31st Voorburg Group Meeting on Services Statistics
Zagreb, Croatia
19-23 September 2016

Guidance Paper: Industry vs. Product

Andrew Baer
Assistant Division Chief, Services Sectors
Economy-Wide Statistics Division
US Census Bureau
andrew.l.baer@census.gov

1. Introduction

This paper describes the practical and methodological considerations National Statistics Offices (NSOs) face when determining if Services Producer Price Indexes (SPPIs) and turnover statistics are to be collected and tabulated by industry or by product. It takes as its starting point the issue papers presented by the US, Germany, Finland, and Hungary at the 2015 meeting of the Voorburg Group.

Consistent with the Eurostat-OECD definitions[1], industry-based surveys are defined here as data collection covering all economic production from statistical units within the same industry classification group. In contrast, product-based surveys collect transaction data only for a particular type of good or service, from those statistical units known to produce it.

2. Industry/Activity Classifications

The International Standard Industrial Classification of All Economic Activities (ISIC), Rev. 4 manual describes the system as follows:

> "ISIC is built on a production-oriented or supply-based conceptual framework that groups producing units into detailed industries based on similarities in the economic activity, taking into account the inputs, the process and technology of production, the characteristics of the outputs and the use to which outputs are applied."[2]

With some subtle variations, the major industrial classification systems used around the world, including the NACE and the North American Industry Classification System (NAICS), follow similar principals[3].

The suitability of these classifications for producing reliable and accurate industrial statistics depends on NSO's abilities to identify discrete and stable operational units engaged in similar activities. Matt Berger described in 2007 "a statistician's utopia" as a world with "homogenous statistical units, each of which produced one (and only one) economic activity."[4] Unfortunately, the real world is instead populated with diverse enterprises producing many products with a variety of production processes. In the

---

[1] OECD-Eurostat Methodological Guide for Developing SPPIs, 3.3.2.3

[2] ISIC, Rev.4, 2.39

[3] An example of a variation is that NAICS prioritizes production process as a determinant of classification, while it is only one of multiple factors that determine primary activity in the ISIC and NACE.

[4] Matt Berger, Australian Bureau of Statistics, "Addressing cross-cutting issues arising in the development of SPPIs for use in the measurement of Service Sector GDP". Prepared for the 2007 meeting of the Voorburg Group.

United States, the proportion of industry receipts derived from secondary production increased by nearly one third from 2002 to 2007[5].

While the use of establishments as statistical units minimizes secondary production, it does not eliminate it.   Secondary activities must be accounted for to ensure comprehensive coverage of economic activity in industry statistics. This is a significant challenge for NSOs, since it is often difficult for responding businesses to provide accurate transaction information for secondary goods and services. In practice, SPPI data collectors sometimes truncate secondary production to minimize burden and increase the likelihood of getting cooperation for the primary services.

3. Product Classifications

The Central Product Classification (CPC) is a comprehensive classification of all goods and services transacted or placed in inventory. Products are classified "based on the physical properties and the intrinsic nature of the products as well as on the principle of industrial origin.[6]" The concept of considering industry when determining product is maintained even more strongly in the European CPA product classification. There each product is assigned to a single industry where it is "characteristically produce(d).[7]" The North American Product Classification System (NAPCS), in contrast, does not consider industrial origin. It is a purely demand-based system, "in which products are grouped according to how they are principally used…[8]"

4. Evaluation of Suitability for SPPIs and Nominal Output Statistics

In her 2015 paper, Dorothee Blang evaluated whether industry or product statistics are most appropriate for each of the primary uses of SPPIs[9]. Except where noted, that analysis is extended here to also include turnover measures.

- Short-term economic indicators

    Economic trends are driven by interactions of supply and demand in competitive markets. Market competition routinely occurs between firms that use different production processes and may be classified in different industries. For example, NAICS provides distinct industries for wired, wireless and satellite telecommunications providers, as these firms use varying methods to produce telecommunications services. Since prices and turnover for telecommunications

---

[5] A rough estimation based on Economic Census figures allocated as primary or secondary by BLS staff shows 6.24% of all industry receipts from secondary production in 2002 and 8.28% in 2007.

