EESW15, the Fourth European Establishment Statistics Workshop

Linking Business Registers with Trade Statistics
Nancy Snyder* and Ronald Jansen
United Nations Statistics Division, New York, United States

1. Introduction

A major development in official statistics in recent years has been the establishment and use of national statistical business registers (SBRs) that not only allow for integration of business surveys, but also for linking of administrative and other data sources, potentially leading to significant gains in efficiency and quality of data collection and processing. The *Guidelines on Statistical Business Registers*, as published by the UN Economic Commission for Europe (UNECE),[1] acknowledge that while traditionally, the function of the SBR has been to provide sample frames from which economic surveys are drawn, they can also crucial in the integration and use of data from administrative and other sources and the business demographics contained therein can be used as a source of economic statistics in their own right. In fact, the integration of the SBR with economic statistics from different statistical domains can provide new information that would not otherwise exist. For example, the SBR can be linked with trade statistics at the micro-level, and an increasing number of countries have undertaken projects to do so. Such integrated datasets can, for instance, indicate which firms (characterized by industry, size class, foreign ownership, and geographic location) are engaged in international trade as part of global value chains and measure the importance of those firms in the overall economy.

In order to assess the status globally of linking trade statistics to the SBR, and at the same time to take stock of the availability and use of SBRs worldwide, UNSD administered the global *Survey on national practices in linking trade statistics and business registers* ("the survey") in the summer of 2015 to all NSOs, requesting information on their national practices in this area. This paper presents the initial findings from this survey, based on a relatively small sample of responses received to date;[2] updated survey results will be presented at the EESW15 workshop held in Poznan in September 2015 and the final results of the survey will be reported in late 2015. One finding of the survey is that a well-maintained statistical SBR is essential for such a linking exercise, as well as strong institutional arrangements among the different agencies that may be involved in maintaining the SBR and in the compilation of business, trade and other economic statistics.

2. The role of the SBR in an Integrated Economic Statistics Programme

The UN Statistical Commission ("the Commission) recommends national statistical offices (NSOs) to adopt an integrated approach to economic statistics in order to ensure the efficiency of the statistical process and increase the consistency and coherence of economic statistics.[3] The Commission has identified several benefits of integrated economic statistics programmes, including consistency between short-term, annual and benchmark statistics; consistency in measuring economic activity across different sectors, industries and regions; consistency across countries for key economic indicators (such as real GDP, inflation, and international trade); greater accuracy in the economic data through the reconciliation of discrepancies across data from different sources; and reduction in the reporting burden for business respondents and increased

---

[1] UNECE, *Guidelines for Statistical Business Registers*, Geneva, June 2015, p. 60.
file:///C:/Users/Nancy.Snyder/Downloads/Guidelines%20on%20Business%20Registers,%202015.06.04.pdf.
[2] As of the writing of this paper on 31 July 2015.
[3] Based on the report of the Secretary-General on integrated economic statistics (E/CN.3/2006/5).

efficiency in the production of data.[4] In order to create an integrated economic statistics programme, the Commission outlines the importance of the following key features:[5]

    i.   The use of harmonized terminology, definitions, concepts, standards and classifications;
    ii. The central role of SBRs in providing a central sampling frame for all business surveys;
    iii. The standardization of surveys, including survey design, sample frame, and questionnaire design;
    iv. Matching of the concepts of administrative source data with statistical records, allowing for the utilization of using administrative records to promote more efficient data collection and reduction of the burden on respondents;
    v.   Editing, linkage and integration of data across various statistical domains;
    vi. Integration of dissemination and communication in order to provide user-friendly presentations of data and explanations of concepts and to ensure consistent formats across dissemination platforms.

Thus, SBRs have a central role in integrated economic statistics. They provide a common universe for sampling and a uniform classification of firms, which promote consistency in collected data across surveys. Moreover, they allow for the integration of data from administrative sources and surveys and serve as an important tool for stratification, imputation and editing of basic source data. As a result, the timeliness in producing statistics can be improved and respondent burden reduced.

3. Linking the SBR to trade statistics

The integration of data from different sources can provide new information for many analytical purposes that would not otherwise exist. Such data can be obtained from linking the SBR to economic statistics. Importantly, a well-maintained SBR is essential for such a linking exercise, as well as strong institutional arrangements among the different agencies that may be involved in maintaining the SBR and in the compilation of business, trade and other economic statistics. One effective application is the linking of the SBR to trade statistics in order to compile more detailed data on trade by enterprise characteristics (TEC).

The survey on national practices in linking trade statistics to the SBR asked NSOs if they are currently linking statistics of international trade in merchandise (IMTS) and services (SITS) and foreign direct investment (FDI) to the SBR. Of responses received,[6] 40 percent of respondents report that they currently link IMTS to the SBR; 27 percent report currently linking SITS; and 29 percent report linking FDI to the SBR. The results of the survey also highlight several key issues that are shown to be instrumental in successfully linking the SBR to trade statistics.

