# Malaysia Statistical Business Register (MSBR) Innovates New Insight of Malaysia Trade Statistics

## Abstract

The Malaysia Statistical Business Register (MSBR) plays a central role in the production of Malaysian economic statistics and it has become an essential part of the system of official economic statistics in Department of Statistics Malaysia (DOSM). As the situation changed, users and policy makers keep demanding on additional information on economic statistics. With the aim of reluctance to increase the burden on survey respondents as well as working with certain budget limitation, integrating MSBR with trade database is one of the solutions. The visible outcome of the integration is the international trade by enterprises characteristics (TEC) in which it provides an opportunity to develop new statistics and indicators. The linkage between these two databases can describe the structure and characteristics of trading enterprises; for instance by their economic activities, their employment size, and export performance of SMEs etc. which is unlike traditional trade statistics. In addition, the project can be an efficient tool to improve the quality of MSBR which it can reveal any inconsistencies and contradictions. The unique identifier in MSBR and trade database is the reliable matching approach together with the string matching algorithms approach. Although the benefits of the project are promised, some hurdles still remain in term of data quality, task complexity and confidentiality in micro data linking.

Keywords: Statistical Business Register, Trade by enterprises characteristics, Integrate
# Contents

## I. Contents

I. Contents ............................................................................................................................. 2

II. Introduction ......................................................................................................................... 3

III. MSBR Innovates New Insight of Malaysia Trade Statistics .................................................. 5

   A. Trade by Enterprise Characteristics (TEC) at a Glance ..................................................... 5
      1. TEC by OECD and various NSO ....................................................................................... 5
      2. Benefits of TEC ................................................................................................................. 9
      3. Potential Output of TEC ...................................................................................................10
   
   B. Malaysia Experience on MSBR & Trade Database Linkage ..............................................13
      1. Malaysia Statistical Business Register (MSBR) ...............................................................13
      2. Trade Database ..............................................................................................................14
      3. Integration MSBR with Trade Database ...........................................................................15
      4. Challenges of Success ....................................................................................................19
   
   C. Future Perspectives .........................................................................................................21

IV. Conclusion ....................................................................................................................... 22

V. References ....................................................................................................................... 233
II. Introduction

Setting the foundation for the harmonization of economic statistics, the statistical business register (SBR) has been designed, developed and enhanced by National Statistical Offices (NSOs). SBR plays an essential role in the construction and maintenance of an integrated economic information system. It is used as the source of frames for all business surveys. The growing demand for better and more detailed business statistics has put focus on SBRs and their role in the production of business statistics. In other word, the statistical business register is an essential part of the system of official economic statistics and one of the most important and efficient statistical tools.

Malaysia Statistical Business Register (MSBR) is a fundamental property in maintaining the comprehensive list of businesses and companies operating in Malaysia. The source of this register is mainly from Companies Commission of Malaysia. However, there are other sources such as Construction Industry Development Board, Employees Provident Fund, SME Corporation Malaysia etc. MSBR’s function within Department of Statistics Malaysia (DOSM) is to establish a common framework in the creation of harmonized registers for statistical purposes. It also serves as a tool in the preparation and coordination of surveys as well as a source of information for statistical analysis of population and business demographics. The function of MSBR is further expanded by integrating it with trade database to be able the trade information connects to the traders’ characteristics.

By integrating these two databases, DOSM can be in the better position to provide statistical indicators on trade by enterprise characteristics. Micro-data linking of MSBR and trade is to gain more data insights without initiating new survey. Ultimately, it enriches the international trade
statistics by providing closer views of traders who are actively engaged in importing and exporting. This project is proposed to complement conventional trade statistics for an in-depth trade statistics of which the traditional trade statistics are unable to portray its business characteristic, economic activity, employment size, SME status etc.

