Presented at 29th Meeting of Voorburg Group, Dublin, Ireland, 22nd – 26th September, 2014

# Maintaining Representative Turnover and SPPI: Re-stratification, Re sampling, Rebasing, and Updating Weights

Sayeeda Kamaruddin,

Department of Statistics, Malaysia

(DOSM)

**ABSTRACT**

Services in Malaysia have been growing in importance for the economy in the past few years. In 2013, Services was responsible for 55.2 percent of the GDP compared to only 46.8 percent in 1980. The concerted development of the service industry is part of the national development strategy to venture into new growth areas and broaden the economic base for exports. It is also expected to provide the basis for sustained growth in the economy in order to achieve the vision of becoming a developed nation by 2020. Recognizing the growth potential in the services sector, Malaysian Government has decided to immediately liberalize 27 services subsectors with no equity condition imposed to attract more foreign investments and bring more professionals and technology as well as strengthen competitiveness of the sector. Thus, a comprehensive compilation on services statistics is very crucial to access economic condition, identify any emerging trends or structural changes in the economy. In response to this growing needs, Department of Statistics, Malaysia has developed price indices for thirty five services activities, which help to form a comprehensive SPPI at national level. This paper provides an overview of methodology engaged in SPPI surveys and challenges arising during the index development.

1. **INTRODUCTION**

SPPI was compiled back in 2008 with coverage limited to Postal & courier and Freight transport by road activity using 2008 as the base year. Since then, it has undergo a series of development where the coverage is extended to Air & water transport in 2010, Information & Communication Technology (ICT) in 2011, Business activities comprised of Legal, Accounting, Engineering and Real estate agents in 2012 and followed by Accomodation, Private Health & Education and Arts, entertainment & recreation in 2013. For the purpose; price concept, methodology used and index calculation are based on OECD manual (Methodological Guide for Developing Producer Price Indices for Services). In the following sections, this paper will describe on objective, scope of the survey and methodology adopted in the sampling process. Technical aspects on the data analysing which includes quality control of the data are highlighted in the next section. Finally, the paper closes with a conclusion and areas for improvement.

2. **OBJECTIVE OF SPPI**

SPPI is defined as the price index, which measures the average change in prices of services produced by domestic producers. The purpose of the quarterly survey for the services producer price index (SPPI) is to facilitate a better understanding of macroeconomic situation and trace impact of price variation. It is also used as a deflator for recalculation of real values of gross output in services, after removing price factors from fluctuations in nominal values and served as an important macroeconomic indicator used for monitoring the price movements of local outputs particularly for services products and is often viewed as a leading indicator of Consumer Price Index (CPI).

3. **SCOPE AND COVERAGE**

SPPI only covers services produced by local/ domestic producers. Hence, imported services produced by enterprises which are registered in other countries, are not covered in this survey. It covers service products that are provided by businesses to other businesses (BtoB). Although service products provided to individuals are beyond the scope of SPPI, the index covers some service products (postal and courier and telecommunication services) that are provided to both individuals and businesses (BtoAll). For the purpose of this survey, industrial classification is using Malaysian Standard Industrial Classification (MSIC) 2008 (Ver.2) which is in conformity with the ISIC

Rev.4. Meanwhile, product classification is using Malaysian Classification of Product by Activity (MCPA) 2009 which is compiled based on MSIC 2008 and United Nations Central Product Classification (CPC) 2.0.

The survey covers the following services sectors where samples are using the same sample as in Quarterly Services Survey (QSS).

| Sectors | MSIC 2008 | Description | Contribution of Output (%) |
|---|---|---|---|
| Transport | 49230 | Freight transport by road | 78% |
| | 50121 | Transport of freight overseas and coastal waters | 83% |
| | 51201 | Transport freight by air | 87% |
| | 53100 | National post services | 100% |
| | 53200 | Courier activities | |
| ICT | 61101 | Wired telecommunication services | 83% |
| | 61102 | Internet access providers by the operator of wired infrastructure | |
| | 61201 | Wireless telecommunication services | |
| | 61202 | Internet access providers by the operator of wireless infrastructure | |
| | 62021 | Computer consultancy | 35% |
| | 63112 | Data processing activities | 49% |
| | 63120 | Web portals | |
| Real estate agents | 68201 | Real estate agents and brokers | 61% |
| Professional | 69100 | Legal activities | 50% |
| | 69200 | Accounting | |
| | 71102 | Engineering services | |
| Accommodation | 55101 | Hotels and resort hotels | 83% |
| Education | 85102 | Pre primary education | 100% |
| | 85104 | Primary education | |
| | 85212 | General school secondary education | 88% |
| | 85302 | College and university education | 100% |
| | 85421 | Music and dancing school | 40% |
| | 85491 | Tuition centre | |

| | 86101 | Hospitals | 99.8% |
|---|---|---|---|
| Health | 86201 | General medical services | 100% |
| | 86202 | Specialized medical services | |
| | 86203 | Dental services | |
| | 86901 | Dialysis centre | 62% |
| | 86902 | Medical laboratories | |
| Arts, Entertainment & Recreation | 93113 | Equestrian clubs | 66% |
| | 93116 | Golf courses | |
| | 93117 | Bowling centre | |
| | 93118 | Fitness centre | |
| | 93210 | Amusement parks and theme parks | 72% |
| | 93297 | Cyber cafe/ internet café | |

