

**28<sup>th</sup> Meeting of the Wiesbaden Group, 2-6 October 2023**

## **Experimental Use of New Data Sources for Prompt Identification of Changes in Firms' Status**

Peh Li Lin, Seng Li Cheng and Cui Hui Min<sup>1</sup>

Business Statistics Division, Singapore Department of Statistics

### **Abstract**

Singapore Department of Statistics (DOS) manages the Statistical Business Register (SBR), which serves as the foundational statistical database for the compilation of business and economic statistics. Firm characteristics from the SBR are also frequently used in statistical compilation and analytical studies to examine Singapore's enterprise landscape and to support data-driven policymaking.

In response to increasing data demands, DOS has been developing its capabilities in the use of new data sources to supplement administrative and survey data in updating of the SBR and producing economic indicators. This paper will share DOS's pilot project on "Early Triggering", which aims to promptly identify changes in firms' economic status through leveraging on various higher frequency data. In this pilot project, DOS has tapped on higher frequency administrative data sources of employment and revenue and explored the experimental use of meter data on electricity usage and stamp duty data. Qualitative data such as firms' future restructuring plans from business expectations survey and web-scraped text from business news and company announcements were also used to supplement the administrative data.

---

<sup>1</sup> Email of the authors: [peh\\_li\\_lin@singstat.gov.sg](mailto:peh_li_lin@singstat.gov.sg), [seng\\_li\\_cheng@singstat.gov.sg](mailto:seng_li_cheng@singstat.gov.sg) and [cui\\_hui\\_min@singstat.gov.sg](mailto:cui_hui_min@singstat.gov.sg).

# 1 Introduction

Significant changes in business activities of firms, especially those of large firms, may have a sizeable impact on Singapore's business statistics, National Accounts and Balance of Payment statistics. These statistics are collected and compiled by different divisions in the Singapore Department of Statistics (DOS). It is important that the various DOS divisions detect and capture such changes in a timely manner to ensure consistent and reliable statistics across subject domains. This requires close coordination between the relevant DOS divisions.

To facilitate the early detection of significant changes in firm's business activity, DOS embarked on an Early Triggering<sup>2</sup> pilot project. The project aims to detect entries and exits of large firms or firms with substantial changes in revenue and/or employment. The identified firms are then shared with the relevant DOS divisions for a coordinated assessment of the potential impact on short-term surveys and economic indicators to achieve a consistent statistical treatment.

This paper will share DOS's efforts in this pilot project on "Early Triggering", which leverages on the Statistical Business Register (SBR) and higher frequency data, including qualitative data from online sources.

## 2 Coverage and Maintenance of the SBR

DOS manages the Statistical Business Register (SBR), which serves as the foundational statistical database for the compilation of business and economic statistics. The SBR is updated primarily using administrative data, supplemented with statistical survey returns from DOS and Research & Statistics Units (RSUs) in government ministries and statutory boards.

The coverage of the SBR includes all entities registered in Singapore. Basic identification (e.g., Unique Entity Number (UEN)<sup>3</sup>, enterprise name) and firm characteristics (e.g., registration date, status) are readily available from the various registration authorities. In the SBR, firms are uniquely identified by their UENs, which are issued to firms upon their registration in Singapore. UENs are used by firms in their interactions with the Government, such as the application of business licenses and permits or filing of tax returns. Therefore, the UEN enables DOS to process and integrate firm-level data from various sources efficiently and accurately.

The primary data sources for SBR are administrative data which have comprehensive coverage and good data quality. However, some administrative data are only filed once a year, with the information available more than one year after the firm's financial year ending.

Increasingly, data users are requesting for faster indicators to detect economic changes early which will facilitate policy agencies to prepare timely responses. In response to these data demands, DOS has been exploring new data sources in updating the SBR and development of more timely economic indicators. In particular, the following data are well suited to promptly identify changes in firms' business activities:

- 1) Employment from the Central Provident Fund Board (CPF Board) and Ministry of Manpower (MOM), and

---

<sup>2</sup> It is also known as Early Warning Systems (EWS) in other countries, as explained in "European Business Profiling Recommendations Manual".

<sup>3</sup> For more information, please refer to <https://www.uen.gov.sg>.

- 2) Goods & Services Tax (GST) data from the Inland Revenue Authority of Singapore (IRAS)

### **3 Early Triggering with Employment and GST Data**

The primary administrative data sources for Early Triggering are higher frequency administrative data sources of employment from the Central Provident Fund Board (CPF) and Ministry of Manpower (MOM) as well as Goods & Services Tax (GST) data from the Inland Revenue Authority of Singapore (IRAS).