The purpose of the paper is to highlight the initiative on trade by enterprise characteristics by DOSM and its challenges. The next section, Section A, provides an overview on trade by enterprises characteristics; production of TEC statistics by other National Statistical Offices (NSOs); benefits; as well as TEC’s potential to meet users and stakeholder needs. Section B presents the Malaysia experience on MSBR & trade database linkage which includes a brief explanation on MSBR and trade database; the linking process and challenges faced by DOSM. Finally, the paper ends with the future perspective on TEC (Section C) and a short conclusion.
III. MSBR Innovates New Insight of Malaysia Trade Statistics

A. Trade by Enterprise Characteristics at a Glance

Exporting and importing is not everyone’s business. Traditional trade statistics record what types of goods are trading across borders between countries but they do not describe the characteristics of the businesses that are behind of these trade flows. In order to know the actor actually engaged in cross border trade, trade data should be linked to the information of enterprises. These identification information can be obtained from SBR, such as name and address, main economic activity of businesses, type of ownership, employment size class, turnover etc. The linkage of trade statistics with business registers allows us to describe those who are engaged in global market, and what are their characteristics. On the other hand, TEC is to complement business data with detailed information on trade.

1. Trade by Enterprise Characteristics by OECD and various NSOs

The development areas of integrating SBR with external trade database is to link trade statistics more closely to other statistics. TEC can provide an opportunity to develop new statistics and indicators by using existing data sets. Organization for Economic Co-operation and Development (OECD) promotes the work on Trade by Enterprise Characteristics (TEC) by disaggregating trade flows according to the characteristics of trading companies in a very efficient and cost effective way (“An integrated approach of trade in goods and services and business statistics”, n.d.).
In OECD.Stat website, the OECD-Eurostat TEC database is able to present trade by size class, by top enterprises, by partner zones and countries, by number of partner countries, and by commodities. Data on exports or imports values and number of exporting or importing enterprises are available for 19 EU member states (Czech Republic, Denmark, Germany, Estonia, France, Italy, Cyprus, Latvia, Lithuania, Luxembourg, Hungary, Austria, Poland, Portugal, Romania, Slovenia, Slovakia, Finland and Sweden), plus Canada, Norway, Israel and the Unites States.

Figure 1: Example of TEC Statistics Interface by OECD

Source: Organization for Economic Co-operation and Development
TEC produce by Statistics Canada focuses on exporters in Canada by providing aggregate information on the number of exporting enterprises in Canada by industry and the value of exports attributable to these exporting enterprises. The information is available for reference periods after November 2000. The latest release was on December 2015 for reference year of 2014. The information contained in the export records was linked to the information available for the enterprise in the Statistics Canada Business Register, namely business characteristics such as industry engaged in exporting by creating profiles of exporting enterprises (“Trade by Enterprise Characteristics (TEC)”, 2015).

In Statistics Denmark, TEC statistics provides the information by imports and exports, number of enterprises, trade value, partner country, concentration, enterprise size, group of countries and main industries. The data is available from 2010 till 2014 and the data can be accessed from Statistics Denmark’s website.

Figure 2: Example of TEC Statistics Interface by Statistics Denmark

![Image of TEC Statistics Interface by Statistics Denmark]

Source: Statistics Denmark
Statistical Office of the Republic of Slovenia published its first release on Trade in Goods by Enterprise Characteristics, 2013 in May 2015. In Slovenia, ‘manufacturing’ enterprises generated nearly two-thirds (65.6%) of the total value of exports in 2003. The largest share of total imports (43.8%) was generated by ‘trade enterprises’ (wholesale and retail trade, repair of motor vehicles and motorcycles). In term of number of enterprises, the highest exporters were enterprises that involved with ‘other activities’ (33.2% of all exporters), followed by ‘trade enterprises’ (28.4% of all exporters) and ‘manufacturing’ (22.1% of all exporters). The structure of importers was the same as exporters; where the enterprises involved with ‘other activities’ (36.5% of all importers) was the highest, followed by ‘trade enterprises’ (24.7% of all importers) ("Trade in goods by enterprise characteristics, Slovenia, 2013", 2015).

Table 1: Number of exporters and importers according to section of activity and values of their trade in goods, Slovenia, 2013.