[1] Real estate services with owned or leased property is not covered since this industry is not covered in QSS

The selection of industry to be covered is based on their output contribution to the total output for that subsector from latest available census data. Rule of thumb used is any industry which cumulatively contributes a significant share of output will be covered. The service products that are not covered are Financial Intermediation, Distributive Trade, Food & Beverages, Administrative & Support Services and Other Services.

## 4. SAMPLING DESIGN

### 4.1 Selection of Primary Sampling Unit (PSU)

In favour of cost-effectiveness and logistic purpose, SPPI is using the same sample from Quarterly Services Survey (QSS). QSS adopted a single stage stratified sampling design with two main strata. The primary strata comprise of establishments with revenue of above certain cut-off value. For this survey, cut-off value applied is not less than 75 per cent of the cumulative revenue at industry level. These establishments will be completely enumerated. Subsequently, establishments below the cut-off value will be further subdivided into substrata so that the variation within strata is small relative to the variation between strata. In opposition with primary strata, establishments in these strata will be sampled. The basis of sample size estimation is 95% confidence level by specifying a coefficient of variation of not more than 10% at national

level. Once the sample size has been specified, it will be disaggregated into their substrata using proportional allocation.

## 4.2    Selection of Secondary Sampling Unit (SSU)

A sample of fixed items per business unit constitutes the second stage sample. A second stage using probability sampling requires a frame of all transactions for each establishment selected at the primary stage along with their contribution to the total value added. However, to construct such a frame is time-consuming, costly and imposed a higher burden of the selected respondents. Hence, the selection of items is conducted using one of non-probability sampling method which is cut-off approach where respondents are asked to quote price from the listed items based on the measurement unit given. Items for every industry are pre- identified through a preliminary survey before the actual survey. This survey is carried out to identify the major services produced under each of the selected as well as detailed description/ specifications of the services offered and unit of measurement of items produced for price collection purposes. Questionnaires are dispatch to selected establishments at the initial stage to inquire the following information:

(a)     Regularly service produced
(b)     Detailed description/ measurement unit of the service produced of the services stated in (a)

The sample size of items varies across the industry. It will be reviewed every 6 months based on respondent's feedback to ensure that any new services offered is covered and non-significant or discontinued services is removed.

## 4.3    Resampling

At present, SPPI is duplicating QSS sample. Hence, resampling will only be done once sampling design for QSS is changed using a newly updated sampling frame. Current practice is to revise the sampling design once the sampling frame is updated, either from latest annual survey or census.

## 5. QUALITY CONTROL

### Non-sampling error

Aside from the sampling error associated with the process of selecting a sample, a survey is subject to a wide variety of errors. These errors are commonly referred to as "non-sampling errors". Non-sampling errors can be defined as errors arising during the course of all survey activities other than sampling process, may occurs during sample or census and are difficult to measure. In order to minimize the errors, a series of checks on the prices data received are carried out to ensure their reasonableness, consistency and accuracy. The checking procedures include comparison of prices supplied for the month against those of the previous quarter of the same respondents as well as price reported by other respondents for the same items. Prices that display unusual variation are clarified with respondents to ensure the accuracy. Prices quoted in the forms are verified on the ground by field supervisors and this are subjected to further scrutiny in the headquarters. Significant price changes are referred to the field supervisors for verifications and in certain cases, the respondents are directly contacted for further clarification. The data are subjected to thorough checking before captured. As far as the data collection is concerned, the field operators are given training at least once a year regarding procedures, concepts and definitions as well as any changes in the selected items or item's specifications.

### Outlier treatment

Outlier detection and treatment is a very important part in any data analysis. A failure of detecting outliers or their incorrect treatment can have serious ramifications on the validity of the inferences drawn from the analysis. For the moment, any values changing by more than 10% in either direction is considered as outlier and need to be clarified by the respondents. Often, the treatment for outliers is using the basic methods, which are data dropping, carry forward imputation and trimmed mean method.

### Treatment on missing prices

If no price report is received for that particular quarter, the sample price is assumed unchanged from the last observed price. Other correction actions taken are imputing the missing prices using the average price change for the prices that are available in the elementary aggregates and imputing the missing prices using the average price change for the identical items from other establishments in the same industry.