Issues to consider regarding the SBR

Effective maintenance and updating of the SBR is a crucial component in facilitating the linking of trade statistics to the SBR. UNSD encourages statistical compilers to ensure shared access to the SBR for all agencies involved in the compilation of economic statistics; to establish a shared vision and commitment to integrating trade and business statistics among the agencies involved; and formation of clear policies on access to and use of confidential information. In the vast majority (87 percent) of responding NSOs, the maintenance and use of the SBR is centralized to one agency – typically the NSO itself. Having a centralized system for the maintenance of the SBR facilitates consistent updating and recording of information. Half

---

[4] *Guidelines on Integrated Economic Statistics,* Series F No. 108, New York, NY, 2013, p. 2. http://unstats.un.org/unsd/publication/seriesf/Seriesf_108e.pdf.
[5] *Ibid., p. 9.*
[6] As of 31 July 2015.

of the respondents report updating the SBR continuously, while most others report updating it annually. Sixty-nine percent of respondents report using tax records and enterprise surveys to update the SBR. Half of the respondents report also using the economic census.

Implementing a unique identifying number for enterprises that is common throughout the statistical system is another key step needed to conduct the linking exercise. While nearly all survey respondents indicated that they use a unique identifying number for each enterprise, if such a number is not available, enterprises can and should be identified with unique legal names and addresses that can be later linked to identifying numbers used in other parts of the statistical system or in administrative data, such as tax identification numbers.

The matching of statistical units from the SBR to trade data sources is also of critical importance. The *Guidelines on Statistical Business Registers* recommends that the SBR should include the legal unit, the enterprise, the local unit,[7] and the enterprise group.[8] A majority of survey respondents (71 percent) report that the enterprise[9] is the most common statistical unit maintained in the SBR, while some report also using establishments.[10] The enterprises in the SBR must then be linked to the businesses recorded in the trade data source, which is most often customs declarations. Based on the survey results, trade data can just as commonly be recorded by enterprise or by establishment. If the trade data are recorded by establishment, the data must be aggregated to the enterprise level (or enterprise group, which is a set of enterprises controlled by a group head), based on some other identifying feature, such as legal name or identification numbers used in administrative sources like tax ID numbers. UNSD notes that while unmatched enterprises (that are typically marked to an "unknown" category) may be a concern, evidence from European countries shows that such unmatched enterprises are mostly small and medium sized enterprises (SMEs)[11] that represent a relatively small portion of the aggregate trade value.[12]

It is also desirable to include other variables in the SBR, including legal name; address; economic activity; number of employees; turnover; date of entry; and active/non-active status. Nearly all survey respondents indicated that their SBRs include these variables. Identification of a principal economic activity (that which contributes most to the value added of the unit, or to the activity the value added of which exceeds that of any other activity undertaken by the unit); turnover and number of employees are relevant for compiling detailed statistics of trade by enterprise characteristics, which are highly policy-relevant (and will be discussed further below). These variables are also important pieces of information when using the SBR for establishing the sample for a given survey.

Much less commonly-maintained variables in the SBR are identification of trader/non-trader (with 24 percent of respondents reporting this variable is maintained) and percentage of foreign ownership (cited by only 18 percent of respondents). Such information would be relevant to analyzing the behavior of multinational corporations and the relationship between foreign direct investment and a country's international trade. An alternative to maintaining such variables on the SBR is to conduct an enterprise FDI

---

[7] A "local unit" is defined as an enterprise or a part of an enterprise (for example, a workshop, factory, warehouse, office, mine or depot) which engages in productive activity at or from one location.

[8] UNECE, *Guidelines for Statistical Business Registers*, Geneva, June 2015, p. 60. file:///C:/Users/Nancy.Snyder/Downloads/Guidelines%20on%20Business%20Registers,%202015.06.04.pdf.

[9] An "enterprise" is defined as an organizational unit producing goods or services which can take decisions, especially for the allocation of its resources.

[10] An "establishment" is defined as an enterprise or part of an enterprise which is situated in a single location and with a principal productive activity.

[11] SMEs are defined as having less than 250 employees.

[12] According to Eurostat, unmatched enterprises represent about 8 to 9 percent of trade values. http://ec.europa.eu/eurostat/statistics-explained/index.php/International_trade_by_enterprise_characteristics#Data_sources_and_availability.

survey; such a survey could also be linked to the SBR in order to validate responses or to update information in the SBR. Twenty-nine percent of survey respondents report linking FDI statistics or surveys to the SBR.

Issues to consider regarding trade statistics

In order to link the SBR to trade statistics, the data source for trade statistics must be well understood. Customs declarations are reportedly used in all responding countries to compile IMTS. Such data may also be complemented by enterprise surveys (cited by 29 percent of respondents), tax records (cited by 18 percent), and to a lesser extent, other administrative data (such as from ministries of energy or agriculture). As noted above, based on the survey results, merchandise trade data can just as commonly be recorded by enterprise or by establishment, which then must be linked or aggregated to the enterprise unit (or other unit) as maintained in the SBR. Moreover, merchandise trade data are recorded by business for each transaction (export or import), with detailed product description and classification, value, exporting and importing country, and date. All of the survey respondents that currently link IMTS to the SBR report doing so at the level of the exporter/importer (rather than at the transaction level).