<table>
<thead>
<tr>
<th>Section of activity of enterprise</th>
<th>Exports</th>
<th>Imports</th>
<th>Exports</th>
<th>Imports</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>21,548,704</td>
<td>22,114,150</td>
<td>21,855</td>
<td>46,318</td>
</tr>
<tr>
<td>A Agriculture, forestry and fishing</td>
<td>97,868</td>
<td>53,108</td>
<td>282</td>
<td>617</td>
</tr>
<tr>
<td>B_D_E Mining and quarrying, electricity, gas, steam and air conditioning supply, water supply, sewerage, waste management and remediation activities</td>
<td>754,743</td>
<td>859,249</td>
<td>198</td>
<td>486</td>
</tr>
<tr>
<td>C Manufacturing</td>
<td>14,133,548</td>
<td>8,451,066</td>
<td>4,833</td>
<td>6,914</td>
</tr>
<tr>
<td>G Wholesale and retail trade, repair of motor vehicles and motorcycles</td>
<td>3,714,427</td>
<td>9,696,791</td>
<td>6,210</td>
<td>11,459</td>
</tr>
<tr>
<td>Other activities (other than A, B, C, D, E or G)</td>
<td>798,490</td>
<td>1,014,381</td>
<td>7,248</td>
<td>16,904</td>
</tr>
<tr>
<td>Unknown</td>
<td>2,049,627</td>
<td>2,039,556</td>
<td>3,084</td>
<td>9,938</td>
</tr>
</tbody>
</table>

Statistical Office of the Republic of Slovenia
2. Benefits of Trade by Enterprise Characteristics

TEC project is an engine to change the conventional trade statistics to disaggregate trade value according to the characteristics of trading enterprises. Particularly TEC can provide an additional and better statistical information for users with the aims of describing trade flows from the viewpoint of enterprises (Muryawan, 2014). For instance, TEC can generate the statistics with regard to the number of importers or exporters and the trade value by economic activities, by employment size, by small and medium enterprise (SME) status etc.

External trade statistics by enterprise characteristics has significant potential in compiling new statistics without causing an additional burden on businesses (Tilewska et al., 2013). TEC is an example of micro data linking on how statistical data collected in different statistical frameworks can be used effectively as new indicators.

The integration approach at enterprise level between international trade data and business register is very likely to generate great benefits ‘at virtually zero costs’ in terms of new statistical indicators and new firm level databases for the analysis of globalization (Giovannin, 2013). It is a cost-effective way for national statistical offices to undertake microdata linking and ensures harmonized, comparable results across countries (Tilewska et al., 2013).
According to Alajääskö (2015) from Eurostat, micro data linking is also an efficient tool to improve the quality of existing statistics linking data from different sources reveal inconsistencies and abnormalities. It is an additional way to ensure data quality and consistency between related data sets. In this case, the contradict information between SBR and trade database can be traced and necessary adjustment can be made.

TEC is relevant for policymakers as countries’ economic activity is becoming more interconnected and globalized. TEC data can address policy questions about what types of firms are involved in international trade; the principal economic activities or industrial sectors of traders; trading firms’ contribution to employment; and generally these firms’ role in global value chains and the associated structure of imported and exported products (Snyder & Jansen, 2015).

3. Potential Output of Trade by Enterprise Characteristics

TEC is not only allow the integration of SBR with trade database, but also potentially leading to significant gains in efficiency and quality of data collection and processing. TEC can provide new information that would not exist in stand-alone statistical domains. The integrated datasets can indicate which enterprises are engaged in international trade as part of global value chains and measure the importance of those firms in the overall economy.
Snyder & Jansen, 2015 describe TEC data can be compiled based on trade value, trade quantity, and by number of enterprises for the following variables (and combinations thereof):

i. Trade by economic activity (i.e. International Standard Industrial Classification (ISIC))

ii. Trade by enterprise size as measured by number of employees

iii. Trade by enterprise size as measured by sales or turnover

iv. Trade by foreign ownership (e.g., domestically-controlled; foreign enterprise; presence of affiliates abroad)

v. Trade by partner countries (by zone of partner countries and/or by number of partner countries)

vi. Trade by (sub-national) geographic location

vii. Trade by sectors’ traded commodities

viii. Trade by export intensity (i.e., exports as percent of total sales)

Eurostat has developed the TEC with five indicators which are available for imports and exports and for intra-EU and extra-EU trade. All indicators use enterprise as the statistical unit and are expressed in terms of number of enterprises and trade value. The indicators are:

i. Trade by economic activity and enterprise size class: to allow analysing the impact of external trade on employment and estimating the importance of small and medium-sized enterprises.

