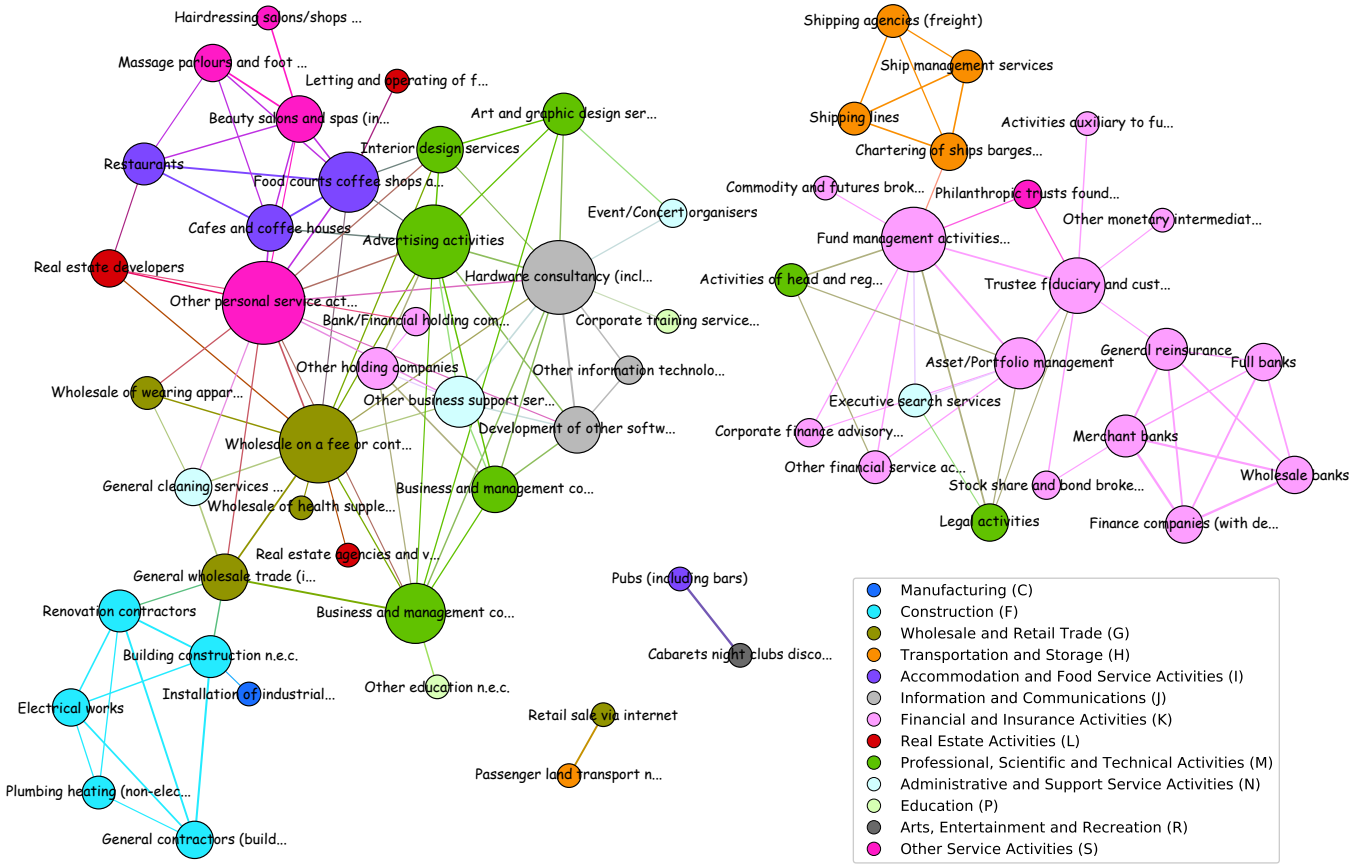

FIG. 8: Co-occurrence network of Singapore's industries in 2019. Only industries with more than 10 firms are included. Here, a threshold of $Co_{ij} \geq 0.85$ is used to determine whether two nodes are linked. Sub-sectors in the same section are shown as nodes of the same color. Note that the legend gives the name of the section with its category in bracket. Node size is proportional to node degree.

D. Co-Change

Next, we compute the tendency of sub-sectors that make correlated change in time between 2009 to 2019. The evaluation is performed between each pair of sub-sectors, and if the calculated co-change index between the pair is greater or equal to the threshold of 0.85, we place a link between the nodes of the two respective sub-sectors. The result of the process is a co-change network which is shown in Fig 9.

In the obtained co-change network, we observe the predominant co-change of two sectors: Finance and Insurance Activities sector (section K) and Information and Communications sector (section J). The expansion of the economy from 2009 to 2019 seems to have driven concerted changes in these two sectors. In this respect, it is interesting to note also the co-change of sub-sector "Retail sale via internet" with a set of other miscellaneous sub-sectors. We perceive that the synchronized changes of all these sub-sectors are indicative of a shift towards the digital economy.

E. Co-Change, Co-locate and Stick or Un-Stick to CA

Urban planners are interested to understand the spatial morphological changes of industrial firms in both time and space for their planning purposes. This includes how firms decentralize from the Central Business District (CBD), or how urban resurgence occurs as firms re-consolidate back into the CBD [24]. They are also interested to identify the locale where unique economic district, such as the Recreational Business District, could potentially emerge [25]. The tools developed in this paper provide techniques that not only help urban planners decipher these morphological

