Startup Location and Business Attributes in Japan

Chigusa Okamoto

Faculty of Economics, Chuo University

2024/3/29 @CREI International Forum

This study was conducted as a joint research project between for Startups, Inc. and the Center for Real Estate Innovation (CREI) of the University of Tokyo, and is also a result of the collaborative research with the Center for Spatial Information Science (CSIS) of the University of Tokyo under project number 1234.

Startups as Drivers of Innovation

- **Startups** are recognized as critical drivers of innovation.
- Acknowledging the role of startups in maintaining economic vitality, the Japanese government has placed a significant emphasis on nurturing the startup ecosystem.
 - In a landmark move, the Japanese government has declared 2022 as the "First Year of Startup Creation" and launched the "Startup Development Five-Year Plan."

Where are Startups in Japan?

 Startups are densely agglomearted in Tokyo, outnumbering small & medium enterprises (11.6%) and large enterprises (41.1%) in 2016.¹



Top 10 Prefectures in 2022

Concentrated in the central part of Tokyo's 23 Wards



What Local Environments Foster Startups?

Among various potential local factors, we focus on the composition of startups within specific regional units (esp. hexagonal grids with an edge length of about 300m), with a particular emphasis on analyzing the effects of specialization versus diversity.

Agglomeration of startups in specific fields

- e.g., ***tech incubation facilities
- e.g., ***tech innovation districts

Agglomeration of startups in diverse fields

e.g., Multi-field incubation facilities

e.g., Multi-field innovation districts

Which types of clusters are more suitable for startup creation and growth?

Long Debate about Specialization vs. Diversity

- Urban economists have debated which leads to greater urban growth: agglomerating similar industries or agglomerating different industries. (Glaeser et al., 1992).
- In both types, knowledge spillovers within agglomerations are thought to drive productivity, innovation, and urban growth.

Agglomeration type	Knowledge spillover mechanisms
Agglomeration of similar industries (i.e., Specialization)	Casual conversation, imitation, and labor mobility between firms con- tribute
Agglomeration of diverse industries (i.e., Diversity)	Interactions among people with di- verse backgrounds

- We aim to integrate this "specialization versus diversity debate" into the startup ecosystem, focusing on more localized areas.
- To index the diversity of each grid in terms of business attributes, we first (1) detect business fields of startups and then (2) examine the distribution of startups across these fields within each grid.²
- Building on this, we (3) re-examine diversity indices and (4) present part of my ongoing project on this debate.

Data of Startups

- Main data source: STARTUP DB (for Startups, Inc.)
 - This DB contains details about the services offered by each startup, and each service is tagged with business attributes.
 - Tag pairs that are frequently registered together for the same service are considered to be highly relevant.



How to Classify Startup Business Fields

 By using community detection (especially modularity optimization of Newman and Girvan (2004)) in co-occurrence network analysis, tags are classified into several groups.



Tag pairs with a co-occurrence frequency of 100 or more

Detected Business Fields

• Name each business field by the top 2 most frequent tags.

	Business Field	Top 5 Tags Based on Individual Frequency		
1	Media & Entertainment	Media, Entertainment, Education, Learning,		
		Video		
2	Medical & Healthcare	Medical, Healthcare, Construction, Manufac-		
		turing, Environment		
3	IT & Consulting	IT, Consulting, Marketing, SaaS, AI		
4	Retail & EC	Retail, EC, Food, Transportation, Fashion		
5	Finance & Payments	Finance, Payments, Blockchain, Asset Man-		
		agement, Administration		
6	Leisure & Real Estate	Leisure, Real Estate, Travel, Sports, Booking		
7	HR & Recruitment	HR, Recruitment, Job Change, Labor, Crowd-		
		sourcing		

Where are the Agglomeration of Each Business Field?

- The Local Moran I statistic (Anselin, 1995) is used to identify the location of agglomerations.
 - This method tests the presence of local spatial autocorrelation with statistical tests. Significant regions are classified into one of the following categories:

Category	Characteristics
High-High	Both the region and its adjacent regions have high values (Hotspots).
Low-Low	Both the region and its adjacent regions have low values.
Low-High	The region has a low value, but its adjacent regions have high values.
High-Low	The region has a high value, but its adjacent regions have low values.

Weight matrix: 1st order queen type, Significance level: 1%

Location of Clusters in Tokyo's 23 wards in 2022 ALL





Fields 1: Media & Entertainment





Location of Clusters in Tokyo's 23 Wards in 2022

Fields 2: Medical & Healthcare





Fields 3: IT & Consulting





Location of Clusters in Tokyo's 23 Wards in 2022

Fields 4: Retail & EC





Fields 5: Finance & Payments





Location of Clusters in Tokyo's 23 Wards in 2022

Fields 6: Leisure & Real Estate





Fields 7: HR & Recruitment





Areas Colored Red :

Hubs of multiple business fields

- Shibuya & Minato wards, including Shibuya St.
- Chiyoda & Chuo wards, including Tokyo St.