#### Employment data

The Central Provident Fund (CPF) is a mandatory social security savings scheme in Singapore funded by contributions from both employers and employees. Employers of local workers<sup>4</sup> are required to declare wage information and pay CPF contributions for all local workers on a monthly basis. On the other hand, firms employing foreign workers must apply relevant work passes with the MOM. Employers of foreign employees are also required to update MOM if there are changes such as revisions in the employee's salary or updates to employee's personal particulars. As such, data from CPF and MOM comprise timely information on local and foreign employment and wages which can be used in the updating of the SBR.

#### Goods & Services Tax (GST) data

Goods & Services Tax (GST) is a broad-based consumption tax levied on the purchases of goods and services in Singapore<sup>5</sup>. Firms with annual taxable turnover exceeding S\$1 million must register for GST while firms with annual taxable turnover not exceeding S\$1 million can voluntarily register for GST. GST data are timely as firms are required to file GST on a quarterly basis, within one month after the end of the accounting period. Data collected in GST returns comprise information such as firms' revenue and taxable purchases, which can serve as a reliable and timely data source to update firm's economic status and performance in the SBR.

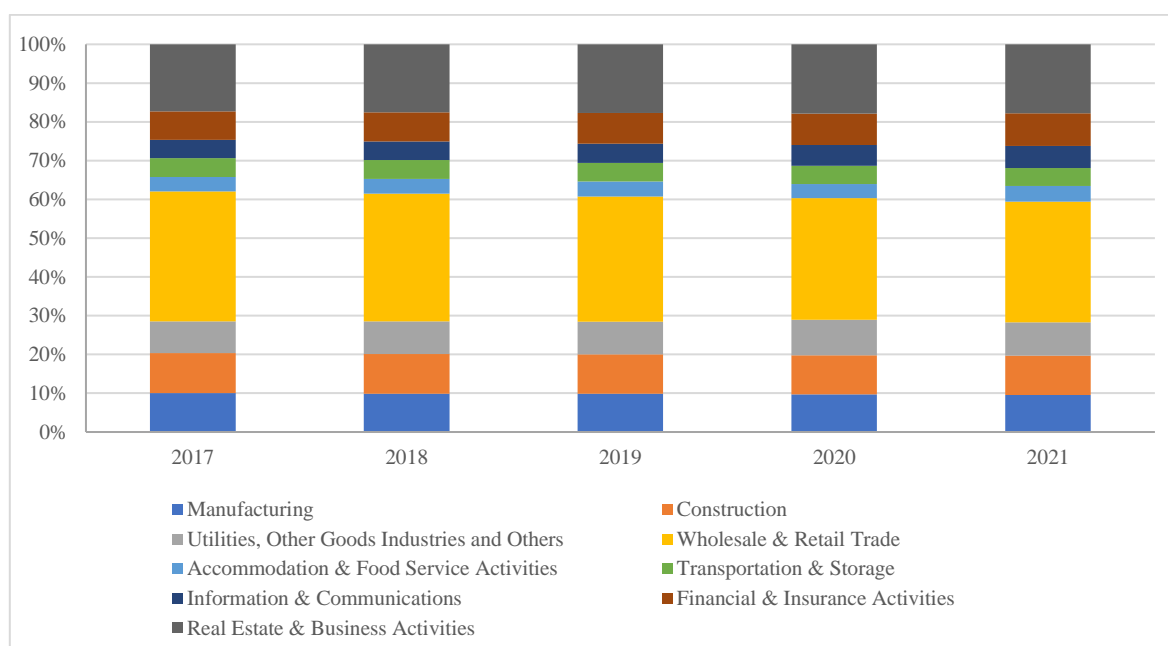
Between 2017 to 2021, Wholesale & Retail Trade accounted for the largest share of GST registered business entities, followed by Real Estate & Business Activities. On average, the two sectors collectively made up about half of all GST registered business entities [Exhibit 1].

---

<sup>4</sup> Refer to Citizens and Permanent Residents of Singapore.

<sup>5</sup> GST is also known as the Value-Added Tax (VAT) in other countries. For more information, please refer to IRAS website: [https://www.iras.gov.sg/taxes/goods-services-tax-\(gst\)](https://www.iras.gov.sg/taxes/goods-services-tax-(gst)).