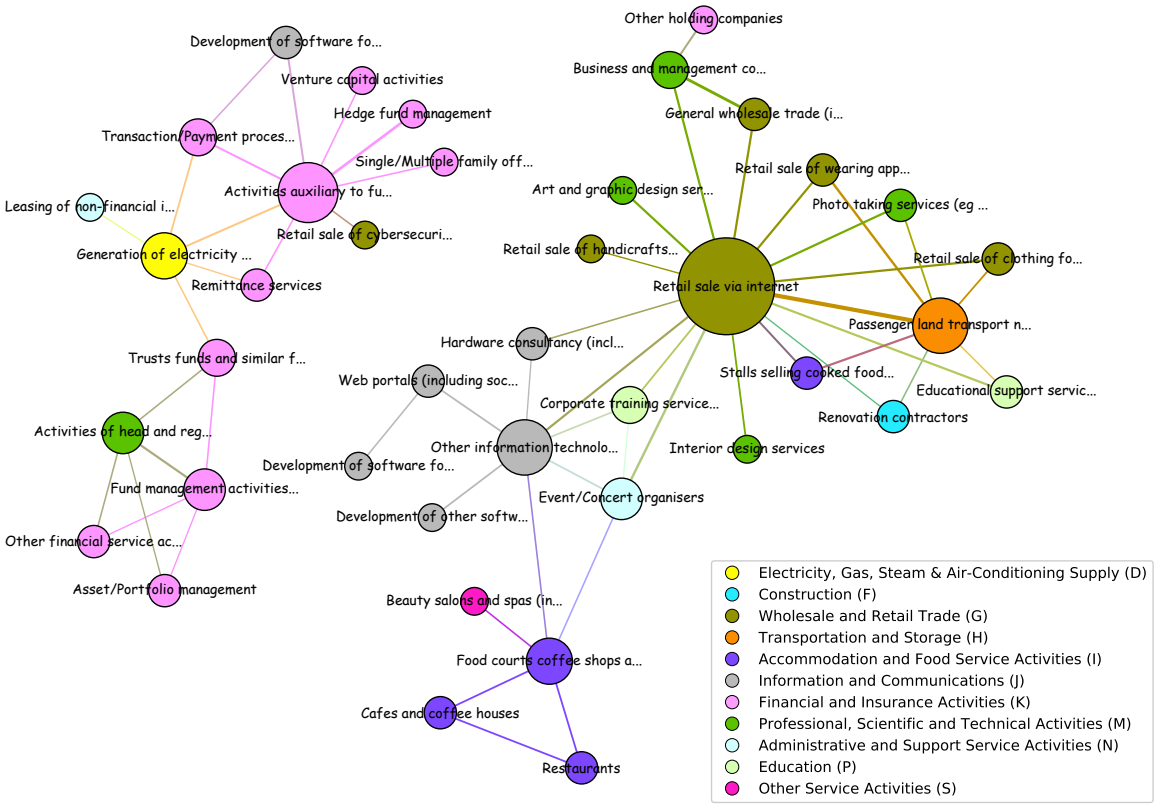

FIG. 9: Co-change network of Singapore's industries from 2009 to 2019. Only industries with more than 10 firms are included. Here, a threshold of $Cc_{ij} \geq 0.85$ is used to determine whether two nodes are linked. Sub-sectors in the same section are shown as nodes of the same color. Note that the legend gives the name of the section with its category in bracket. Node size is proportional to node degree.

changes, they can also elucidate the co-movement of industrial firms as well as their tendency to co-locate and to stick to a certain region of space. By combining the use of our developed indices, we are able to probe not only movement and spatial preference of individual firms, but also the spatial and temporal propensity of a *collection* of firms in their location choice. The insights gained from the observed collective behaviour of the firms could serve to validate the planning intent of the urban planners and to guide their decisions towards planning schemes that yield optimal economic benefits.

In this section, we shall demonstrate the combined use of the co-change, co-occurrence, and CA indices to uncover sub-sectors that co-change, co-occur, and stick or unstick, to CA. To achieve this, we first apply the co-change index to construct the co-change network before employing the co-occurrence index to filter out sub-sectors that do not spatially co-locate with each other. The outcome is a co-change cum co-occur network that reduces from the co-change network. CA index is then applied to the sub-sectors in this co-change cum co-occur network to elucidate sub-sectors that are sticky or unsticky to CA.

Our analysis found that among the industries, "Wholesale Banks" and "General Reinsurance" have similar location preference. The two sub-sectors predominantly located together in CA in 2009, and remained in CA between 2009 – 2019, with an increase of count that happened mainly in CA ($\approx 90\%$). In other words, these two sub-sectors are not only sticky to each other, they are also sticky to CA (see Fig. 10). For these two industries, efforts to decentralize them out of CA would be challenging. In contrast, firms in the sub-sectors "Other Activities Auxiliary to Financial Services" and "Other Financial Service Activities Except Insurance" would be more amenable to the decentralization effort. These two sub-sectors are sticky to each other but not as sticky to CA. The two sub-sectors were predominantly located together in CA in 2009, but shifted in a similar pattern between 2009–2019 (see Fig. 11). Although their preference to locate in the CA remains strong, we start to see an increase of firms in other areas like Bedok, Tampines, Kallang, and Geylang. Another pair of sub-sectors, "Development of e-commerce applications" and "Accounting and auditing services" are also found sticky with each other. A substantial number of firms in these two sub-sectors located together in CA in 2009. However, they are much less sticky to CA subsequently and were observed to co-locate in

alternative locations instead of the CA when their industries expanded between 2009–2019 (See Fig. 12).