ii. Concentration of trade by economic activity: to show the share of the total trade accounted for by the top 5, 10, 20, etc. enterprises.

iii. Trade by partner country and economic activity: to show how many enterprises were trading with certain partner countries or country zones,
and the value they accounted for. This enables to identify most typical exports or imports markets.

iv. Trade by number of partner countries and economic activity: to show how geographically diversified the exports markets are. For imports, it shows the number of countries from which goods are imported.

v. Trade by commodity and economic activity: to show which sectors were involved in trading of each product group.
B. Malaysia Experience on MSBR & Trade Database Linkage

1. Malaysia Statistical Business Register

Like any other NSOs, Malaysia Statistical Business Register (MSBR) has been developed according to the requirement of statistical purposes and serves as central register in DOSM. It is used as a tool for the preparation and coordination of the surveys. Currently, MSBR contains about 1.2 million active establishments and 2.0 million inactive establishments; it has been in operation since end of 1980s. Earlier, the register was known as Central Register System (CRS), however in 2010, the expansion of the CRS; it has been known as MSBR. There are two statistical unit widely used in Malaysia compilation statistics i.e. establishment and enterprise. The establishments operating in Malaysia are categorised based on industry classifications (MSIC 2008 which is comparable to ISIC 4 of the United Nations). The maintenance of MSBR are based on multiple sources; censuses, surveys and administrative database. The essential identification information of establishment in MSBR are:

i. Unique ID number that serves as the primary identifiers (maintained by DOSM);

ii. Business registration number;

iii. Contact information such as registered name, business name, postal and location address and others;

iv. Activity or industry classifications;

v. Legal status;

vi. Basic financial information such as revenue, sales, salary and others;

vii. Number of employees;

viii. Operational status;
ix. SME status; etc.

2. **Trade Database**

Malaysia’s external trade statistics is to provide information on Malaysia’s trade performance compared with other countries in the world in terms of volume and value of merchandise goods. In order to ensure the comprehensiveness of coverage, several sources of information are collected to produce Malaysia’s trade statistics. On average, 98 per cent of the Malaysia import and export declarations are sourced from Royal Malaysian Customs Department and the rest are from Free Zone Authorities; Malaysian Postal Services; Malayan Railway Limited; Malaysia Airlines System Berhad; petroleum companies; operators of duty free shops; statements and export declarations on bunkers and stores supplied to ships; and statements on the sale and purchase of electricity issued by power producer company. The available information in Customs declaration are:

i. Business registration number (importers or exporters);

ii. Forwarding agent code;

iii. Importers or Exporter name & address;

iv. Forwarding agent name & address;

v. Port of loading or Place of discharged;

vi. Mode of transport;

vii. Import Export value;

viii. Import or Export quantity;

ix. Product code (HS or AHTN);

x. Country of origin or country of final destination; etc.
3. **Integrated MSBR with Trade Database**

MSBR plays an important role in bringing trade statistics closer to the business statistics. MSBR provides a tool to link detailed external trade micro data with the statistical units used in business statistics. With enterprise level information available in MSBR; which contain information on size and turnover, activity (industry), legal status and SME status, it can be matched with Customs data that provide volume, value and product traded at 9 or 10 digit levels together with the identification of enterprise entities involved in international transaction.

In Malaysia, the identification of enterprise entities in both MSBR and Customs declaration is business registration number which is maintained by Companies Commission of Malaysia (CCM). CCM is a statutory body to serve as an agency to incorporate companies and register businesses as well as to regulate companies and businesses in Malaysia. This business registration number has been used as unique identifier for Malaysia’s TEC project.

The TEC project is carried out in three separate phases beginning October 2015 and expected to be completed at the end of 2016. For a start, August 2015 export data (one month data) is used as pilot data linkage. The project coverage is about 97.8 per cent of the total export declarations records and 85.4 per cent of total exports value in August 2015. This is due to the unavailability of business registration number information in other than Customs declarations. The very reliable matching approach for the project is business registration number. Before matching process, the business registration number in trade database needs to be cleansed and standardized. The quality of statistics is very much depends on the matching rates between two source data sets.
Phase 1 (Linkage by Unique Identifier):

i. The business registration number from trade database will be cleansed to proceed with the matching process. The first and last character of business registration number will be removed.

