Nowcasting Economic Activity with Mobility Data

Kohei Matsumura (Bank of Japan*), Yusuke Oh (BoJ), Tomohiro Sugo (BoJ), and Koji Takahashi (BoJ)

ABFER 8th Annual Conference

June 1, 2021

*Views expressed are those of authors and do not necessarily reflect those of the Bank of Japan. All errors are ours.

Motivation

- Due to the covid-19 pandemic, evaluating economic conditions in a shorter period of time has become extremely important.
- However, traditional statistics have 1 to 2 months lags about current economy.
- High-frequency data may partly resolve this issue.

What we do

- We focus on global positioning system (GPS) data, and show how to use them to nowcast economic conditions from macroeconomic perspectives.
- Specifically, by combining the GPS data with information on coordinates of facilities including shops and factories, we closely examine to which sectors the data can be applied for nowcasting.
- We find that for some sectors, we can nowcast household consumption and firm production with high accuracy.

Outline

1. **Literature review**
2. Data
3. Mobility data and service industries
   - Amusement parks
   - Shopping center
   - Food service industry
4. Mobility data and manufacturing industry
5. Conclusion

Related literature

➢ Our paper is mainly related to two strands of literature.

1. Application of human mobility based on smartphone GPS data
   - Couture et al. (2021)
     - ✓ develop a location exposure index and a device exposure index using mobility data.
   - Watanabe and Yabu (2020)
     - ✓ "stay-at-home" index and quantitatively investigate to what extent Japanese shelter-in-place behavior is explained by intervention effect and information effect.
2. Studies on the development of nowcasting indexes
   - Cajner et al. (2019)
     - ✓ develop the ADP-FRB active employment index to understand the labor market conditions
   - Aprigliano et al. (2019)
     - ✓ B-to-B and B-to-C payment data help to improve the accuracy of nowcasting GDP, business fixed investment, and consumption.

Contribution

- Our study extends these two bodies of literature and contributes to the development of nowcasting indexes of economic conditions, showing the usefulness of mobility data as a tool for nowcasting macroeconomic activity.

Outline

1. Literature review
2. **Data**
3. Mobility data and service industries
   - Amusement parks
   - Shopping center
   - Food service industry
4. Mobility data and manufacturing industry
5. Conclusion

Data

- Main data records number of people who stayed in each mesh.
- Mesh defined as **100m×100m** square.
  - ✓20 mil. distinct meshes.
  - ✓The size of meshes, conceptually equivalent to the resolution of microscopes, plays a crucial role. Targets smaller than this resolution difficult to see.
- Hourly data from Jan. 2017 to Mar. 2020.
- However, raw data won't tell us anything. Have to relate each mesh to economic activities.

Mesh ID	Year	month	day	hour	population
XX	2017	1	1	0	1
XX	2017	1	1	1	2
XX	2017	1	1	2	30
⋮	⋮	⋮	⋮	⋮	⋮
34 bil. records in total					
⋮	⋮	⋮	⋮	⋮	⋮
ZZ	2020	3	31	20	1
ZZ	2020	3	31	23	10

Data (cont.)

- ➢ How can we combine mesh with economic activities?

1. Combine with other data
   - Economic Census
     - ✓ Plant level data including addresses, number of employees, sales, etc.
   - National Land Numerical Information
     - ✓ Data on coordinates of stations and airports, and use of lands.
   - Information provided by private companies
     - ✓ Data about coordinates of facilities of interest (e.g. restraurants) based on their names and addresses.
2. Infer from the raw data themselves

   - ✓ The population distributions help us infer the type of economic activities
   - ✓ Machine learning methods (e.g. clustering) can be useful.

Construction of indices

- We construct Economic Indicator from GPS data (EIG) as the following:

1. For sector J, we calculate the weighted total population $TotalPop_t^J$

$$TotalPop_t^J = \sum_{i \in I_t^J} w_{it} \times Pop_{it}$$

I_t^J ⋯ the set of meshes related to sector J at time t
w_{it} ⋯ weight obtained from other data source (e.g. Census micro data)

2. Define the normalized index EIG_t^J for reference point of time s.

$$EIG_t^J = \frac{TotalPop_t^J}{TotalPop_s^J} \times 100$$

Note that Pop_{it} is available on hourly basis, so is the EIG index.

Outline

1. Literature review
2. Data
3. **Mobility data and service industries**
   - ➢Amusement parks
   - ➢Shopping center
   - ➢Food service industry
4. Mobility data and manufacturing industry
5. Conclusion

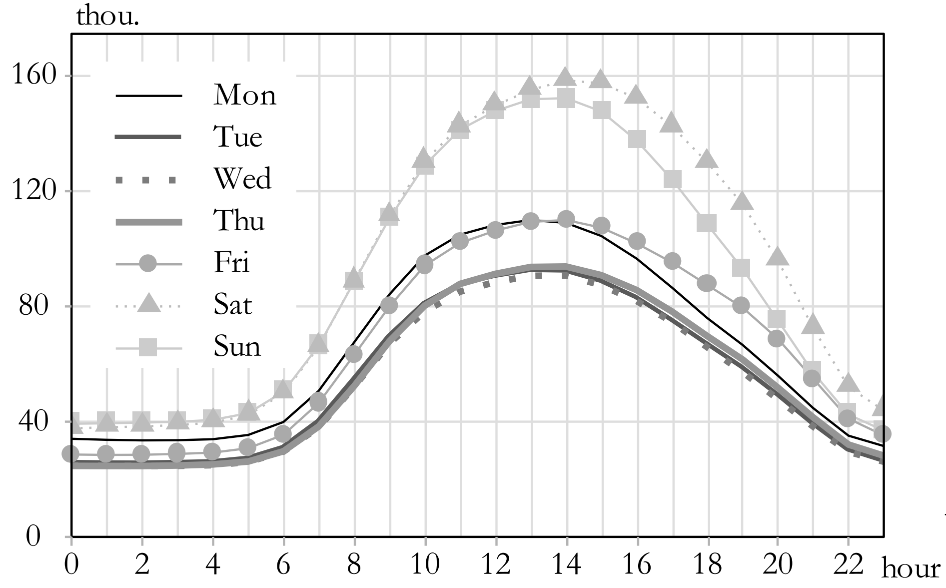
Service industries

- In service industries, the number of people in a commercial facility is likely to represent the number of its customers.
- Focus on three sectors with three different approaches
  1. Amusement park : prior knowledge
  2. Shopping center : prior knowledge ＋ Additional data
     - ✓ Facilities often larger than the size of mesh (100m × 100m).
     - ✓ Naïve aggregation expected to provide good results.
  3. Food service : prior knowledge ＋ Additional data + Statistical method
     - ✓ Facilities usually smaller than the size of mesh.
     - ✓ Need to disentangle activities.