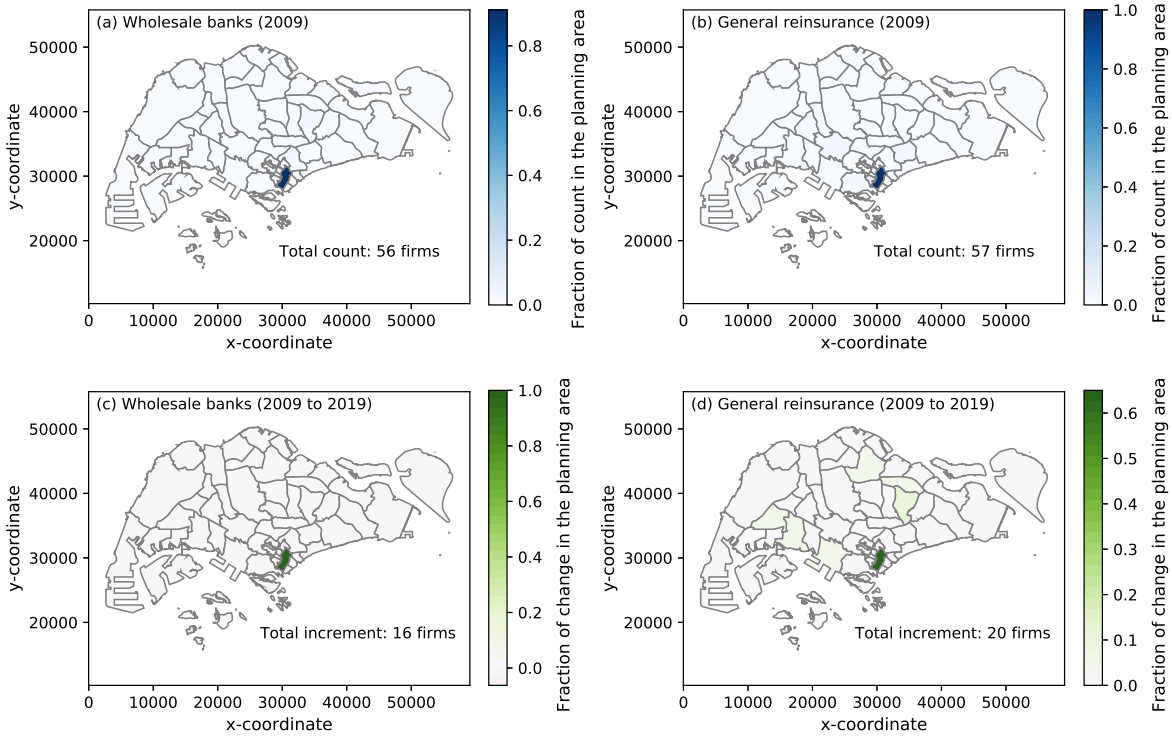

FIG. 10: Distribution of firms ((a) and (b)) and distribution of change of firm count ((c) and (d)) in the 55 planning areas of Singapore for "Wholesale Banks" and "General Reinsurance".

F. UK and the London Region

We have performed equivalent analysis on UK by applying our developed indices and also the approach of complex networks. First, we compute the region index of the London region to uncover which category of industry is sticky to the region concerned. Fig. 13 shows the region index of the London region where we again observe that businesses in section J, K, L and M prefer to locate in the region-of-interest just like the CA of Singapore. Interestingly, we also found common sticky industries to that of the CA of Singapore. These are the Information and Communication Industry, the Financial and Insurance Activities industry, the Real Estate Activities industry, and the Professional, Scientific and Technical Activities industry since trades within these industry tend to have positive region index. However, the set of sub-sectors that consistently concentrates in the London region or consistently diffuses out of it from 2015-2021 (see Fig. 14 and 15) seem to follow opposite trends compared to that of Singapore CA. For London, we observe a concentration of a set of IT and digital service firms into the area, but certain trades in this same category are found to diffuse away from Singapore CA. On the other hand, we notice a diffusion of firms in the media, financial services, and life insurance sub-sectors out of the London region, while firms in these same sub-sectors are observed to concentrate in the CA of Singapore. Nonetheless, we observe the interesting phenomenon that firms in the sub-sector "Other activities auxiliary to financial services" show a consistent decline in both London as well as the CA of Singapore.

Figs. 16 and 17 illustrate the co-occurrence networks of UK for the year 2015 and 2021 respectively, which are both of relatively the same size. In the two years, there occurs one large sub-network and a few small subnetworks. The large sub-network displays a diversity of trades across sections without forming co-occurrence of sub-sectors in the same section like that of Singapore. Nonetheless, we observe co-occurrence of the sub-sector "Construction" in the small network of UK, where we note that spatial co-occurrence of construction firms is also observed to happen in Singapore (see Figs. 7 and 8). Compared to Singapore, larger spatial co-occurrence across sections is observed. In addition, strong co-occurrence between sub-sectors in the Finance and Insurance Activities sector is absent. We believe this to result from our use of the region level of UK which is a coarse-grained analysis. If a smaller geographical scale

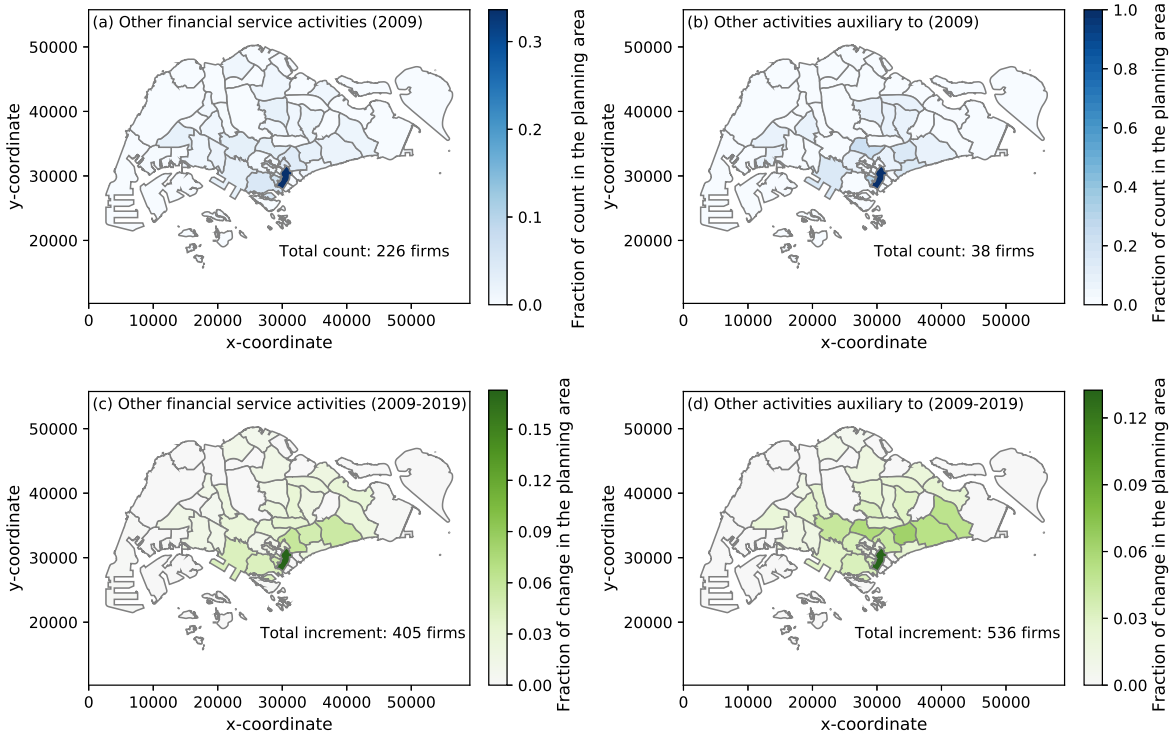

FIG. 11: Distribution of firms ((a) and (b)) and distribution of change of firm count ((c) and (d)) in the 55 planning areas of Singapore for "Other Financial Service Activities Except Insurance" and "Other Activities Auxiliary to Financial Services".

