

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

by Peh Li Lin, Seng Li Cheng and Cui Hui Min
Business Statistics Division
Singapore Department of Statistics

Introduction

Significant changes in business activities of firms, especially those of large firms, may have a sizeable impact on Singapore's business, National Accounts and Balance of Payments statistics. It is important that such changes are promptly detected and captured to ensure consistent and reliable statistics across subject domains.

To promptly identify significant changes in business activities of firms, the Singapore Department of Statistics (DOS) initiated the 'Early Triggering'¹ pilot project. The project aims to detect entries and exits of large firms or those with substantial revenue and/ or employment changes. Identified firms are then assessed by the relevant subject domain teams to gauge the potential impact of their changes on short-term surveys and economic indicators for consistent statistical treatment.

This article shares DOS's efforts in the pilot project on 'Early Triggering' which leverages the Statistical Business Register (SBR) and higher frequency data that includes qualitative data from online sources.

Coverage and Maintenance of the SBR

The SBR serves as the foundational statistical database for compiling business and economic statistics. The SBR, managed by DOS, is primarily updated using administrative data and supplemented with statistical survey returns from DOS and Research & Statistics Units (RSUs) in government ministries and statutory boards.

The SBR encompasses all entities registered in Singapore. Basic identification [e.g., Unique Entity Number (UEN)², enterprise name] and firm characteristics (e.g., registration date and 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 the 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 the SBR are administrative data which have comprehensive coverage and good data quality. However, some administrative data are filed annually and are only available more than one year after the firm's financial year ending.

Increasingly, users are seeking indicators with more recent data to detect economic changes early and facilitate timely policy responses. This urges DOS to explore new data sources to update the SBR and develop more timely economic indicators.

Early Triggering with Employment and Goods & Services Tax (GST) Data

The primary administrative data sources for the Early Triggering project are higher frequency administrative data sources of employment from the Central Provident Fund Board (CPF Board) and Ministry of Manpower (MOM), as well as 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 employers and employees. Employers of local workers³ are required to declare wage information and pay CPF contributions to them

1 Also known as Early Warning Systems (EWS) in other countries, as explained in the 'European Business Profiling Recommendations Manual'.

2 For more information, please refer to the [UEN website](#).

3 Refer to Citizens and Permanent Residents of Singapore.

monthly. Conversely, firms employing foreign workers must apply for relevant work passes from the MOM. Employers of foreign employees are also required to inform MOM of changes such as revisions in employee’s salary and updates to personal particulars. Therefore, data from CPF and MOM provide timely information on local and foreign employment and wages which can be used to update the SBR.

GST Data

GST is a broad-based consumption tax levied on the purchases of goods and services in Singapore⁴. Firms with annual taxable turnover exceeding S\$1 million must register for GST, while those under 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 on GST returns comprise information such as firms’ revenue and taxable purchases, serving as a reliable and timely data source to update their economic status and performance in the SBR.

Early Triggering with Employment and GST Data

Early Triggering is more effective when the triggering criteria is tailored to industry characteristics.