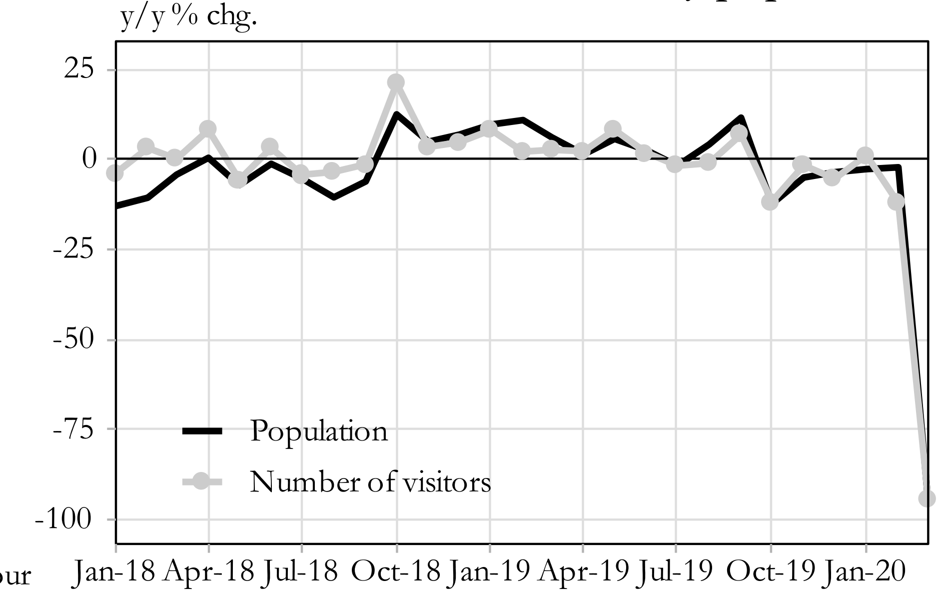
Amusement Parks

- Population approximates the number of visitors
- Small number of major parks and possible to spot meshes corresponding to them on the map.

Hourly population by day of the week



Indicator based on the hourly population



Sources: Agoop; Ministry of Economy, Trade and Industry

Shopping Center

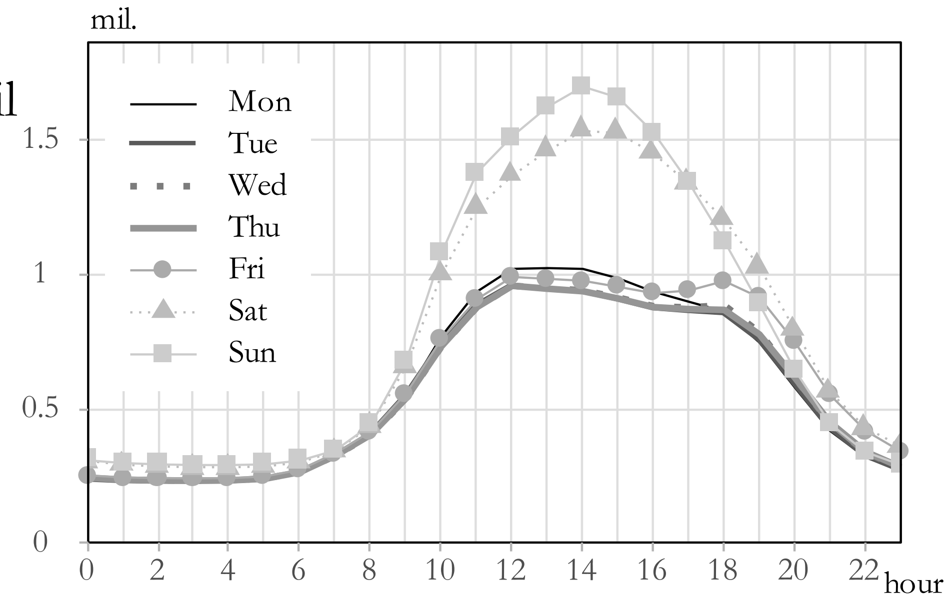
- Cannot manually check all SCs on the map. However, address information of major SCs is available. Some meshes may contain irrelevant facilities (e.g. stations).

<u>Selection of meshes:</u>

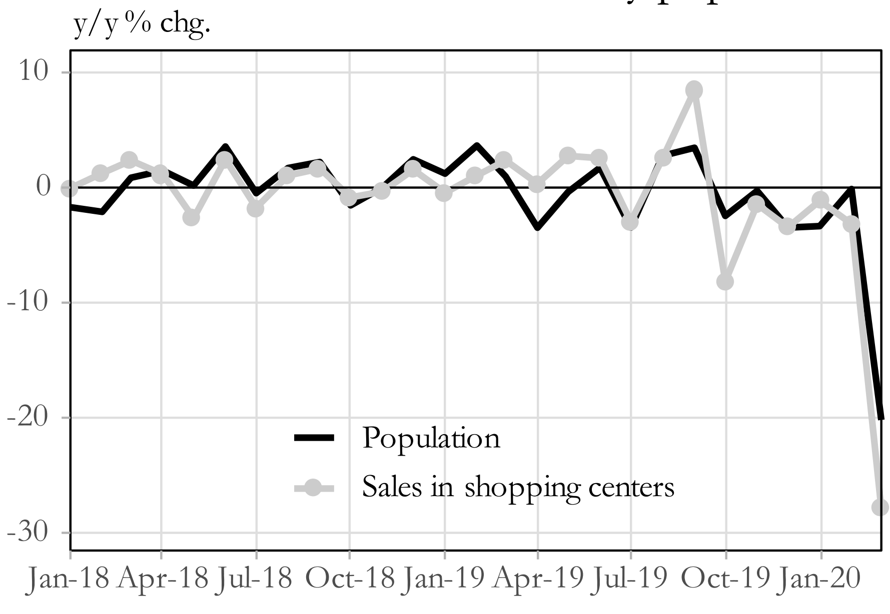
1. Extract meshes based on addresses of SCs listed by Japan Council of Shopping.