## Exhibit 1: Sectoral Composition of GST Registered Business Entities in Singapore, 2017 - 2021



Source: Inland Revenue Authority of Singapore (IRAS)

### Early Triggering with Employment and GST data

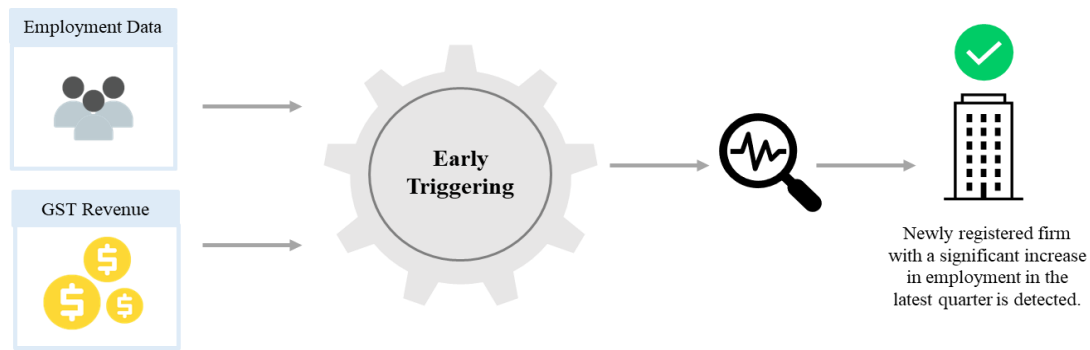
Early Triggering is more effective when the triggering criteria is tailored to industry characteristics. Sector-specific quantitative thresholds are adopted instead of having only a single quantitative threshold. The sector-specific quantitative thresholds are developed based on the revenue and employment percentile distributions of firms by industry and discussions with relevant DOS divisions.

On a quarterly basis, the project team processes the employment and GST data and compares the employment and revenue data in the current quarter with the previous quarter. Firms with changes in employment and/or revenue exceeding the sector-specific thresholds are identified and shared with economic and business statistics divisions for their information and further assessment.

The diagram below shows an example of detecting entry of a large firm from employment data in the Early Triggering project [Exhibit 2]. A newly registered firm recorded a significant increase in employment in the latest quarter but did not report any GST revenue yet. This could occur when a firm is at the pre-production phase and recruitment and training of employees is being conducted prior to the commencement of production activities. The increase in employment is thus a key indicator or an early alert of impending production, where the revenue would be reported much later after the actual sales of goods and services.

The project team would provide relevant DOS divisions with information on the firm and its characteristics (e.g., registration date, economic sector) as well as quarterly employment and/or revenue data over available time periods.

## Exhibit 2: Detect the Entry of a Large Firm in Early Triggering



## 4 Use of qualitative data as supplementary indicators

### Business Outlook

DOS collects qualitative information on changes in firms' business activities, to supplement the quantitative information from employment and GST data. DOS's Quarterly Business Expectations Survey<sup>6</sup> collects information from larger firms about their business outlook and major restructuring plans in the next six months. This information is also shared with the relevant DOS divisions.

### Weekly Economic Brief

In addition, DOS works together with a service provider to sieve through and select online business news based on keywords covering a range of topics such as firm expansion and restructuring plan, mergers and acquisitions, business headwinds and tailwinds. The selected news articles are then classified into pre-defined categories such as retail trade, tourism and internationalization. A summary of the news articles is compiled into a Weekly Economic Brief and shared with DOS officers involved in the data collection and compilation of economic and business statistics.

### Indicator on firm's internet presence

DOS has also implemented a project to text mine web-based data and extract key features from firms' corporate websites. Supervised machine learning is then used to derive an indicator on firms' internet presence. The indicator on firms' internet presence is integrated with firm characteristics available in the SBR to derive new insights for further analyses. One such use case is that firms with corporate websites can be an early indication of firms commencing business activities and will be identified for inclusion in the sampling frame of business surveys.

## 5 Experimental Use of Electricity Consumption and Property Stamp Duty Data in Early Triggering

Most administrative sources are available at the enterprises level, including the employment and GST data used for the Early Triggering project. DOS has also been exploring new data

<sup>6</sup> For more information on Business Expectations Survey, please refer to: [www.singstat.gov.sg/find-data/search-by-theme/industry/business-expectations/latest-data](http://www.singstat.gov.sg/find-data/search-by-theme/industry/business-expectations/latest-data)