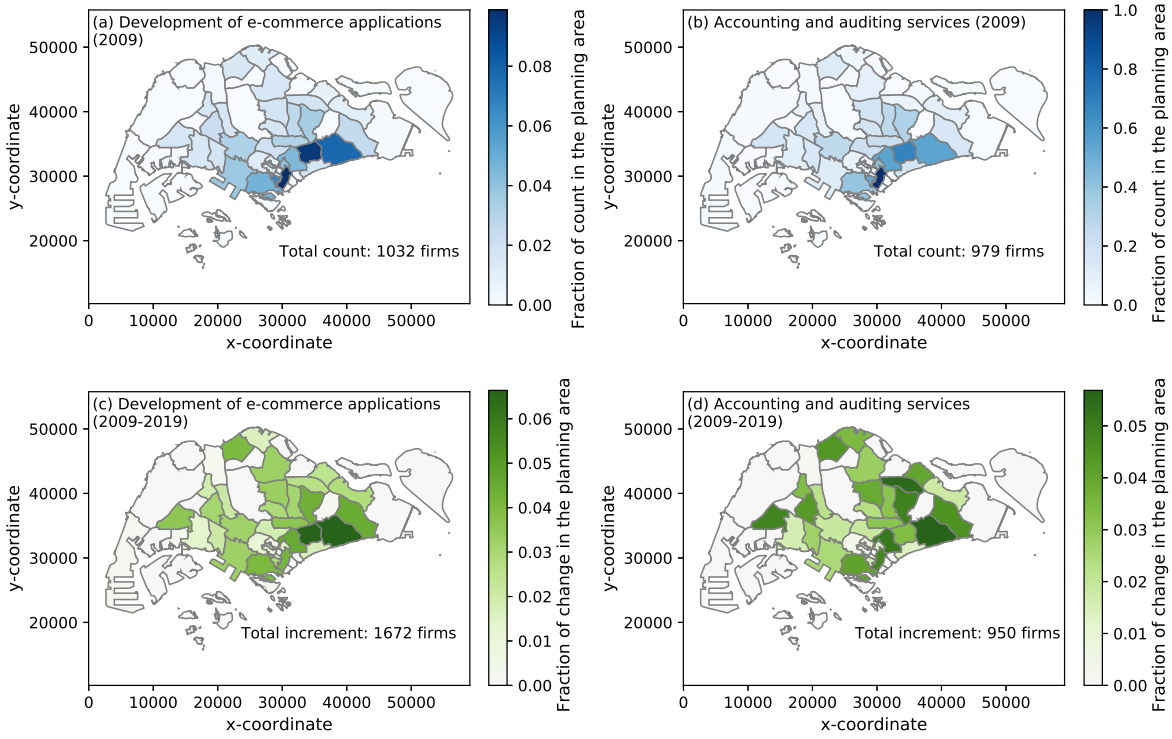

FIG. 12: Distribution of firms ((a) and (b)) and distribution of change of firm count ((c) and (d)) in the 55 planning areas of Singapore for "Development of e-commerce applications" and "Accounting and auditing services".

is applied, co-occurrence between sub-sectors in the Finance and Insurance Activities sector should become evident. In the same vein, we expect strong preference of firms to co-locate in the prime areas of a city instead of co-locating at the region or state level. It is interesting, however, to highlight the counter-example of the construction sector, whose co-occurrence behavior appears to be scale-invariant. We detect spatial co-occurrence among a small cluster of sub-sectors in this industry, which may suggests an operational need for these sub-sectors to co-locate.

Fig. 18 shows the co-change network of UK from 2015 − 2021. Synchronized changes happen mainly in the Construction, Real Estate Activities, and Professional, Scientific and Technical Activities sections. It is noteworthy that the construction sub-sector demonstrates a strong tendency to co-change and co-occur in UK, which is different from that of Singapore where we observe a tendency of this industry to co-occur but not co-change. Finally, we notice that the sub-sectors "Real estate agencies" and "Management of real estate on a fee or contract basis" has a propensity to co-occur and co-change. Although they are also sticky to the London region, we have detected their gradual movement out of this region-of-interest.

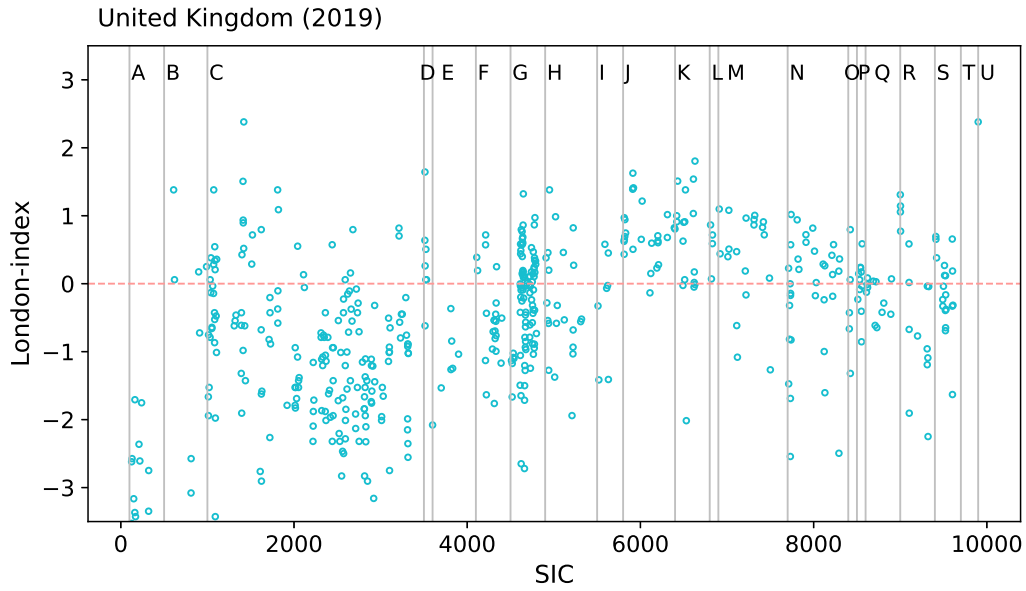

FIG. 13: Region index of the London region of United Kingdom in 2019. Business activities are further grouped into a broad industry group structure consisting of 21 sections (i.e. A to U). Interestingly, most businesses in sections J, K, L and M are found to position themselves preferably in the London region, similar to how businesses in these sections possess positive CA-index in Singapore.