2. Exclude meshes containing stations, and meshes where the population in weekends is smaller than that in weekdays (unlikely to be SCs).

Hourly population by day of the week



Indicator based on the hourly population



Sources: Agoop; Japan Council of Shopping Centers; Ministry of Land, Infrastructure, Transport and Tourism.

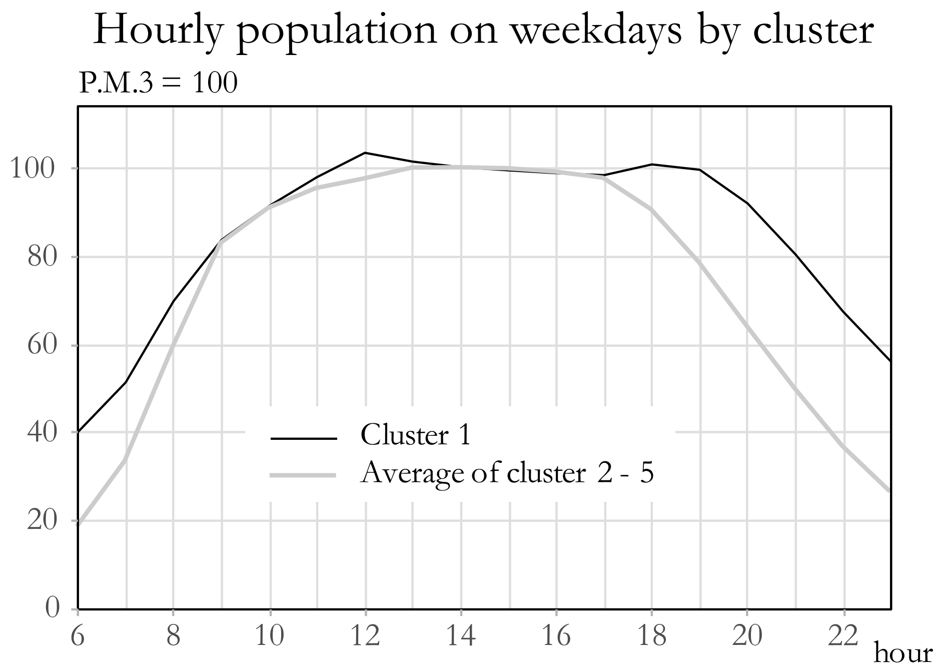
Food service

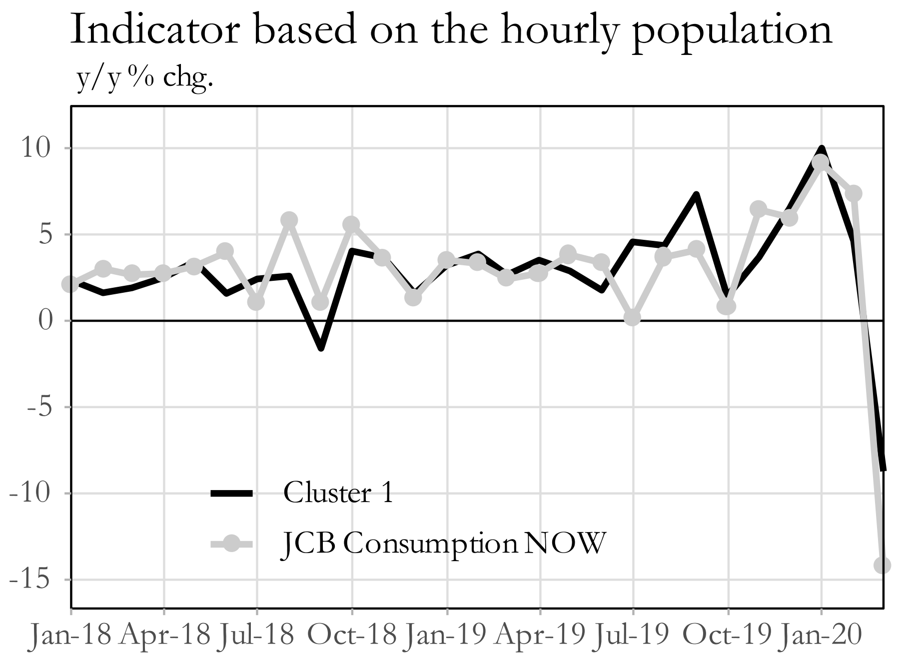
- In general, a restaurant is small compared to the size of meshes (100m × 100m) and the number of restaurants is enormous.
- Usually located in areas where many various different activities are conducted. Difficult to remove unrelated meshes

<u>Selection of meshes:</u>

1. Focus on non-residential areas based on the ratio of population in daytime and nighttime (following Mizuno et al. (2020), $daytime \geq 0.8 \times nighttime$).
2. Using Grunavi (restaurant guide service) API, we extract meshes for which more than 300 restaurants are located within 300m from its center.
3. Cluster the selected meshes into 5 groups, using k-medoids.

Food service (cont.)





Sources: Agoop; Gurunavi; Ministry of Land, Infrastructure, Transport and Tourism; NOWCAST, Inc./ JCB, Co., Ltd., "JCB Consumption NOW."

Service industries: wrap up

- Since the pandemic hit service industries severely, a high correlation in turbulent times does not guarantee high performance in normal times.
- However, we only use data up to March 2020, and show that even before the serious deterioration of the pandemic, the indicator demonstrates a high correlation with other statistics.