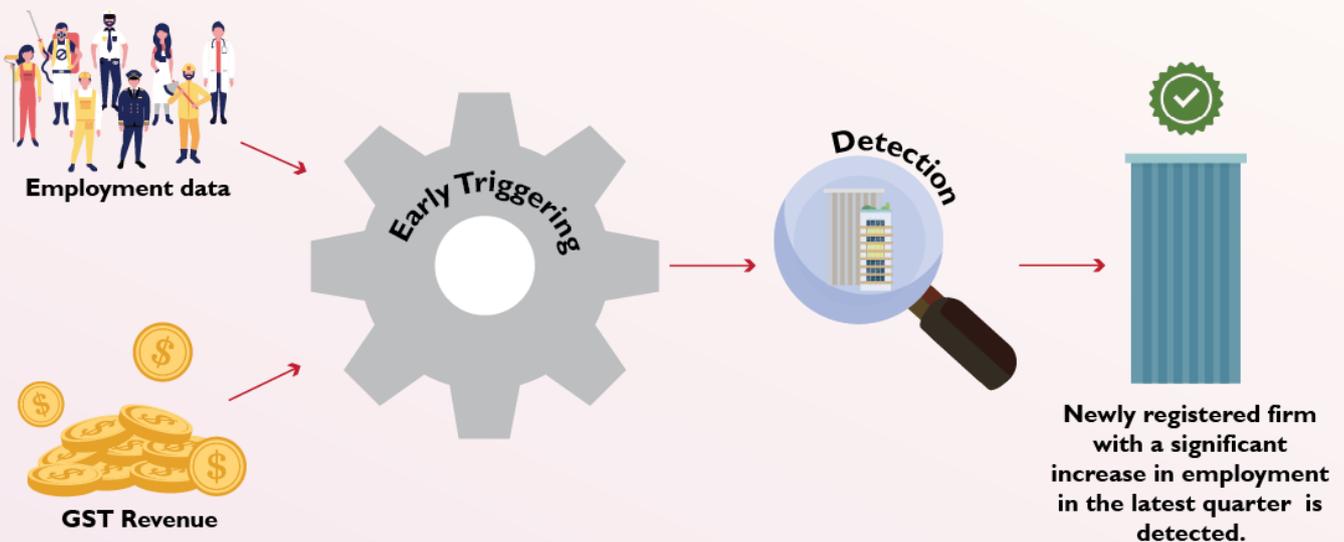
Quantitative thresholds specific to the sector are adopted instead of having a single quantitative threshold to cover all industries. The sector-specific quantitative thresholds are developed in consultation with the relevant subject domain experts, based on the revenue and employment percentile distribution of firms by industry.

Figure 1 illustrates the detection of the entry of a large firm using employment data in the Early Triggering project. The newly registered firm recorded a significant increase in employment in the latest quarter but did not report any GST revenue. This could occur when a firm is at the pre-production phase, where recruitment and training of employees precede the actual production activities.

The increase in employment activities serves as an early indication of impending production, with its revenue expected to be reported after the commencement of sales.

The project team provides the firm’s information such as characteristics (e.g., registration date and economic sector), quarterly employment and/ or revenue data over available time periods to the relevant subject domains teams for developing of timely economic indicators.

FIGURE 1 DETECT THE ENTRY OF A LARGE FIRM IN EARLY TRIGGERING



4 GST is also known as the Value-Added Tax (VAT) in other countries. For more information, please refer to the [IRAS website](#).

Use of Qualitative Data as Supplementary Indicators

Business Outlook

DOS collects qualitative information on changes in firms' business activities, to supplement quantitative data on employment and GST. DOS's Quarterly Business Expectations Survey⁵ collects information from larger firms about their business outlook and major restructuring plans in the next six months.

Weekly Economic Brief

DOS has access to a curation of business news from major media outlets based on keywords covering a range of topics such as firms' expansion and restructuring plans, mergers and acquisitions, as well as business headwinds and tailwinds.

The selected news articles are categorised into pre-defined topics such as retail trade, tourism, and internationalization for easy reference.

Indicator on Firm's Internet Presence

DOS utilises text mining to extract keywords such as 'Shop', 'Cart' and 'Price' from corporate websites. Using these keywords and their frequency, DOS leverages supervised machine learning to classify the firms' websites into different categories (i.e., those with websites and generate revenue directly from their website, or those with websites but do not generate revenue from them) and derive an indicator on firms' internet presence.

This indicator is integrated with firms' characteristics in the SBR to provide new insights for further analyses. For instance, the identification of firms with corporate websites are early indications of business activities to be included in the sampling frame for business surveys.

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

Most administrative sources are available at the enterprise level, such as the employment and GST data used for the Early Triggering project. DOS has been exploring new data sources of electricity consumption and property stamp duty which provide information at the establishment level, to be incorporated into the SBR and used as supplementary indicators in the Early Triggering project.

Electricity Consumption Data

Experiences from other National Statistical Offices demonstrated that electricity meter data can be used in the production of statistics on the electricity consumption of businesses and households⁶. In Singapore, meters are installed in all households and establishments to measure electricity usage, hence monthly consumption data for households and business accounts are available. Business accounts data includes details like UEN, establishment address and the electricity consumed in kilowatt-hours (kWh) for the billing month. Therefore, electricity consumption data offers the advantage of data timeliness and establishment details, which could be useful for the project.