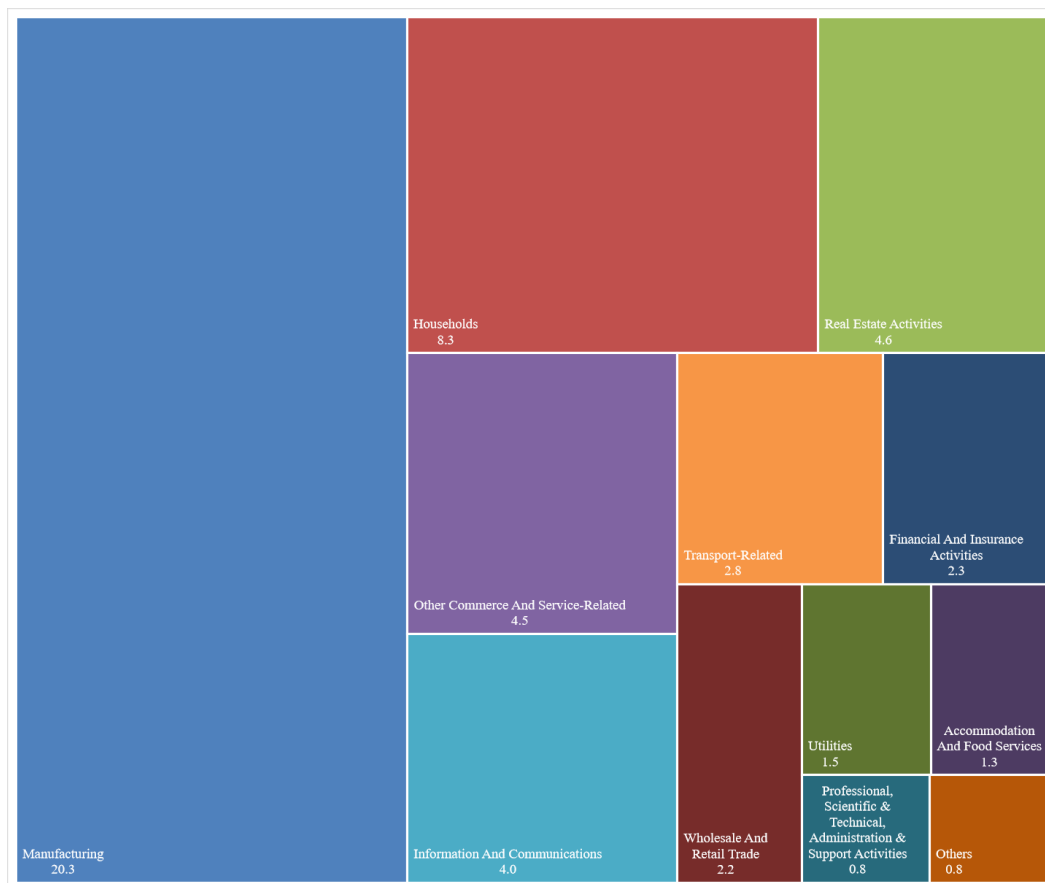
sources of electricity consumption and property stamp duty for incorporation into the SBR and to serve as additional indicators in the Early Triggering project. These new data provide information at the establishment level.

### Electricity Consumption Data

Experiences from other National Statistical Offices demonstrated that electricity meter data can be used in producing official statistics such as electricity consumption of businesses and/or households<sup>7</sup>. Meters are installed in all households and establishments in Singapore to measure electricity consumption. As such, monthly electricity consumption data are available for household and business accounts. Data of electricity consumption for business accounts contain information such as UEN, address of the establishment and amount of electricity (in kWh) consumed in the billing month. Hence, we explored the data for Early Triggering as it offers the advantage of both data timeliness and establishment details.

Singapore’s total electricity consumption stood at 53.5 TWh in 2021. Manufacturing sector was the largest consumer of electricity, accounting for 37.9 per cent of total electricity consumption in the year. This was followed by Household (15.5 per cent) and Real Estate sectors (8.6 per cent)<sup>8</sup> [Exhibit 3].

**Exhibit 3: Singapore’s Electricity Consumption (Terawatt Hours) by Major Sector, 2021**



Source: Energy Market Authority (EMA)

<sup>7</sup> For more information, please refer to EuroStat’s report on “Implementation of Smart Meter Data in the Production of Official Statistics”.

<sup>8</sup> Please refer to <https://www.ema.gov.sg/resources/singapore-energy-statistics> for energy statistics.