Outline

1. Literature review
2. Data
3. Mobility data and service industries
   - Amusement parks
   - Shopping center
   - Food service industry
4. **Mobility data and manufacturing industry**
5. Conclusion

Labor input and Production

- In general, production is determined by many different factors including capital utilization, technology, etc., and the impact of each factor differs across sectors.
- Still, labor input is an important factor for almost all sectors.
- We try to capture labor input by population in factories. For sectors where labor input is a dominant factor for production, our index can be useful.

Labor input and Production (cont.)

- The correlation between industrial production and labor input differs a lot across industries: high in Production machinery and low in Electronic parts, devices and electronic circuits.

- Seem promising for Production machinery and Transportation equipment.

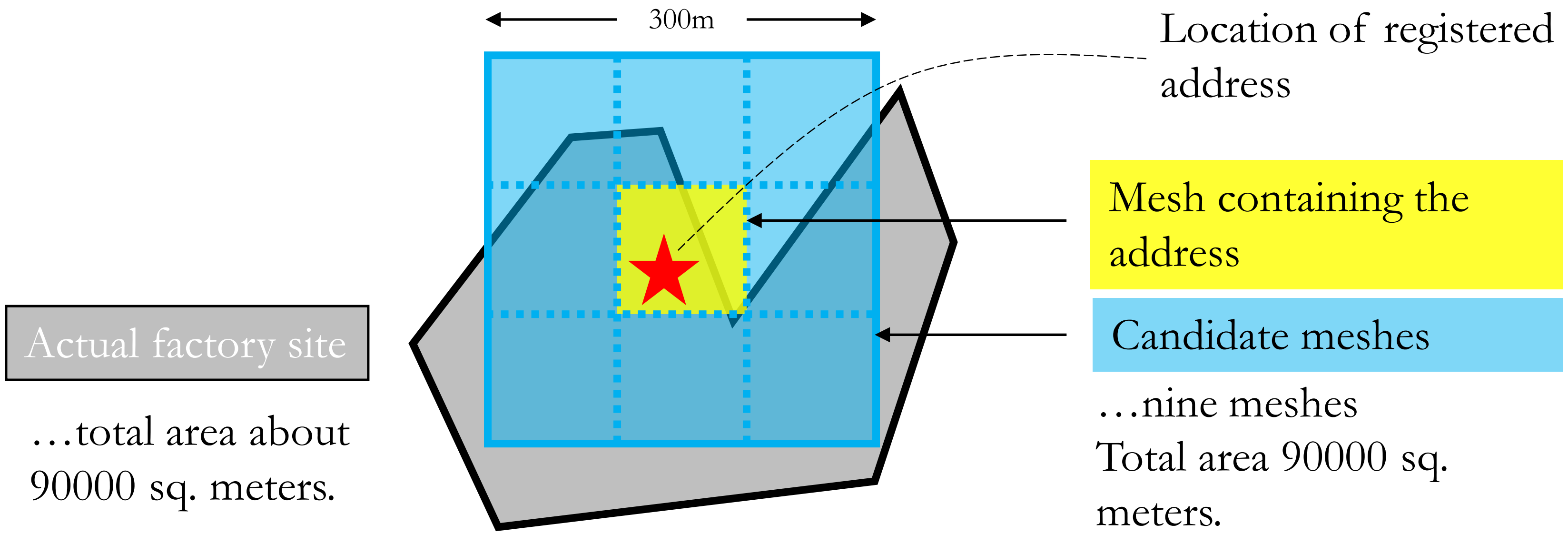
Correlations between labor input and production

Industry	Correlation
Production machinery	0.75
Transportation equipment	0.72
Fabricated metal products	0.43
Plastic products	0.23
Pulp, paper and paper products	0.18
Ceramic, stone and clay products	0.15
Electronic parts, devices and electronic circuits	0.12
Food, beverages, and tobacco	0.05

Note: correlation based on y/y % chg. of period Jan. 2014 - Dec. 2019. Labor input = Total hours worked × Regular employees (from the Monthly Labour Survey).
Sources: Ministry of Economy, Trade and Industry; Ministry of Health, Labour and Welfare.

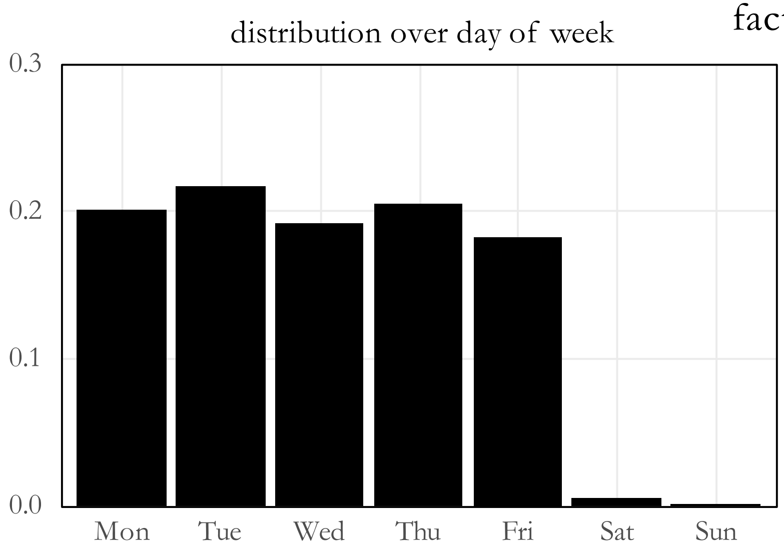
Identifying meshes representing factories