Property Stamp Duty Data

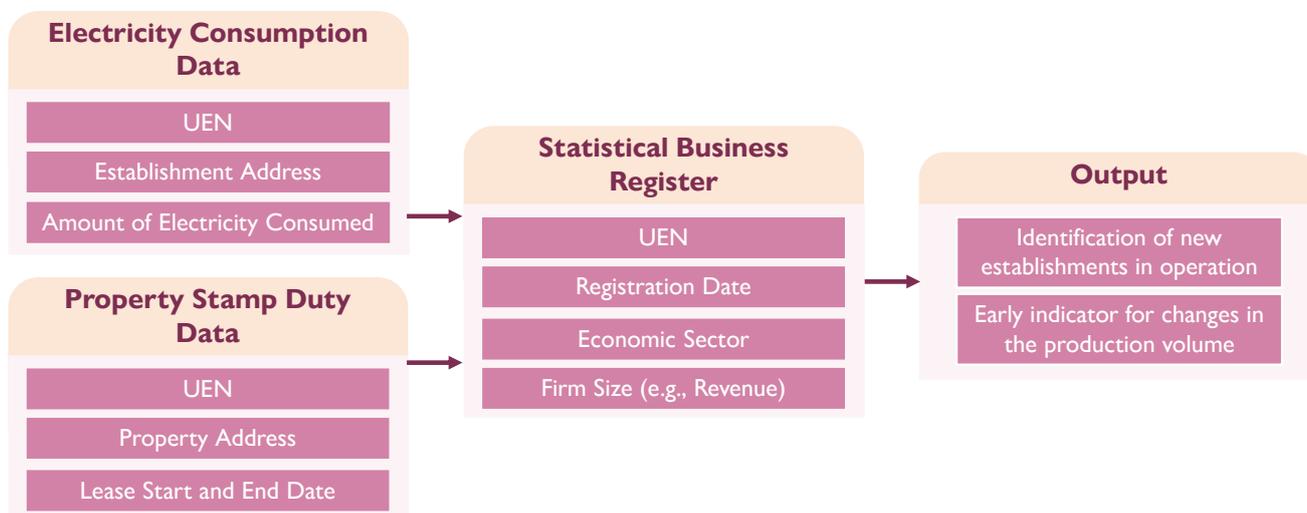
Property stamp duty refers to the tax on documents relating to the transactions of properties in Singapore. Parties involved in property transactions are required to e-stamp the documents of sales and purchase agreement or tenancy agreement and pay stamp duty⁷. Stamp duty data from IRAS contains sales and purchase information such as UEN, property address, the selling price of the property, as well as tenancy information like monthly rents and the start and end date of leases.

5 For more information on the Business Expectations Survey, please refer to the [SingStat Website](#).

6 For more information, please refer to EuroStat's report on "Implementation of Smart Meter Data in the Production of Official Statistics".

7 For more information, please refer to the [IRAS website](#).

FIGURE 2 USE OF ELECTRICITY CONSUMPTION AND PROPERTY STAMP DUTY DATA IN EARLY TRIGGERING



Experimental Use in the Early Triggering Project

Integrating electricity consumption and stamp duty data at the establishment level with firm characteristics in the SBR (e.g., registration date, economic sector and firm size) could be additional sources for Early Triggering.

Specifically, information on firms’ new electricity meter accounts and leases or purchases of commercial properties could facilitate identification of new establishments or outlets in operation. Furthermore, significant changes in electricity consumption can be an early indicator for changes in production volume (Figure 2).

Conclusion

The SBR plays an important role in the data collection, compilation and analyses of economic and business statistics. It aids in early detection of significant changes in firm’s business activities using higher frequency data. Additionally, qualitative data can provide an early sensing of events that may impact the firm’s performance and could be used as supplementary indicators where available.

The experimental use of electricity consumption and property stamp duty data showed the possibility of using them to update establishment information in the SBR. DOS will continue exploring new data sources to update the SBR and produce comprehensive 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](#)