## 6 DISSEMINATION/ PUBLICITY AND TRANSPARENCY

DOSM disseminates quarterly information on PPI in services 30 days after the reference period. The information contains basic SPPI and also indices in comparison with the previous quarter and the corresponding quarter of the previous year. At the present time, SPPI is only meant for internal circulation mainly as a deflator for Index of Services (Ios) and calculating the real value of output in services from nominal values.

## 7 AREAS FOR IMPROVEMENT

### 7.1 Reweighting

Currently, weights at industry level are using 2010 Economic Census. The reweighting process will be applied using the latest available census data, which is 2015 Economic Census which would be readily used by 2016.

### 7.2 Treatment of Missing data

The collection of data requires consistency quarterly price data for the items measured. In practice, however this condition always not satisfied where missing data will always occur. This may happen due to non-respond cases where no information is provided for several items (missing items) or no information is provided for a whole unit (missing observation). Currently, imputation method using last available price is applied. However, the imputation process is quite complicated especially if there is permanent data missing.

### 7.3 Sample Selection

Due to certain limitations, SPPI is using the same sample as in Quarterly Services Survey (QSS). QSS aims to measure turnover whereas SPPI measures the average change of prices across industries. The sample size generated in every stratum will largely depends on the variation of turnover where stratum with larger variation will be selected more than those with smaller variation. Nevertheless, a large variation in revenue does not necessarily to have an impact of the price stability. Thus, a dedicated sampling frame and sampling plan for SPPI should be developed once the time series data for SPPI is available. The initial sample sizes would need to be revised when sufficient survey data of SPPI become available,

response patterns have been analyzed, and measures of precision have been computed. A stratum may have a low response rate, but if the price movement within is stable, the corresponding sample size would not have to be augmented. Similarly, highly variable index may still require an increase in the sample size to offset the variability, while a stable index could lead to a reduction in the sample size.

7.4    Enhancement of sampling plan

SPPI is a fixed panel survey where the same sample is used for a period of time (minimum of 3 years). In Malaysia case, the same sample has been using since 2012 and resampling process will take place after the establishment frame is updated by latest information during the next 2015 Economic Census. This practice may increase respondent burden and when this happen, it will reflect the data quality as respondents may be not willing to respond or keep reporting the same number. As an alternative, other survey sampling strategy such as rotating panel design maybe introduced. Rotating panel is a survey sampling strategy sometimes used when estimates are produced regularly over time. Under such a design, equally sized sets of sample units are brought in and out of the sample in some specified pattern. These sets, often called rotation groups, may be composed of households, business firms, or other units of interest to the survey. Rotating panel designs are used to reduce the variances of estimators of level or change and often to reduce the survey costs associated with introducing a new unit into the sample. Currently, rotating panel design only applied to Labour Force Survey. However, there should be a further study on the applicable of the design and to the price estimation process before the implementation.

7.5    Outlier Treatment

The treatment for outliers is using the basic methods, which are data dropping, carry forward imputation and trimmed mean method. Extension of more dynamic outlier detection dedicated for price index survey such as Quartile Method, Tukey Algorithm, Resistance Fence and Mean-Absolute Deviation (MAD) together with the appropriate outlier treatments should be explored in future and requires further study.

7.6     Building Smart Partnership

In Malaysia, DOSM has moved towards intensifying data tapping on secondary and administrative data from relevant agencies instead of relying on primary data to cater to the increasing demand for statistics. This is vital in order to avoid duplication of efforts and to reduce the respondents' burden as well as to use the resources more effectively. Thus, more engagement with data providers on the availability of SPPI indices which are not collected via SPPI surveys is in needed.

7.7     Representativeness of item's specification

It is essential to identify the item's specification for which price data are needed as it is not possible to include all the services produced in the country. Hence, the items are review on regularly basis to ensure that it covers any new services and withdraw any discontinue services. This updating is heavily relied on respondent's feedback and normally, respondents would only specified the price for the listed items without reporting any new services produced. Thus, the available item's specification need be reviewed frequently so as to ensure its representativeness.

7.8     Extension of coverage

At present, SPPI covers 8 subsectors with a total of 35 industries. Meanwhile, study on the enhancement on scope of survey has been done continuously. Currently, two industries have been undertaken for further study which is Travel agent & Advertising. As for other sectors, industries that have significant output value will be included for further analysis and study.

## 8   CONCLUSION

Our future plan is to disseminate the SPPI to public since there is an increasing demand on numerous short term indicators on services by the stakeholders and public. Since SPPI is a new product by Department of Statistics, Malaysia and still under development, there are still many areas need to be further improvised in order to produce and generate quality index.