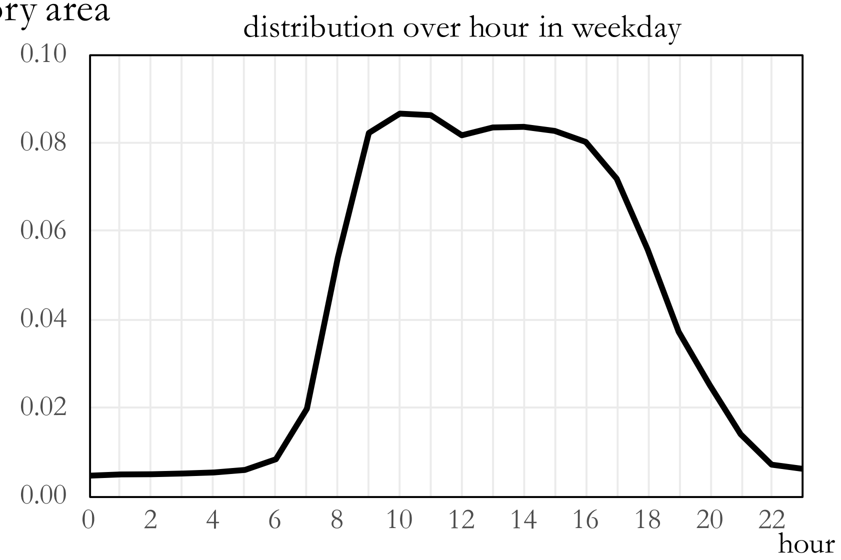
1. Among factories listed in Economic Census for Business Activity whose area are above 10,000 m^2, we select top 10,000 factories ranked by value-added.
2. To determine the set of meshes on which we aggregate population, we set candidate meshes around the registered address according to its area.
3. Focus on several features (Sunday ratio, etc.) to remove unrelated ones.



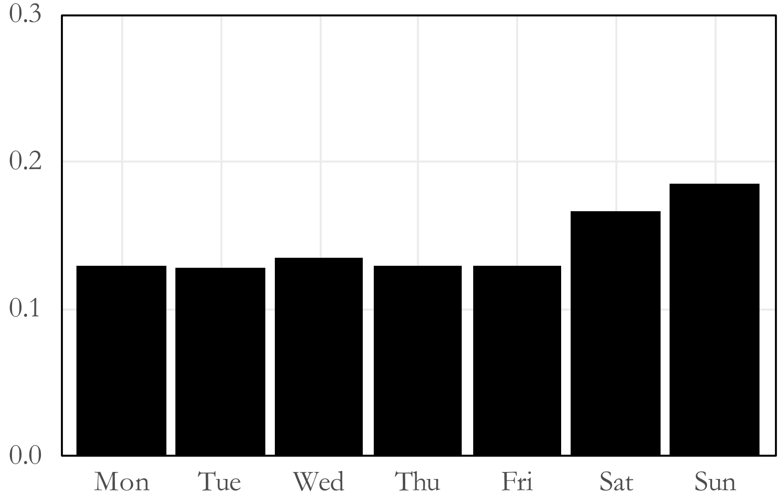
Identifying meshes representing factories (cont.)

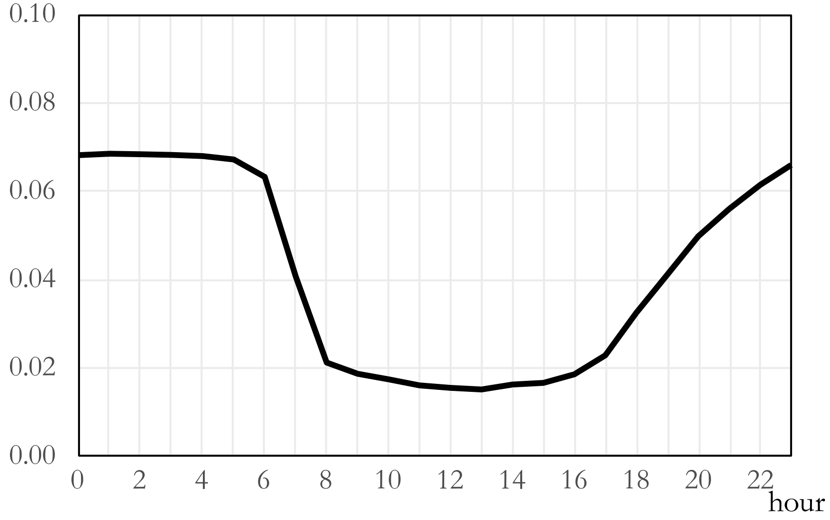
factory area





residential area





Source: Agoop.

Identifying meshes representing factories (cont.)