[6] CPC Ver 2.1, 2.B.27

[7] CPA 2008, 3.1

[8] Overview of NAPCS Objectives, Guidance, and Implementation Strategy and Goals: A United States Perspective, April 2003

[9] http://www.voorburggroup.org/Documents/2015%20Sydney/1013.pdf

services in the US are determined by competitive market conditions *between* each of these distinct types of producers, product-based cross-industry data that include all of these transactions will provide more clarity about business competition and market pricing. For purposes of analyzing short-term economic trends, **product data** provide a clearer picture of market behavior. They are also more straightforward for data users to interpret since industry data include a disparate group of products, some primary and some secondary.

- Deflation of national accounts

  The OECD-Eurostat SPPI Guide notes that **product data** are best suited to facilitate the calculation of supply and use tables, which are recommended as the building-block accounting framework by the System of National Accounts[10]. Supply and use tables record the industries that produce particular products and the users who purchase them either as final demand or as intermediate inputs. Real product output data are needed to compute these tables.

- Escalation of business contracts (SPPIs)

  Parties that enter into long-term contracts often use SPPIs to determine fair market value in later contract periods. Since contracts specify the nature of the good or service transacted, rather than the industry classification of the party that produces it, **product data** is most appropriate for these users.

- Analysis of price transmission (SPPIs)

  Economic analysts may investigate lagged relationships between SPPIs and CPIs to monitor how inflationary pressures are passed from producers to consumers. Since CPIs are calculated on a product basis, **product data** allows for the easiest comparison between the two measures.

- Measuring productivity

  By organizing data based on *how* services are produced, **industry data** are optimally designed for measuring productivity[11]. Wherever-made product data cannot be easily linked with specific labor and capital inputs. These inputs are most easily observed at the establishment level, where industry data are ideally collected.

---

[10] OECD-Eurostat Methodological Guide for Developing SPPIs, 1.1.3
[11] OECD Guide for Measuring Productivity, 3.4.61

Product measures are most appropriate for a clear majority of the uses of services price and output data. As a result, it is important for NSOs to manage the many practical challenges of producing wherever-made product data.

5. Strategies to Address Challenges with Product Data

Challenge: Insufficient sampling frames for product surveys

A significant impediment to creating product-based surveys is lack of adequate sampling frames. Most business economic survey samples are selected from business registers or national tax records. In the US, these data sources do not provide information on business activity by product. This is because most business respondents do not organize their records to allow for enumeration of employment, turnover, or wages for production of a specific good or service. Without these data, probability proportionate to size sampling for product-based surveys can only be achieved through alternative data sources and/or extensive imputation of producer size measures. These alternative methods typically do not provide the same comprehensive and accurate coverage as tax or business register sources.

Strategies

The most promising solution to this challenge in the US is to include product data collected from an existing industry-based survey directly on the business register. The US Economic Census is a comprehensive industry-based data collection of business activity that occurs every five years. If the product information collected or imputed in the tabulation of Economic Census were to be applied onto the business register, suitable product-based sampling frames would be available for future use. This challenge is more significant, however, in countries where comparable comprehensive data collections do not exist.

Value-added tax (VAT) collection records are a potentially suitable source for sampling frames in countries where they exist. Of course this would be available only for those product areas where VAT is collected, making this a partial solution that would need to be combined with a different approach to achieve comprehensive coverage.

Where the sampling frame challenge is not solvable with these two approaches, it may be necessary to produce approximate product-based data from industry-based data collections[12].

---

[12] OECD-Eurostat Methodological Guide for Developing SPPIs, 3.3.2.3

Challenge: Coordinating product data collection schedules (SPPIs)

In the US, SPPI industry-based samples are selected every five to eight years on a rotating basis. For example, in 2017 a new industry-based sample may be selected for NAICS industry 721110, Hotels and Motels. Bureau of Labor Statistics (BLS) data collectors will then visit the sampled hotels and conduct an hour-long meeting in which they perform a statistical selection of the transactions to be priced, giving all goods and services offered by the hotel a chance to be selected. This would include not only lodging services, but also food and beverage services provided by these establishments. Data collectors would not visit the hotel again until 2022-2025, and then only if the property was selected in the subsequent SPPI industry-based sample.