Example:

Original business registration number in Customs declaration = SLA9999999W

Right business registration in MSBR = LA9999999

ii. There are two attempts in doing cleansing and standardizing process:

<table>
<thead>
<tr>
<th>Attempt</th>
<th>Enterprises</th>
<th>Cleaned Records</th>
</tr>
</thead>
<tbody>
<tr>
<td>First</td>
<td>Companies</td>
<td>66.4%</td>
</tr>
<tr>
<td>Second</td>
<td>Companies &amp; Businesses</td>
<td>78.7%</td>
</tr>
</tbody>
</table>

The rate is highly depending on what have been filled-in by traders or forwarding agents on the business registration numbers.

iii. The cleaned trade business registration number will be matched with MSBR business registration number.

- The matching rate of the total exports’ records (August 2015) is as follows:

<table>
<thead>
<tr>
<th>Attempt</th>
<th>Enterprises</th>
<th>Exports</th>
<th>Number of records</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Value</td>
<td>Percentage</td>
</tr>
<tr>
<td>First</td>
<td>Companies</td>
<td>58.5%</td>
<td>61.9%</td>
</tr>
<tr>
<td>Second</td>
<td>Companies &amp; Businesses</td>
<td>65.4%</td>
<td>73.8%</td>
</tr>
</tbody>
</table>
- The matching rate of the cleaned exports’ records (August 2015) is as follows:

<table>
<thead>
<tr>
<th>Attempt</th>
<th>Enterprises</th>
<th>Exports</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Value</td>
<td>Number of records</td>
</tr>
<tr>
<td>First</td>
<td>Companies</td>
<td>95.2%</td>
<td>93.2%</td>
</tr>
<tr>
<td>Second</td>
<td>Companies &amp; Businesses</td>
<td>91.9%</td>
<td>93.8%</td>
</tr>
</tbody>
</table>

iv. The overall summary of TEC statistics (exports) for companies and businesses in August 2015:

<table>
<thead>
<tr>
<th>Sector</th>
<th>No. Of Exporters</th>
<th>Export Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manufacturing</td>
<td>44.4%</td>
<td>68.8%</td>
</tr>
<tr>
<td>Distributive Trade</td>
<td>33.8%</td>
<td>8.6%</td>
</tr>
<tr>
<td>Services</td>
<td>13.6%</td>
<td>13.1%</td>
</tr>
<tr>
<td>Multiple Activities</td>
<td>3.1%</td>
<td>7.4%</td>
</tr>
<tr>
<td>Construction</td>
<td>2.1%</td>
<td>0.9%</td>
</tr>
<tr>
<td>Agriculture</td>
<td>1.9%</td>
<td>0.4%</td>
</tr>
<tr>
<td>Mining &amp; Quarrying</td>
<td>0.6%</td>
<td>0.5%</td>
</tr>
<tr>
<td>Water Supply; Sewerage, Waste Management &amp; Remediation Activities</td>
<td>0.4%</td>
<td>0.2%</td>
</tr>
<tr>
<td>Electricity, Gas, Steam &amp; Air Conditioning Supply</td>
<td>0.1%</td>
<td>0.1%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>100.0%</strong></td>
<td><strong>100.0%</strong></td>
</tr>
</tbody>
</table>

**Phase 2 (String Matching Algorithm):**

Since the matching rate by unique identifier approach (of total exports’ records) is not achieved DOSM acceptance level or less than 80 per cent, the matching technique will be further improved by string matching algorithm approach. The approach is currently being studied to identify the best string matching methods and tools to generate more accurate statistics on TEC.
Phase 3 (Expand the Coverage and Automate the Linking Process):

This feasibility study will be extended to the full scale project once the suitable approach and system is determined. The one month exports data will widen to the one year exports imports data and this business process plans to be automated.
4. Challenge of Success

In progressing to integrate these two databases, there are challenges that required attention especially in the aspect of data quality, task complexity and confidentiality.