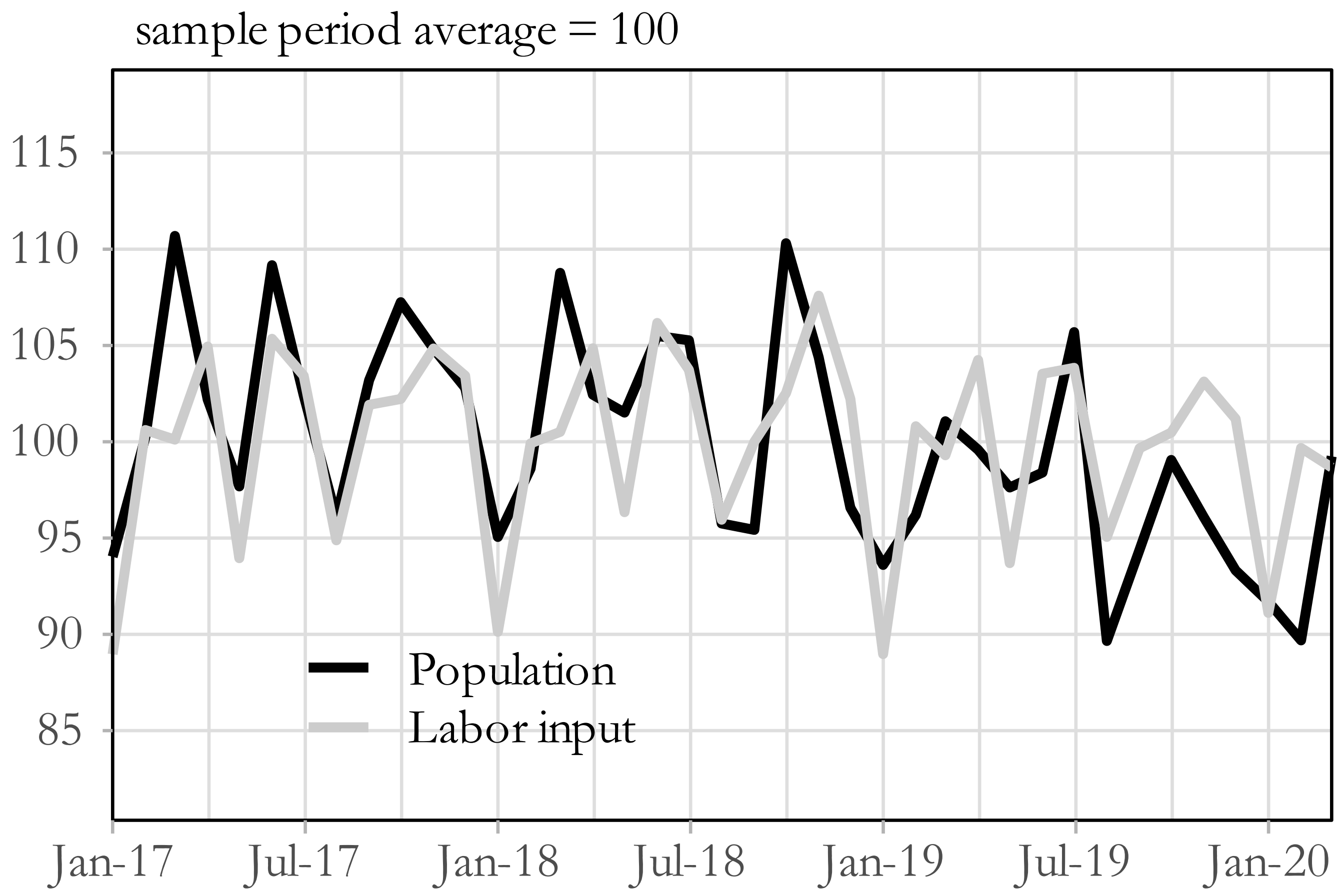
- For each industry, we grid search three thresholds such that the correlation between population and labor input (from official statistics) is highest.

Criteria for Sunday and daytime ratios, and average population

Industry	Sunday ratio upper threshold	Day time ratio lower threshold	Average population lower threshold	Correlation
Transportation equipment	0.5	0.6	10	0.86
Production machinery	0.1	2	40	0.69
Information and communication electronics equipment	0.9	1	80	0.68
General-purpose machinery	0.5	1.8	40	0.63
Food, beverages and tobacco	0.1	1	60	0.60
Business oriented machinery	0.5	2	10	0.54
Non-ferrous metals and products	0.1	1.4	20	0.49
Electrical machinery, equipment and supplies	0.1	0.8	10	0.42
Electronic parts, devices and electronic circuits	0.9	1.4	10	0.31
Chemical and allied products	0.9	0.6	10	0.20

Note: Sunday ratio = average population on Sunday / average population on weekdays. Daytime ratio = average population from 9:00 a.m. to 4:59 p.m. / average population from midnight to 4:59 a.m.
Sources: Agoop; Ministry of Health, Labour and Welfare.

Population and labor input



Sources: Agoop; Ministry of Health, Labour and Welfare.

Population and industrial production

- Summing population over all hours and days does not necessarily approximate labor inputs well. Fluctuations can be driven by facility maintenance workers.
- Grid search three parameters to determine the best time window to aggregate population.

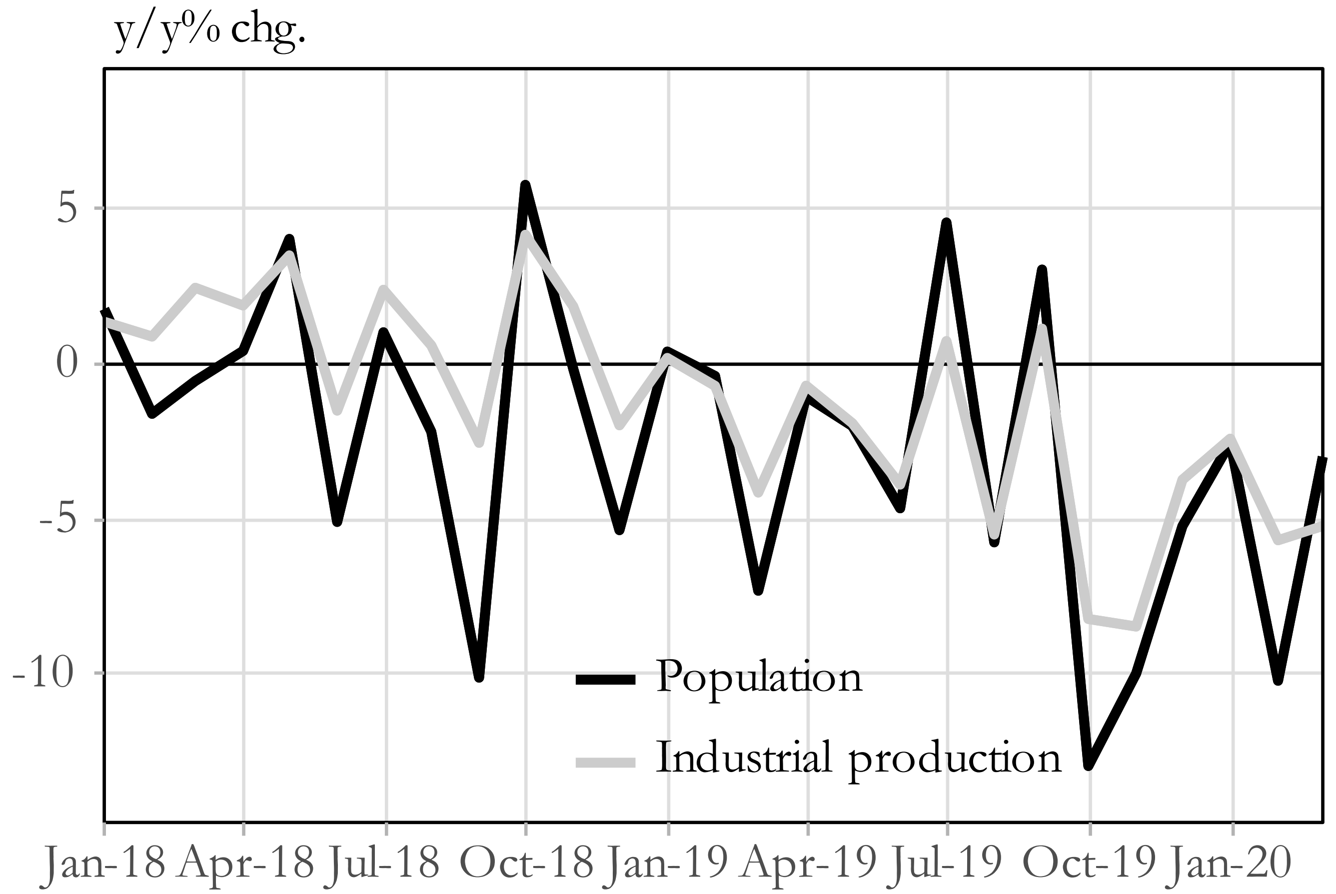
Criterion for holiday inclusion and hours