## Property Stamp Duty Data

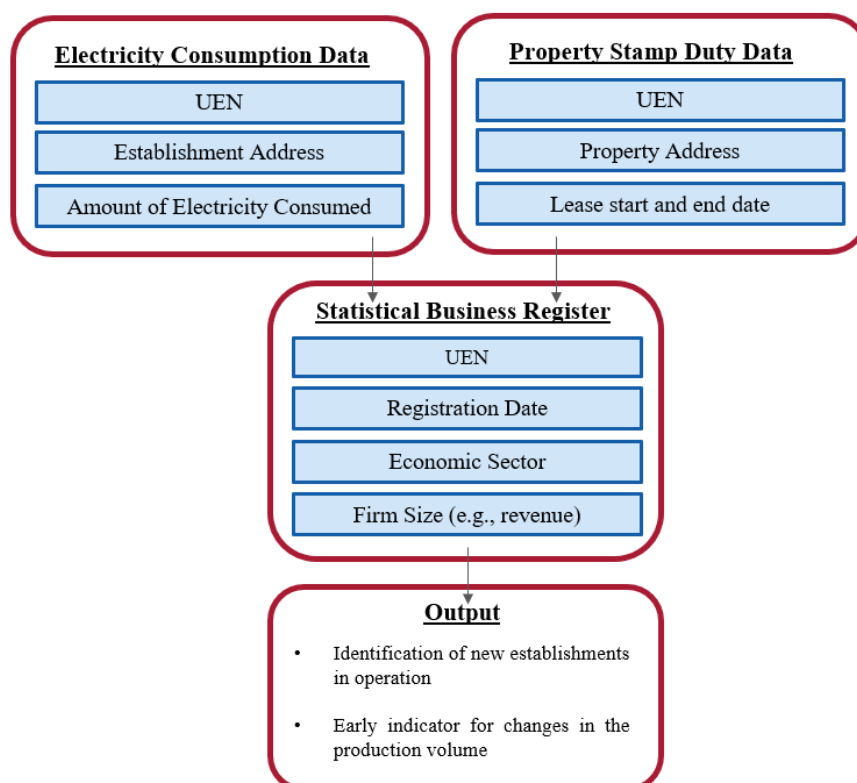
Property stamp duty refers to the tax on documents relating to the transactions of properties in Singapore. The parties involved in the property transactions are required to e-stamp the documents of sales & purchase agreement or tenancy agreement and pay stamp duty<sup>9</sup>. Stamp duty data from IRAS contain sales and purchase information such as UEN, property address and selling price of the property as well as tenancy information such as monthly rental amount and lease start & end date.

## Experimental use of Electricity Consumption and Property Stamp Duty Data in Early Triggering

Integration of electricity consumption and stamp duty data at establishment level with firm characteristics information in the SBR (e.g., registration date, economic sector and firm size) could serve as additional sources for Early Triggering. In particular, information on firms' new electricity meter accounts and leases or purchases of commercial properties could be used to identify new establishments or outlets in operation. If the increase in the number of new establishments exceed a certain threshold, the firm would be identified and shared with relevant DOS divisions by the project team.

In addition, significant changes in electricity consumption can also be used as an early indicator for changes in the production volume [Exhibit 4].

### **Exhibit 4: Use of Electricity Consumption and Property Stamp Duty Data in Early Triggering**



<sup>9</sup> For more information, please refer to IRAS website: <https://www.iras.gov.sg/taxes/stamp-duty>.

## 6 Conclusion

The SBR plays an important role in the data collection, compilation and analyses of economic and business statistics. The SBR is also useful in the early detection of significant changes in firm's business activity using higher frequency data. In addition, qualitative data can provide an early sensing of events that may impact firm's performance and should be used as supplementary indicators where available.

The experimental use of electricity consumption and property stamp duty data also showed that they can be used for updating establishment information in the SBR. DOS will continue to explore the use of new data sources to update the SBR and produce official statistics.

## References

Energy Market Authority (2023): [Singapore Energy Statistics 2022](#)

EuroStat (2020): [European Business Profiling Recommendations Manual](#)

EuroStat (2020): [Implementation of Smart Meter Data in the Production of Official Statistics](#)

EuroStat (2021): [European Business Statistics Methodological Manual for Statistical Business Registers](#)

Singapore Department of Statistics (2022): [Using Big Data to Profile Singapore's Internet Economy](#)

Singapore Department of Statistics (2023): [Experimental Use of Machine Learning and New Data Sources in the Updating of the Statistical Business Register](#)

United Nations (2020): [United Nation Guidelines on Statistical Business Registers](#)

United Nations Economic Commission for Europe (2015): [Guidelines on Statistical business Registers](#)