Areas Colored Blue :

Hubs of specific business fields

- Medical & Healthcare: Area around University of Tokyo
- IT & Consulting: Area around Gotanda St.
- Finance & Payment:
 West side of the Imperial Palace



Are Multi-Fields Hubs More likely to Create Startups?

No clear tendency for multi-field hubs to create more startups



Point: mean value, Bar: 95% confidential interval

Regions could be Diverse without Multi-Fields Hubs



So next, we measure each regional diversity using the entropy index.

Preparation before Calculating Entropy Index

- Before calculating the entropy index, which assumes each observation belongs to one classification, we directly classify startups into several groups using hierarchical clustering based on Hoberg and Phillips (2016).
 - We set the number of clusters to match the 7 groups of business fields.

	lag combinations (business fields in pharenthesis)				
	#1 most common	#2 most common			
1	HR(7),Hiring(7)	HR(7),Hiring(7),JobChange(7)			
2	Learning(1), Education(1)	IT(3),Saa $S(3)$,Learning(1),Education(1)			
3	EC(4),Retail(4),Food(4)	Retail(4),Food(4)			
4	R&D(2),DrugDiscovery(2),Medical(2)	Medical(2),MedicalDevices(2)			
5	IT(3),SaaS(3)	Consulting(3),Marketing(3)			
6	${\sf Entertainment}(1), {\sf Game}(1)$	Entertainment(1),Media(1)			
7	Consulting(3)	Energy(2),Environment(2),ElectricPower(2)			

Will Diverse Areas Have More New Startups?

Conduct a Poisson regression analysis using the following formula:

 $E(n_{i,2016-2019}) = \exp[\beta_0 + \beta_1 \times Diversity_{i,2016} + \beta_2 \times Startups_{i,-1} + \beta_3 \times VCs\&CVCs_{i,-1} + f_j]$

n_{i,2016-2019}: # of new startups of grid *i* in year 2016-2019
 Diversity_{i,2016}: Entropy index of grid *i* in year 2016
 Startups_{i,2016}: # of incumbent startups of grid *i* in year 2016
 VCs&CVCs_{i,2016}: # of VC&CVC of grid *i* in year 2016
 f_i: Fixed effects for the main municipality *j*, to which grid *i* belongs

Will Diverse Areas Have More New Startups?

Dependent var.:	$\#$ of new startups $_{i,2016-2019}$			
Edge length of grid :	300m	400m	500m	600m
Diversity i,2016	0.686***	0.425***	0.376**	0.491**
	(0.083)	(0.107)	(0.166)	(0.228)
$\#$ of incumbent startups $_{i,2016}$	0.031***	0.023***	0.020***	0.014***
	(0.003)	(0.005)	(0.005)	(0.003)
# of VC&CVC _{i,2016}	0.164***	0.207***	0.176***	0.117***
	(0.056)	(0.033)	(0.025)	(0.044)
municipality F.E.	Yes	Yes	Yes	Yes
Observations	3,248	2,408	2,002	1,750
Pseudo R^2	0.43803	0.52716	0.59869	0.64784

Clustered (municipality) standard-errors in parentheses. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1

Greater business diversity in 2016 predicted more startups over the next four years.

Conclusions

- We categorized tags representing startup business attributes to identify their business fields.
 - We detected seven business fields: Media & Entertainment, Medical & Healthcare, IT & Consulting, Reatil & EC, Finance & Payments, Leisure & Real Estate, and HR & Recruitment.
- Next, we identified locations where startups related to each business field agglomerate within Tokyo's 23 wards.
 - Startups statistically agglomerate in central areas including Shibuya, Minato, Chiyoda, and Chuo wards.
 - Some regions act as hubs for multiple business fields, while others act as hubs for specific business fields.

Conclusions

- We examined whether regions with higher diversity metrics had a higher number of new subsequent startups.
 - When we calculate the diversity by the entropy index, which reflects the balance of startups Regions with diverse startups tend to have a higher number of new startups thereafter.

Limitation and future work

Limitation

 This analysis uses cross-sectional data, making it difficult to interpret the results as a definitive causal relationship.

Future work

- Is there another appropriate definition of diversity? Given any definition of diversity, how does diversity influence the creation and growth of startups?
- In addition to diversity of business characteristics, does diversity of age and stage of growth matter?

Thank you very much for listening.

References

- Anselin, Luc, "Local indicators of spatial association—LISA," Geographical analysis, 1995, 27 (2), 93–115.
- Glaeser, Edward L, Hedi D Kallal, Jose A Scheinkman, and Andrei Shleifer, "Growth in cities," *Journal of Political Economy*, 1992, *100* (6), 1126–1152.
- Hoberg, Gerard and Gordon Phillips, "Text-based network industries and endogenous product differentiation," *Journal of political economy*, 2016, *124* (5), 1423–1465.
- Newman, M E J and M Girvan, "Finding and evaluating community structure in networks," *Physical review E*, 2004, *69* (2), 026113.
- **岡本, 千草**, "日本のスタートアップ事業分野とその立地パターンについて (Startup Business Fields and Their Location Patterns in Japan)," CREI Working Paper No.16 2024.