Industry	Conditions			Correlation
	Holidays included	Start hour	End hour	
Transportation equipment	yes	17	18	0.85
Production machinery	yes	18	21	0.82
Iron, steel and Non-ferrous metals	yes	8	12	0.78
Electrical machinery, and information and communication electronics equipment	no	16	17	0.72
Fabricated metals	no	16	19	0.71
Plastic products	yes	17	18	0.69
General-purpose and business oriented machinery	yes	10	11	0.65
Pulp, paper and paper products	no	9	10	0.61
Ceramics, stone and clay products	yes	14	15	0.55
Electronic parts and devices	yes	8	9	0.36
Chemicals	no	16	17	0.35
Foods and tobacco	no	22	23	0.30

Note: "Start hour" and "end hour" indicate the start and end time of the selected sample window, respectively.
Sources: Agoop; Ministry of Health, Labour and Welfare.

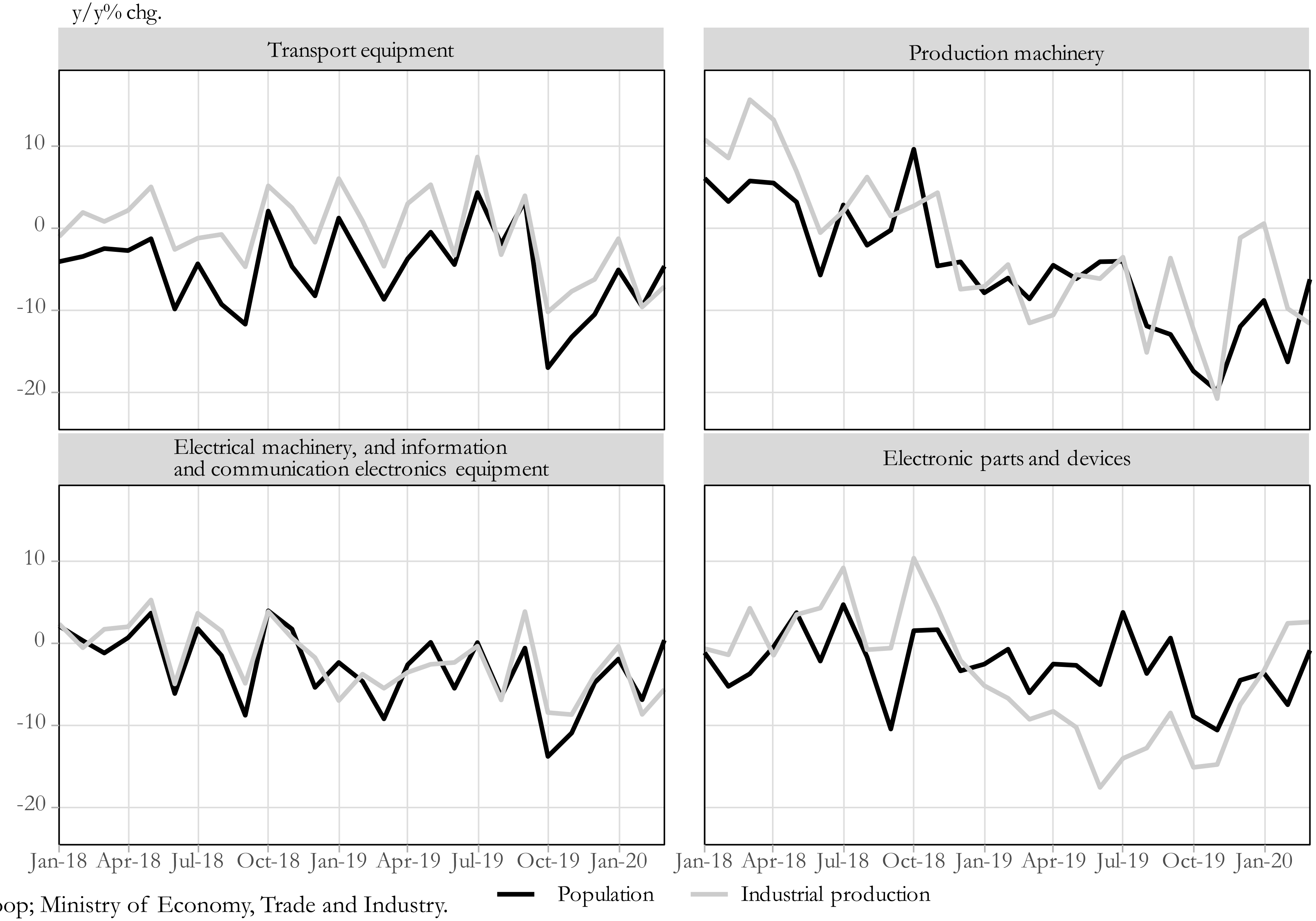
Population and industrial production (cont.)