SIC	Industry	SIC	Industry
5821	Publishing of computer games	6512	Non-life insurance
5829	Other software publishing	6612	Security and commodity contracts brokerage
6209	Other information technology and computer service activities	6622	Activities of insurance agents and brokers
6311	Data processing; hosting and related activities	7010	Activities of head offices
6312	Web portals		

TABLE IV: Description of SSIC shown in Fig. 14.

IV. DISCUSSION AND CONCLUSION

In this paper, we have developed three indices to decipher the spatial-temporal patterns and dynamics of firm location choice. The first index is the region index, which serves to ascertain the level of attractiveness of a region-of-interest to business activities over time. Our second index measures the proclivity of firms to locate in spatial

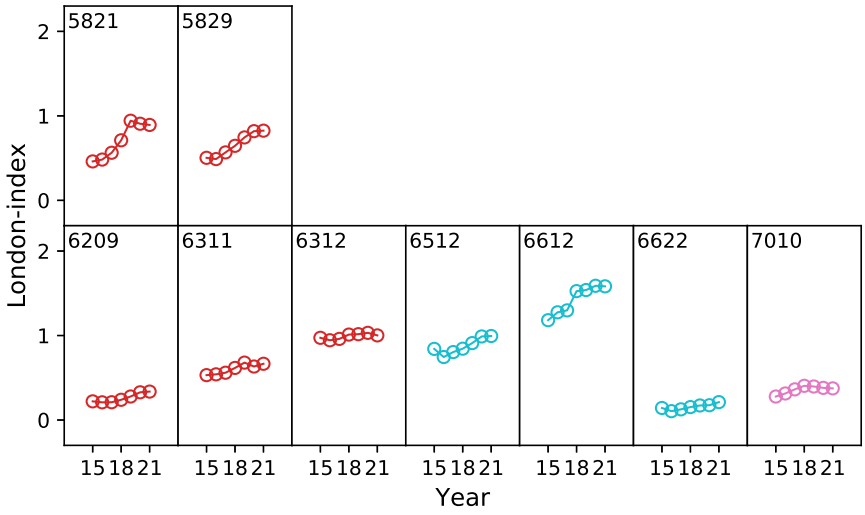

FIG. 14: London-index for industries in the United Kingdom from 2015 to 2021. The industries indicated here are from sections J (red), K (cyan), L (grey) and M (pink), with London-index increases consistently throughout the period under study (i.e., correlation greater than 0.75).

SIC	Industry	SIC	Industry
5911	Motion picture; video and television programme production activities	6619	Other activities auxiliary to financial services; except insurance and pension funding
5912	Motion picture; video and television programme post-production activities	6630	Fund management activities
5913	Motion picture; video and television programme distribution activities	6810	Buying and selling of own real estate
6010	Radio broadcasting	6831	Real estate agencies
6020	Television programming and broadcasting activities	6832	Management of real estate on a fee or contract basis
6190	Other telecommunications activities	6910	Legal activities
6419	Other monetary intermediation	6920	Accounting; bookkeeping and auditing activities; tax consultancy
6420	Activities of holding companies	7111	Architectural activities
6499	Other financial service activities; except insurance and pension funding n.e.c.	7312	Media representation
6511	Life insurance	7420	Photographic activities
6611	Administration of financial markets	7430	Translation and interpretation activities

TABLE V: Description of SSIC shown in Fig. 15.

proximity with each other, and we term it the co-occurrence index. The third index captures the tendency of firms making synchronous temporal change to their location choice, and we call it the co-change index. We showed that the single and combined use of these indices enabled the deciphering of various spatial and morphological behaviour of firm-types against time. This includes the concentration or diffusion behavior of firms, as well as the co-movement and co-location of a set of firms into or out of a certain region of space in a city.

Specifically, we employed the indices to the study of the set of industrial firms in Singapore and UK. Interestingly, we observe certain common firm location choice patterns between the two countries through our analysis. We found that the sectors: "Information and Communication", "Financial and Insurance Activities", "Real Estate Activities", and "Professional, Scientific and Technical Activities", are spatially sticky to the Central Area of Singapore as well as the London region, while the sub-sector "Other activities auxiliary to financial servcies" displays a consistent tendency to diffuse out of CA and London. There are also opposite movement trends, with a proclivity of IT and digital service firms to concentrate in London but diffuse out of CA, while the media, financial services, and life insurance firms to diffuse out of London but concentrate in CA. In the case of spatial co-occurrence, Singapore sub-sectors demonstrate

