DATA GENERATION GUIDELINES

2012 Occupational Wages Survey (OWS)

Republic of the Philippines
DEPARTMENT OF LABOR AND EMPLOYMENT
BUREAU OF LABOR AND EMPLOYMENT STATISTICS
Manila, Philippines

ISO 9001:2008 Certified

May 2013

TABLE OF CONTENTS

OBJECTIVE

This Data Generation Guidelines for 2012 Occupational Wages Survey (OWS) is prepared to facilitate the generation of output tables needed in the preparation of the publication tables.

Preparatory activities should be done by concerned personnel before generating desired statistical tables. In addition, status codes need to be reconciled and necessary adjustments should be made before the generation of such tables.

From the 2012 OWS, the following statistics will be generated:

- occupational wage rates (average of current rates) and employment of time-rate workers on full-time basis in selected industries and selected occupations
- median basic pay and allowances of time-rate workers on full-time basis by industry/region
- distribution of time-rate workers on full-time basis by basic pay and allowance intervals by industry/region

A. PREPARATORY ACTIVITIES

1. All appropriate tabulations, cross-tabulations or disaggregations of the data categories should have been determined during the Pre-Field Operations stage of the survey, in order that,

    a. tabulations of the same data variables/categories are maximized, hence processing time is minimized;
    b. survey results that would be made accessible to the data users are comprehensive, thus, no need for further tabulations;
    c. consistency checks are facilitated; and
    d. reasonableness of the survey results at detailed levels are easily determined which would not be possible at aggregate levels. For one, errors in coding (industry) may be covered up when the data are tabulated at higher levels of disaggregations.

2. Based on the proposed data tabulations, the computer syntax should have been prepared prior to table generation.

B. RECONCILIATION OF STATUS CODES AND SELECTED DATA

1. For DUP establishment, track its duplicate or the establishment to be retained.

    a. If the establishment to be retained has responded, it should have the lower EIN. If it does not have the lower EIN, replace it with the lower EIN while the DUP establishment is assigned with the higher EIN. Revise name, address, industry and geographic codes and status of the involved establishments accordingly in the status monitoring database.

        Example:

        Establishment Y with EIN 426 is DUP of Establishment X with EIN 678. Establishment X has responded.

        The EIN of Establishment X should be replaced with 426 and its status code will be changed to RET. The EIN of Establishment Y should be 678 and its status code becomes DUP of EIN 426.

    b. If the establishment to be retained has not responded, it should still have the lower EIN. The necessary revisions in the EIN and other particulars as mentioned above should be made in the status monitoring database for the DUP establishment and the establishment to be retained.

    c. If there are more than one duplicates of an establishment, the principle remains the same. The establishment to be retained has the lower EIN, the duplicates have the higher EINs and the status codes of the involved establishments are adjusted in the status monitoring database.

2. For CET establishment, track the EIN/s of the establishment/s or the questionnaire/s in which reports are included in the CET questionnaire. The status code of these sample establishments should be CON with EIN ___ (EIN of the CET establishment).

In case the sample establishment/s of the CET establishment has also responded and the status "RET" and data values have been encoded, change the status code of the establishment/s to CON with EIN ___ (EIN of the CET establishment) in the status monitoring database. Also the data values that have been encoded should be deleted in the respondents' database thus retaining the responses of the CET establishment only.

3. Always check consistencies in the EIN, status code, industry code and employment for BITS and OWS questionnaires. Print validation prooflist to reconcile/harmonize variables..

4. Ensure that the totals of retrieved/processed (RET, CET) and "spoilage" (REF, STR, TCL, CBL, PCL, OSE, OSP, DUP, CON, OTH) questionnaires in the database are the same as their corresponding total number of questionnaires recorded in *FM-BLES 03-3.19 Status Monitoring of Returned Questionnaires* (retrieved and "spoilage") and *FM-BLES 04-4.7 Monitoring of Data Processing Activities* (encoded questionnaires). Confirm further that the sum of RFVs and unaccounted questionnaires is the difference of retrieved/processed and "spoilage" questionnaires from the sample size. Should there be any discrepancy, it probably would be due to the adjustments on status codes made for DUP, CET and CON questionnaires.