Indicator based on the hourly population for industrial production



Sources: Agoop; Ministry of Economy, Trade and Industry.

Population and industrial production (cont.)



Sources: Agoop; Ministry of Economy, Trade and Industry.

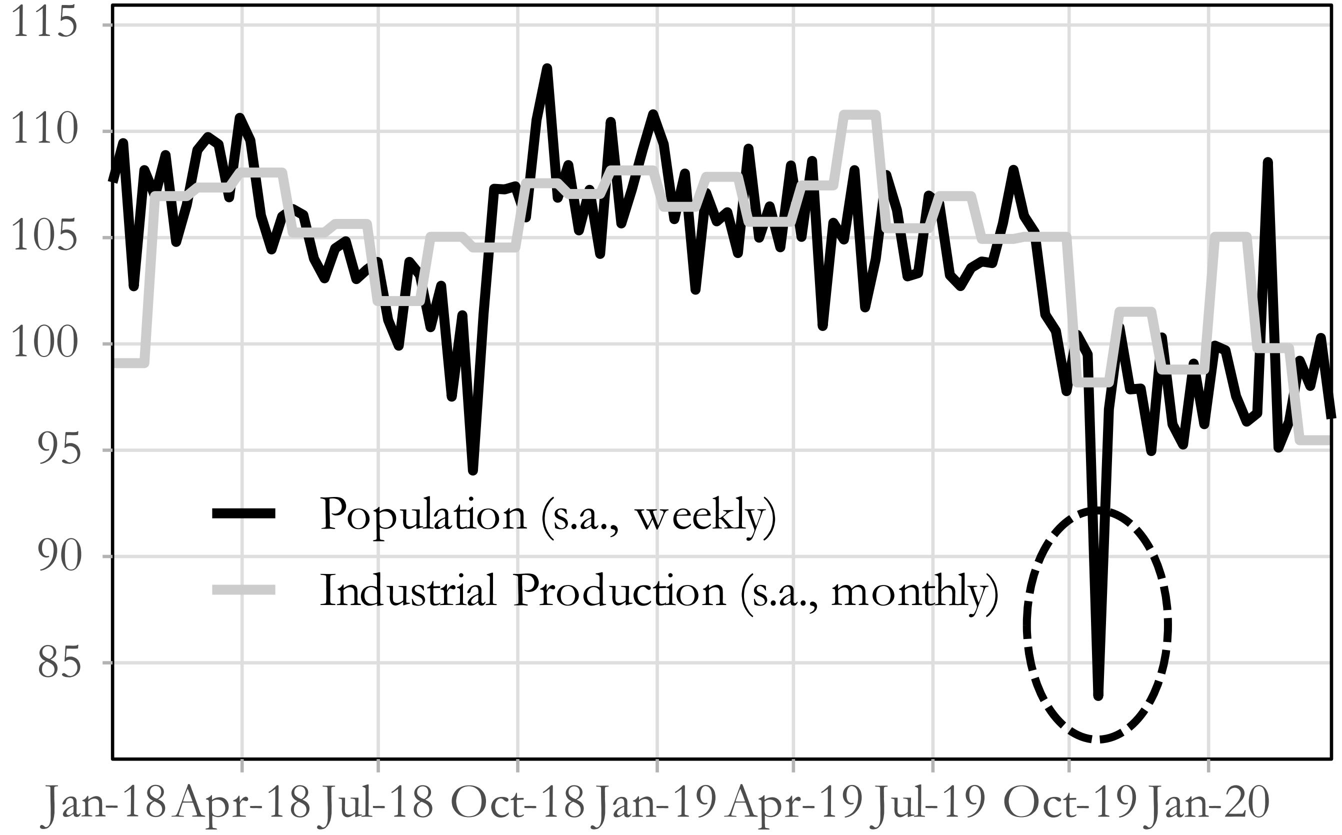
Extension: High-frequency Analysis

- We construct weekly index by re-aggregating our indices at the weekly level and removing seasonal effects.

  ✓Can see the outstanding effect Typhoon Hagibis caused on production in late Oct. 2019.

Weekly production indicator (transportation equipment)

CY 2017=100



Sources: Agoop; Ministry of Economy, Trade and Industry.

Outline

1. Literature review
2. Data
3. Mobility data and service industries
   - Amusement parks
   - Shopping center
   - Food service industry
4. Mobility data and manufacturing industry
5. **Conclusion**

Key takeaways

- We develop indicators to capture sales in the service industries and production activity in manufacturing industries using mobility data.
- In the service industries, our indicators are useful for nowcasting activity not only in sectors where the typical site area is large, but also in sectors where site area is small compared to the size of meshes.
- In the manufacturing industry, the indicator provides good nowcasts for sectors where the labor input is a dominant factor for production.

Caveats

- Panel data of places rather than people: can’t see who were there.
- Cannot reflect the behavior of people who do not use smartphones.
- Sample period limited: difficult to check the plausibility of our analysis.
- Critical to utilize other data to detect and verify structural changes.
- Unable to measure activity not driven by population.

References

- **Aprigliano, Valentina, Guerino Ardizzi, and Libero Monteforte**, "Using payment system data to forecast economic activity," International Journal of Central Banking, 2019, 15 (4), 55–80.
- **Cajner, Tomaz, Leland D Crane, Ryan A Decker, Adrian Hamins-Puertolas, and Christopher Kurz**, "Improving the accuracy of economic measurement with multiple data sources: The case of payroll employment data," 2019. Finance and Economics Discussion Series 2019-065, Board of Governors of the Federal Reserve System.
- **Couture, Victor, Jonathan I Dingel, Allison Green, Jessie Handbury, and Kevin Williams**, "Measuring movement and social contact with smartphone data: a real-time application to COVID-19," Journal of Urban Economics, 2021, 103328.
- **Watanabe, Tsutomu and Tomoyoshi Yabu**, "Japan's voluntary lockdown," 2020. CARF Working Paper.