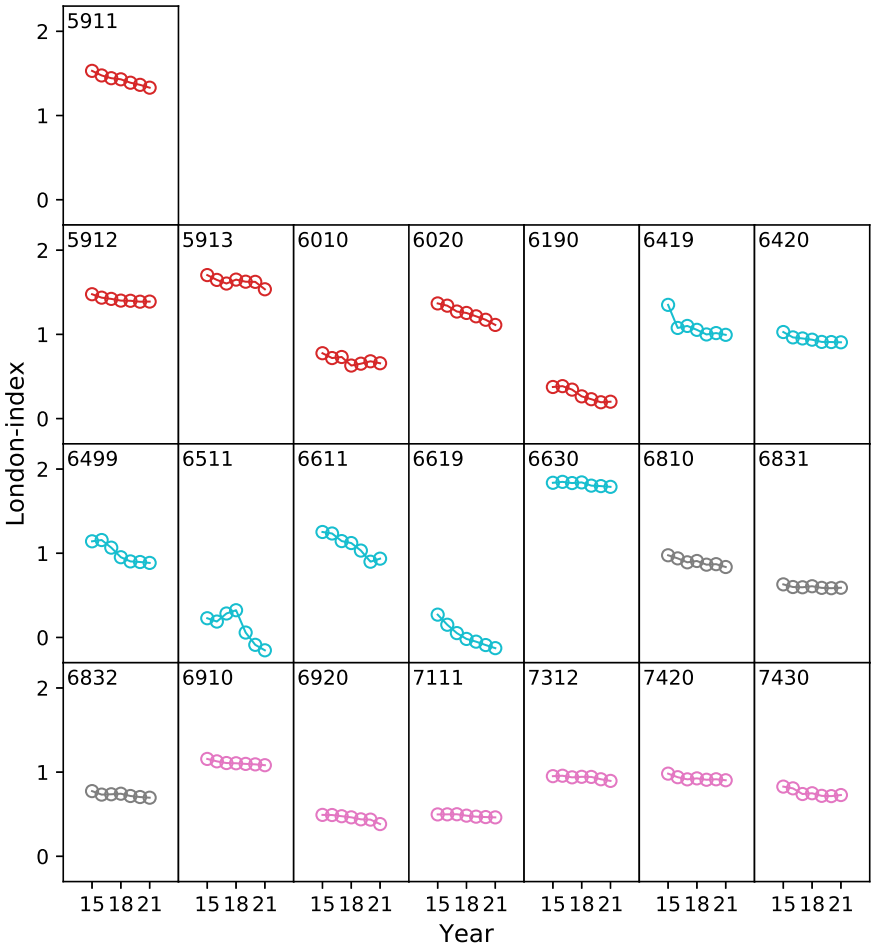

FIG. 15: London-index for industries in the United Kingdom from 2015 to 2021. The industries indicated here are from sections J (red), K (cyan), L (grey) and M (pink), with London-index decreases consistently throughout the period under study (i.e., correlation smaller than -0.75).

an inclination to co-occur within the same section, while UK firms shows more mixing of sub-sectors of the diverse sections in the co-location pattern. Nonetheless, both countries share the common feature of co-location of sub-sectors within the "Construction" section. Finally, our methods have also successfully uncovered sub-sectors that co-occur, co-change, and spatially sticky to the CA of Singapore and the London region: "Wholesale bank" and "General reinsurance" for Singapore; and "Real estate agencies" and "Management of real estate on a fee or contract basis" for UK. The fruitful application of our approach to this comparative analysis of firms in Singapore and UK indicate the applicability of our method to cities in general.

Through our analysis on the spatial-temporal dynamics of firms in this study, we have identified locations which have become more or less attractive for specific sectors over time, and which types of firms are more or less space agnostic. Collectively these insights may be useful to inform the formulation of polycentric strategies by identifying business activities which may be location agnostic and open to moving out of CA. Furthermore, the analysis has enabled us to understand how the locational preference or land use requirements of certain firms are changing over time.

Our methods are useful as a first analysis to decipher potential location's stickiness and co-occurrence behaviors among sub-sectors. Further studies can then be proposed based on the findings to validate these behaviors through

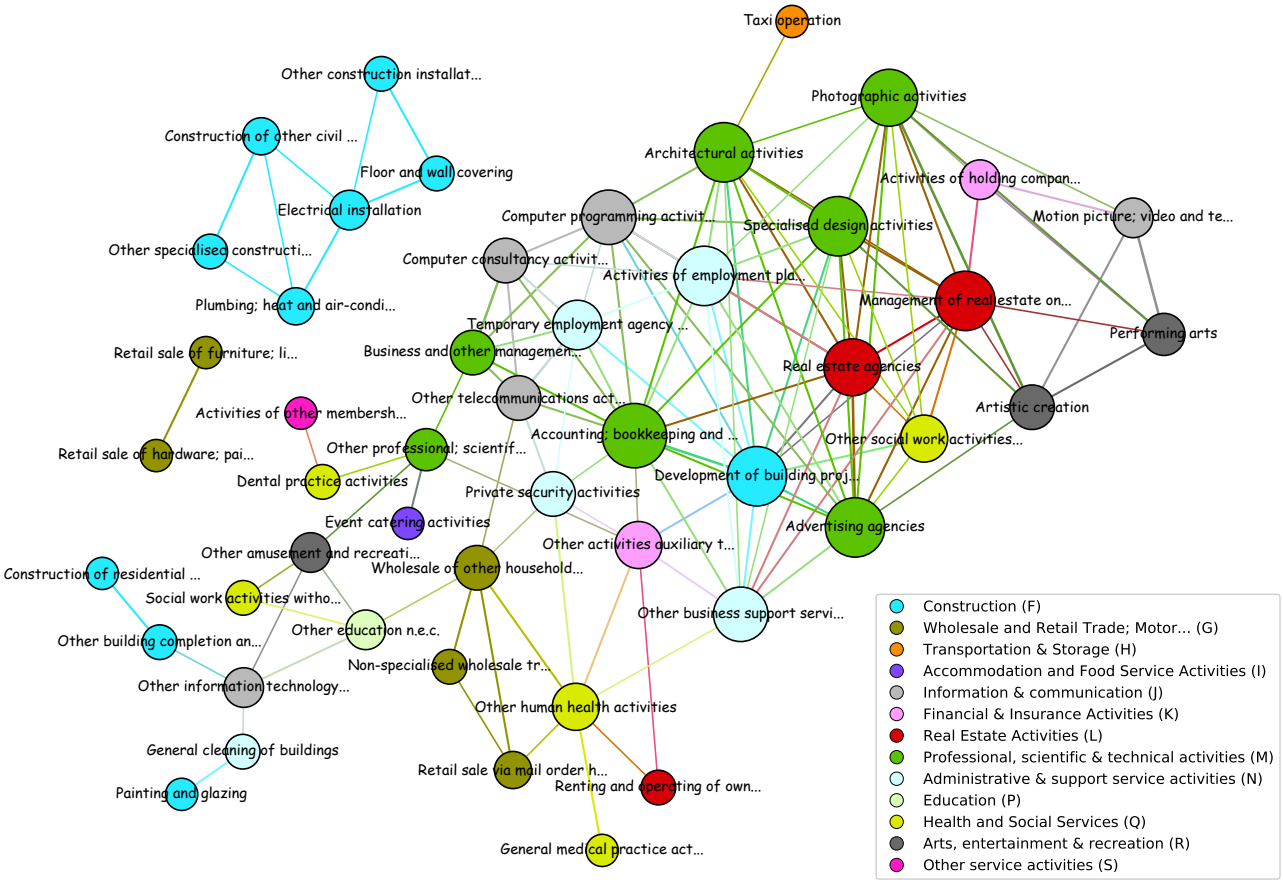

FIG. 16: Co-occurrence network of UK's industries in 2015. Only industries with more than 5000 firms are included. Here, a threshold of $Co_{ij} \geq 0.98$ is used to determine whether two nodes are linked. Note that a higher threshold is used here because of the course-grained's nature of the dataset. Sub-sectors in the same section are shown as nodes of the same color. Note that the legend gives the name of the section with its category in bracket. Node size is proportional to node degree.

ground truths.