If the US instead chose to select product-based SPPI samples, data would be collected for all lodging service providers in 2017. This would include not only firms classified in NAICS 721110, but also those classified in NAICS 721120 (Casino Hotels), NAICS 721191 (Bed and Breakfast Inns), NAICS 721211 (Recreational Vehicle Parks and Campgrounds), and other related industries. In 2018, a product-based sample of food and beverage service providers may then be selected. With this product-based sampling scheme, establishments that provide both lodging and food and beverage service may be asked to schedule meetings with a data collector in subsequent years. This creates an additional reporting burden and may harm response rates for voluntary US SPPI surveys. This is particularly true for the large conglomerate firms that offer a broad variety of products and services and consequently would be visited repeatedly for multiple product surveys.

Since product-based sampling frames are usually of a lower quality, misclassification may also present challenges. With an industry-based sample, if a firm is found to be classified in a different but related industry, data collection may still occur. However, if a firm selected for a product-based sample is found not to transact the specified product, data collection is not possible. This creates an inefficiency of resource usage for product-based samples.

Strategies

If a rotating schedule is employed for data collection, attempts should be made to plan for the collection of complementary products, such as lodging and food services, at the same time to minimize responding burden. This will also reduce the risk of encountering responding units that do not produce any of the services eligible to be collected at a given time.

A transition to product-based data collection may also create an opportunity for NSOs to develop collection strategies that request key respondents to update smaller numbers of items at more frequent intervals. Rather than attempting a comprehensive once every five years in-person interview to select items, a subset of products could be updated each year through shorter telephone or video conference interactions. This has the

advantage of frequent respondent engagement, potentially encouraging the respondents to become active partners in ensuring their requested items are up-to-date with their firm's current activities. This could also support the development of a customer-centric account manager model, where an agency employee is assigned as a single point of contact for an important respondent. This model has been successfully employed by leading NSOs to improve the quantity and quality of data received from larger, more complex enterprises.

Challenge: Collecting product data at an establishment level (turnover surveys)

In the services sector, the concept of allocating product-specific turnover to a specific, local, business establishment has become exceedingly challenging. This is particularly true for the information technology, telecommunications, and professional services industries. Take for example the growing area of cloud computing, where firms primarily offer subscription Internet access to software and computing capacity. A typical industry player would have a headquarters at a single location, sales offices in the major cities closest to the clients, customer support centers anywhere globally, and data storage facilities typically in low-population areas with reliable access to the electric grid. To make it more challenging, an individual customer's connection to the firm's network may at different times route to *different* data storage facilities, based on where capacity is most readily available. So to which establishment should the turnover generated from this customer's activity be allocated? It is hard to argue that there is a clear and correct answer to this question. These firms typically do not maintain records that compile turnover by establishment.

Strategies

By removing the methodological constraints of industry/activity classification, product surveys could be more easily collected at the enterprise level. Organizing data collection by product rather than by industry/activity could here again provide an opportunity to change legacy models that are becoming less in-step with modern business practices.

Establishment-level employment and wage data can in most cases be collected at an enterprise statistical unit. If establishment-level turnover data are also needed for local or regional statistics, directly collected enterprise turnover figures could be allocated to the establishment level based on model-identified characteristics associated with turnover generation.

Challenge: Maintaining continuity of data series

In many countries, industry SPPI and turnover data have been disseminated and widely used for a number of years. Transitioning to product data in place of industry data would create significant disruptions to time-series analysis of developments in the services sectors.

Strategies

For NSOs with extensive histories of producing industry data, this is a significant challenge that eludes easy solution. It is likely that agencies that make this transition would create approximate industry-based data using techniques similar to those currently used to create approximate product-based data (see Section 6 below). The resulting approximate industry-based figures could then serve as continuous data series for the previously tabulated industry data.