Shortcoming in data quality is a condition where there is an obstacle to the user to use that data. The root cause(s) should be determined and the options for remediation should be decided. The incomplete coverage of the MSBR or the poor quality of business registration number declared by exporters or importers in Customs declaration has contributed to the unmatched data. The unmatched data will lead to inadequacy of TEC statistics.

In another aspect, the invalidated business registration number in Customs declaration has allowed the forwarding agents use whatever code they like to represent exporters or importers in which ultimately can defeat the purpose of producing TEC. The efforts of reducing unmatched data will be fruitless as long as the quality of data is not rectified at source. Since the databases of the project come from different statistical domains, the data inconsistency also can become another challenge faced by DOSM. All of these need extra efforts from the team.
The fundamental problem in this TEC project is the poor quality of business registration number in trade database which cause complexity in handling the task. The linking process should be further enhanced by using string matching algorithms where the importers' or exporters' name is matched with enterprises' name in MSBR. Having knowledge in algorithms, algorithmic complexity and intricacies of specific algorithms is useful to accomplish the TEC. Thus, person in charged should be equipped with necessary knowledge, right skills and tools to manage the task complexity.

Another frequently asked question is how to determine the MSIC classification to a particular importer or exporter. Trade database would consider enterprise as statistical unit, if there are multiple exporting plants that all report to one enterprise (use same business register in Customs declaration), the location and MSIC will be multiple based on individual plant registered in MSBR. Therefore, identifying the primary activity to represent enterprise will be tricky and complicated.

Confidentiality issue is one of the challenges where DOSM should not release any information which could identify any person, business, or organization, unless consent has been given by the respondent or as permitted by the Statistics Act. Confidentiality is not a major issue on the aggregated level data, but when it comes to more detailed statistics, the possibility to identify an enterprise or a trader getting higher.
The TEC project is based on linking trade micro data with MSBR information thus the linked micro data files should not be shared with third parties. Identifying ‘risky’ information is relatively easy but to ensuring that they are sufficiently concealed is the challenge. The difficulty of applying optimal confidentiality increases with level of details.

In addition, data gaps due to confidentiality rules in producing TEC may hinder international comparability and may discourage users; even though new information is available but they cannot access to the details. For instance, in some cases TEC data cannot be presented at 5 digit level of MSIC due to few players involved in that particular industry. It may be presented at fairly aggregated level (e.g. 2 digit level of MSIC).

C. Future Perspectives

The TEC’s potential can be extended further once the trade database is linked to business demography information in MSBR. DOSM is in the process of getting the data on enterprises’ ‘incorporated date’ from CCM. This is ensure that all establishments registered in MSBR complete with ‘incorporated date’ information. At present, only 22 percent of establishments equipped with the information in MSBR. The linkage will allow DOSM to produce statistics on export performance of young enterprises which is known as born-global\(^1\) enterprises. The information can be an input to the policy maker in identifying the characteristics of enterprises that operate in international markets from the earliest days of their establishment; for instances the type of industries they involved.

\(^1\)Type of company that from the beginning of its activities becoming global and globalizes rapidly.
the employment size and the location of the enterprises etc. Therefore, any incentives or policies to promote export activities will be reached to the right target.

On top of that, business demography information can help government to detect whether the enterprises are new or persistent exporters. If they are persistent exporters, what are the products that support them to survive longer in the global market and in what industry they are classified to. Generally, bringing trade statistics closer to the frame of statistics could offer many attractive features.

IV. Conclusion

The TEC project has been identified as a pioneer project in Asia and the success of the project is very much expected to inspire other countries to come out with these new statistics without having additional surveys and increasing respondent burden. The initiative is also in line with the international recommendations i.e. Business Registers Recommendation Manual 2010 (Eurostat) and International Merchandise Trade Statistics (IMTS) 2010 (UNSD). The value added to existing international trade statistics project would provide new statistics or new indicators for the purpose of formulating or revising the national policies. In short, the investment of data linking is relatively low and the returns in the form of new economic insights are high. Even we need to face bigger challenge but the benefit generated from the project is even bigger. Hence the TEC will be DOSM new development areas for external trade statistics in order to analyze the effects of international trade to production, employment and enterprises' performances.
V. References