[1] Marshall, A. (1920). *Principles of Economics.* London: MacMillan.

[2] Jaffe, A. B., Trajtenberg, M. & Henderson, R. (1993). Geographic localisation of knowledge spillovers as evidence by patent citations. *Quartely journal of Economics, 108*(3), 577–598.

[3] Doeringer, P. B. & Terkla, D. G. (1995). Business strategy and cross-industry clusters. *Economic development quarterly, 9*, 225–237.

[4] Breschi, S., & Lissoni, F. (2009). Mobility of skilled workers and co-invention networks: An anatomy of localized knowledge flows. *Journal of Economic Geography, 16*(2), 1–30.

[5] Chen, L., Hasan, R. & Jiang, Y. (2021). Urban Agglomeration and Firm Innovation: Evidence from Developing Asia. *Policy Research Working Paper: No. 9861.* World Bank, Washington, DC. http://hdl.handle.net/10986/36637 License: CC BY 3.0 IGO.

[6] Duranton, G. & Kerr, W. R. (2015). The Logic of Agglomeration *NBER Working Paper No. w21452.*

[7] Leitham, S., McQuaid, R. W. & Nelson, J. D. (2000). The infuence of transport on industrial location choice: a stated preference experiment. *Transportation Research Part A, 34*, 515 – 5335.

[8] Porter, M. E. (2000). Location, competition, and economic development: Local clusters in a global economy. *Economic development quarterly, 14*(1), 15–34.

[9] Parr, J. B. (2002). Agglomeration economies: Ambiguities and confusions. *Environment and Planning A, 34*(4), 717–731.

[10] Duranton, G. & Puga, D. (2004). Micro-foundations of urban agglomeration economies. *Handbook of regional and urban*

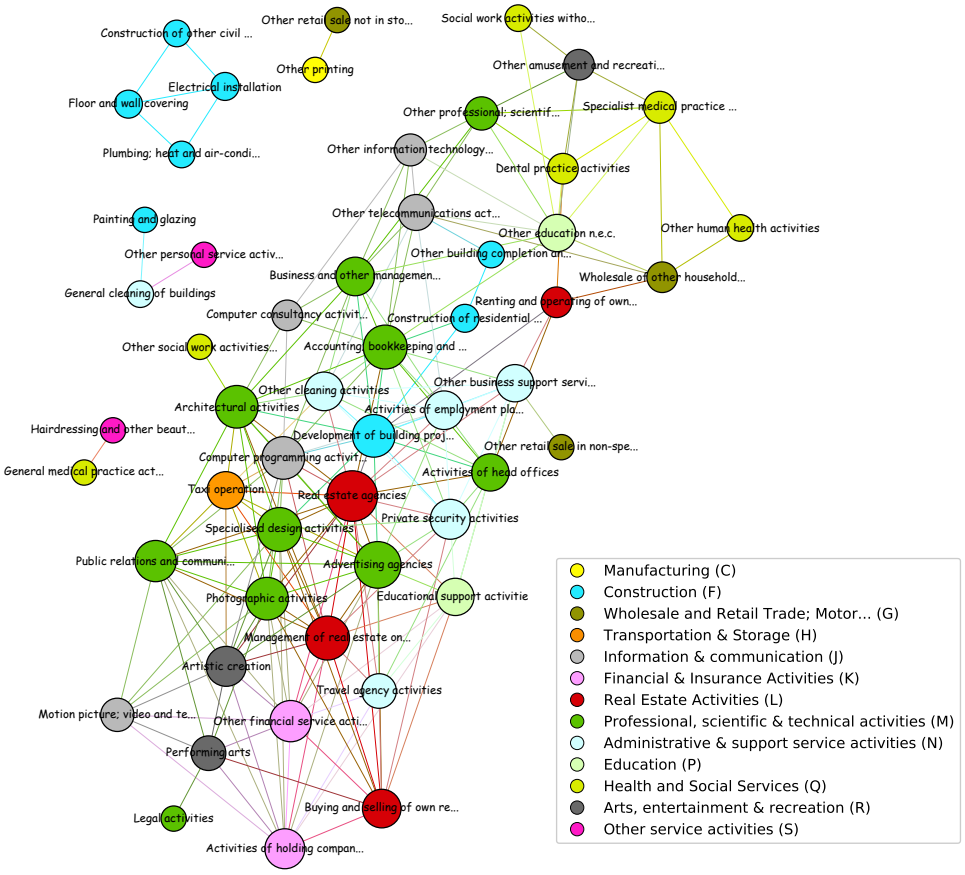

FIG. 17: Co-occurrence network of UK's industries in 2021. Only industries with more than 5000 firms are included. Here, a threshold of $Co_{ij} \geq 0.98$ is used to determine whether two nodes are linked. Note that a higher threshold is used here because of the course-grained's nature of the dataset. Sub-sectors in the same section are shown as nodes of the same color. Note that the legend gives the name of the section with its category in bracket. Node size is proportional to node degree.

economics, 4, 2063–2117.

[11] Flores, M., Villarreal, A. & Flores, S. (2017). Spatial Co-location Patterns of Aerospace Industry Firms in Mexico. *Appl. Spatial Analysis, 10,* 233–251.

[12] Arcaute, E., Molinero, C., Hatna, E., Murcio, R., Vargas-Ruiz, C., Masucci, A. P. & Batty, M. (2016). Cities and regions in Britain through hierarchical percolation. *Royal Society Open Science 3,* 150691.

[13] Cottineau, C., Finance, O., Hatna, E., Arcaute, E. & Batty, M. (2019). Defining urban clusters to detect agglomeration economies. *Environment and Planning B: Urban Analytics and City Science, 46*(9), 1611–1626.

[14] Rigby, D. L. & Brown, W. M. (2015). Who benefits from agglomeration? *Regional Studies, 49,* 28 – 43.

[15] Zheng, X.-P. (2007). Economies of network, urban agglomeration, and regional development: a theoretical model and empirical evidence. *Regional Studies, 41,* 559 – 569.