Another issue to consider is differences in the various domains of trade statistics. For instance, the majority (80 percent) of respondents cite the NSO as the responsible agency for compiling and disseminating IMTS, whereas for SITS and FDI, about half of the respondents cited the NSO as the agency responsible for maintaining a list of relevant enterprises, with the other half citing the central bank. In the latter case, institutional arrangements among the different agencies involved in maintaining enterprise lists, compiling the data and disseminating the official statistics becomes critically important in order to be able to link such data to the SBR. Moreover, data sources for SITS and FDI are often enterprise surveys, which, unlike customs declarations or other administrative data, may be less consistent in their recording of unique identifying numbers for enterprises across sources. Integrated survey design and coordination across statistical domains is therefore needed.

4. Compilation of trade by enterprise characteristics (TEC) data

Linking trade statistics to the SBR enables the compilation of a new statistical domain, called "trade by enterprise characteristics" (TEC). Several countries[13] are already compiling TEC indicators, with international organizations Eurostat[14] and the Organization for Economic Cooperation and Development (OECD) involved in the collection and dissemination.[15] TEC data indicate which types of firms – for example, by industry, size class, foreign ownership, or geographical region – are engaged in international trade. This information is becoming increasingly 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. TEC data can be compiled based on trade value, trade quantity, and by number of enterprises for the following variables (and combinations thereof):
- Trade by economic activity (i.e., International Standard Industrial Classification (ISIC))
- Trade by enterprise size as measured by number of employees
- Trade by enterprise size as measured by sales/turnover
- Trade by (sub-national) geographic location
- Trade by foreign ownership (e.g., domestically-controlled; foreign enterprise; presence of affiliates abroad)

---

[13] A non-exhaustive list of these countries includes the EU member states (except Ireland), Albania, Canada, Norway, Israel, Turkey and the United States; Australia compiles TEC for exporters only.

[14] See http://ec.europa.eu/eurostat/statistics-explained/index.php/International_trade_by_enterprise_characteristics

[15] See http://www.oecd.org/std/its/trade-by-enterprise-characteristics.htm.

- Trade by partner countries (by zone of partner countries and/or by number of partner countries)
- Trade by sectors' traded commodities
- Trade by export intensity (i.e., exports as percent of total sales)

Similar variables can be compiled for trade in services statistics (i.e., Services Trade by Enterprise Characteristics (STEC)) and FDI statistics (FDIEC), which are compiled by combining the information from enterprises surveys on trade in services and FDI with the SBR.

According to the UNSD survey results, 41 percent of respondents disseminate TEC indicators, with the most commonly disseminated indicator being value and number of enterprises by firm's number of employees, followed closely by economic activity and by sub-national geographic region. The fact that data on merchandise trade by firm's number of employees is most commonly disseminated indicates that such data are likely addressing the pressing policy question regarding the effect of international trade on domestic employment. Dissemination of STEC and FDIEC indicators is far less common, with only 12 and 24 percent of respondents, respectively, reporting that they currently disseminate such statistics. Among these respondents, most report disseminating indicators on value by economic activity.

While not all TEC indicators that are compiled are publicly disseminated (and/or are only available upon special request), due to the detailed nature of such data, confidentiality is of utmost concern in order to maintain the trust and confidence of data reporters and to ensure the integrity of the NSO. When asked about their confidentiality policy regarding compilation/dissemination of TEC indicators, most respondents reported that they require at least three to five enterprises per cell as the threshold.

5. Challenges and Ways Forward

When asked about challenges facing the NSO when attempting to link trade statistics to the SBR, all respondents cited matching enterprises or establishments. This step is by far the most time-consuming and often requires some kind of technical know-how to be able to collate, compare and match data from different sources and database formats. Another challenge, albeit only reported by 24 percent of respondents, is the existence of many wholesalers or distributors, making it more difficult to link the firm with its trade activity. Other challenges cited include methodological issues; technical deficiencies; human resource limitations; and institutional barriers.

UNSD encourage NSOs to continue exploring the benefits of integrating data from different sources, especially linking the SBR to trade statistics, in order to compile new statistics that would otherwise not exist. TEC, STEC and FDIEC indicators have the potential to address a wide range of policy concerns in our increasingly globalized and complex world. Although the process of creating a system to link and match statistical units across different data sources may be time-consuming and resource-intensive in the initial stages, UNSD suggests that NSOs consider the issues raised in this paper in order to facilitate the process and reap the benefits of an integrated economic statistics programme. More specifically, strong and effective institutional arrangements; integration of data sources; alignment of key concepts and data analysis tools across statistical domains; and integration of the different steps in data production (from survey design, to data collection, to processing, to dissemination) will not only help modernize and increase the efficiency of statistical systems, but also lead to greater accuracy, quality and relevance of official statistics.