6. Approximate Product-Based Data

In many cases approximate product-based data serve as a practical solution for creating wherever-made product tabulations without committing to full product-based data collection.

In the US, approximate product-based SPPIs are created by re-organizing the transaction data collected from industry-based surveys. For example, the home telephone services transactions obtained from wired, satellite, and other telecommunications providers are combined to create a single home telephone service product-based SPPI. The contributions of the aggregated transactions from each respective industry are weighted based on product-by-industry turnover data derived from US Economic Census publications.

Using product-by-industry turnover data to weight product-based SPPIs

	Wired telecommunications industry SPPI	Satellite telecommunication industry SPPI	Other telecommunications industry SPPI
Industry total turnover	$1 billion	$500 million	$250 million
Home telephone services	$200 million	$100 million	$50 million
Business telephone services	$100 million	$50 million	$25 million
Programming services	$400 million	$200 million	$100 million
Data services	$300 million	$150 million	$75 million

	Home telephone service product-based SPPI
Product total	$350 million
Home telephone services from wired telecommunications firms	$200 million
Home telephone services from satellite telecommunications firms	$100 million
Home telephone services from other telecommunications firms	$50 million

There are two significant challenges with this approach. First, individual establishment weights are derived by multiplying the collected establishment turnover by a "sampling factor". This factor increases the contribution of smaller units so that they also represent the activities of similarly sized units that were present but not selected from the industry sampling frame. Since these similarly sized industry units may not sell the same products, it may not be appropriate to represent their activity in the product-based SPPI data. Sampling factors that are suitable for a product-based SPPI should ideally be obtained through a product sample.

The second challenge is that some firms have leading market positions in a broad number of product areas. This is particularly true for large national retailers, such as those classified in NAICS industry 452910, Warehouse Clubs and Supercenters. For example, a large retailer may report the following sources of turnover:

Retailer A and total market sales by product line

	Retailer A Sales	% of Retailer A Sales	Total Market Sales	Retailer A Market Share
Groceries	$100 billion	71.0%	$1 trillion	10.0%
Pharmaceuticals	$20 billion	14.0%	$1 trillion	2.0%
Apparel	$10 billion	7.0%	$500 billion	2.0%
Electronics	$10 billion	7.0%	$200 billion	5.0%
Books	$1 billion	0.7%	$10 billion	10.0%
Greeting Cards	$500 million	0.3%	$2 billion	25.0%

If Retailer A were selected in an industry-based sample, standard probability proportionate to size item sampling would lead to the selection of grocery products for the majority of the items. Books and greeting card sales, each representing less than 1% of total turnover would typically not be selected at all.

However, in many cases large US retailers have leading market positions for products that generate a small portion of their total sales. While greeting card sales contribute a trivial portion of Retailer A's total turnover, these sales are very important in the context of total US greeting card sales. If a product-based survey was used to collect greeting card sales only, Retailer A would have substantial weight in the sample. To ensure adequate coverage in approximate product-based SPPIs, items that represent large shares of product markets should be judgmentally selected even when they represent a small portion of unit sales.

7. Conclusion

This paper supports the guidance established by the OECD-Eurostat SPPI manual that product-based data are most appropriate for SPPIs and turnover surveys. As secondary production and competition between firms in separate industries increases, the need for product data has become increasingly urgent. Still, the complexities of transitioning established data collection routines from product-based to industry-based present numerous challenges. Countries that are just beginning to produce SPPI and services turnover data may want to consider organizing initial data collection and tabulation by product to avoid the challenges associated with transitioning from industry to product. Lack of sampling frames by product and need for data continuity with existing industry data are also significant challenges that need to be addressed. Product-based data tabulated from industry-based data collection is an attractive option for satisfying users that would like both industry *and* product data. Strategies should be pursued to ensure adequate coverage of a full range of products from large enterprises with large amounts of varied production. This may require a more personalized data collection approach for these enterprises, such as account manager assistance, spreadsheet data reporting, and using large data file transfers in place of survey response.