[16] Andersson, M. & Lööf, H. (2011). Agglomeration and productivity: evidence from firm-level data. *Ann Reg Sci, 46,* 601–620. https://doi.org/10.1007/s00168-009-0352-1

[17] Rosenthal, S. & Strange, W. (2004). Evidence on the nature and sources of agglomeration economies. In: Henderson JV, Thisse JF (eds) *Handbook of regional and urban economics.* Elsevier, Amsterdam, pp 2118–2147.

[18] Hervas-Oliver, J.-L., Sempere-Ripoll, F., Alvarado, R. R. & Estelles-Miguel S. (2018). Agglomerations and firm performance: who benefits and how much? *Regional Studies, 52,* 338 –349.

[19] Nilsen, O. L., Tørset, T., Gutiérrez, M. D., Cherchi, E. & Andersen, S. N. (2020). Where and why do firms choose to move? Empirical evidence from Norway. *Journal of Transport and Land Use, 13(1)*, 207—225.

[20] Niedomysla, T., Källströmb, J., Kosterc, S. & Östhd, J. (2019). Interregional migration of business owners: who moves and how does moving affect firm performance? *Regional Studies, 53(4)*, 503 – 516.

[21] Huynh, H. N., Makarov, E., Legara, E. F., Monterola, C. & Chew, L. Y. (2018). Characterisation and comparison of spatial patterns in urban systems: A case study of U.S. cities. *Journal of Computational Science, 24,* 34–43.

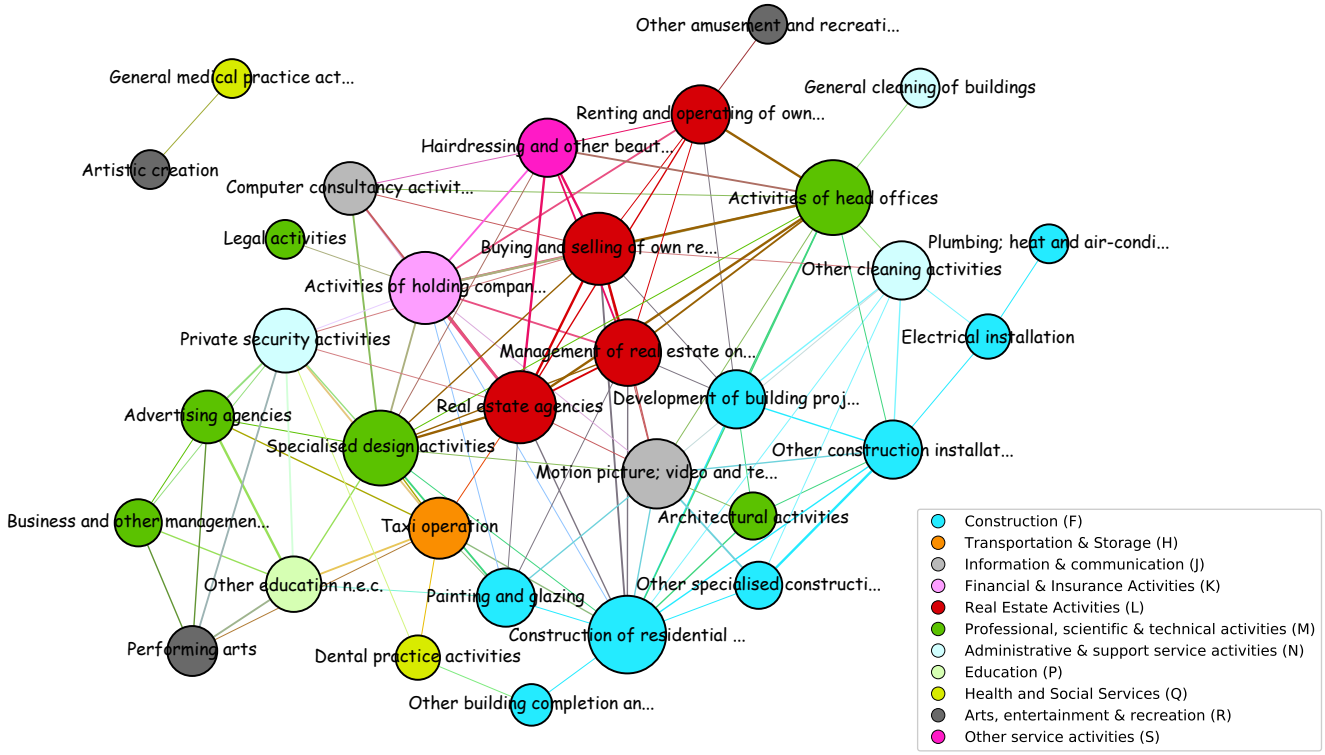

FIG. 18: Co-change network of UK's industries from 2015 to 2021. Only industries with more than 5000 firms are included. Here, a threshold of $Cc_{ij} \geq 0.95$ is used to determine whether two nodes are linked. Note that a higher threshold is used here because of the course-grained's nature of the dataset. Sub-sectors in the same section are shown as nodes of the same color. Note that the legend gives the name of the section with its category in bracket. Node size is proportional to node degree.

[22] Chew, L. Y., Chung, N. N., Sia, W. X., Huynh, H. N., Sim, G., Chua, A. & Huang, Z. (2023). A data analytic elucidation on the spatial agglomeration of Singapore maritime industry. *Accepted for publication by International Journal of Smart and Sustainable Cities.*

[23] Rysman, M. & Greenstein, S. (2005). Testing for agglomeration and dispersion. *Economics Letters, 86*, 405–411.

[24] Osman, T. (2022). A Tale of a Few CBDs: The Uneven Resurgence of Central Business District Economies across U.S. Cities, 1994–2019. *Journal of Planning Education and Research, 0*(0). https://doi.org/10.1177/0739456X221131744

[25] Battino, S., Borruso, G. & Donato, C. (2015). Recreation Tourist Areas. An Exam on Recreational Business Districts in Olbia (Sardinia). In: O. Gervasi *et al.* Computational Science and Its Applications – ICCSA 2015. ICCSA 2015. *Lecture Notes in Computer Science (Part II), vol. 9156*, 662 –674. Springer, Cham. https://doi.org/10.1007/978-3-319-21407-8_47

[26] Accounting and Corporate Regulatory Authority. https://www.acra.gov.sg.

[27] Office for National Statistics. https://www.ons.gov.uk.