C. REQUIRED STATISTICAL TABLES

1. After adjustments in status codes and reconciliation of selected data have been made, generate the *final survey status report (FM-BLES 03-3.17 Assessment on the Implementation of Field Operations of BLES Survey/s)*.

2. To aid in the preparation of the BUFs (Blowing-up Factors), generate the preliminary *Table A - Distribution of Establishments and Weighted Retrieval Rates by Industry Group and Employment Size*. Note that the "transfers to and from" of establishments should have been taken into account in this distribution.

   a. For each industry and employment size, ensure that:

   $$\mathbf{N'_{hk,lm} > eligible_{hk,lm} > n'_{hk,lm}}$$

   where:

   $N'_{hk,lm}$ — estimated population in the initial stratum k and h and in the post-stratum l and m.
   It is estimated based on an eligibility ratio, i.e.,
   $N'_{hk,lm} = N_{hk} \times eligible_{hk,lm}/n_{hk}$
   where N_{hk} is population count in the initial stratum k and h and n_{hk} is sample count in the initial stratum k and h.

   $eligible_{hk,lm}$ — number of eligible samples (RET, CET, REF, STR, TCL, RFV, Unaccounted) in the initial stratum k and h and in the post-stratum l and m

   $n'_{hk,lm}$ — responding samples in the initial stratum k and h and in the post-stratum l and m

b. Evaluate this distribution to determine the stratum (employment size) of the domain (industry) or cell (industry and employment size) that should be collapsed because of low response or non-response relative to the number of eligible samples in the stratum or cell. Make the necessary adjustments in the stratum or cell.

3. Generate the final *Table A - Distribution of Establishments and Weighted Retrieval Rates by Industry Group and Employment Size*. Note that adjustments for CET and CON establishments and collapsing of strata or cells undertaken should have been taken into account in this distribution table.

4. Compute the BUFs (ratio of the $N'_{hk,lm}$ to $n'_{hk,lm}$) per establishment record.

5. Generate the pre-determined *output tables*.

   Evaluate numerical consistency of the data of a variable in a statistical table and across statistical tables where the same variable appears. In particular, check for consistency across totals of the same data variable that were disaggregated into different categories (industry and employment size).

   For industries without monitored occupations, there should be no data for occupations other than the benchmark. If there are, these should be deleted from the database.

6. Prepare *publication tables*

   a. Reflect industry/occupation/region even if without data. Put footnote as applicable e.g., no report, suppressed due to low response rate, etc.

   b. Since the presentation of the publication tables may be different from the output tables (e.g. percentages or averages), reference should be made to previously published survey results for comparability with updated survey data. Recent economic developments or issuances, e.g. wage orders between previous and current survey should also be considered to explain any variations in the data.

   c. Part of the validation process also takes into account coherence checks with related survey data generated by BLES or by other establishment surveys (total employment against similar data from the List of Establishments; trend in total employment against that of the employment index released by the National Statistical Coordination Board from the Quarterly Survey of Philippine Business and Industry, among others).

      Reference to administrative statistics should also be made to determine coherence of survey data (the trend in median monthly basic pay should be looked into for comparability with that of the compensation index per employee at nominal terms from NSCB, among others).

      Coherence does not necessarily mean full numerical consistency. “The coherence of statistical information reflects the degree to which it can be successfully brought together with other statistical information within a broad analytical framework and over time.”

D. FINAL ASSESSMENT OF ACCURACY IN DATA PROCESSING

Though controls are in place during data collection, and editing/validation and data encoding, there may have been lapses that were overlooked during these stages of survey operation. It is essential then that inaccuracies in data processing, to the extent possible, be finally detected and corrected during output table generation.

Should inconsistencies in output tables be noted, backtracking should start from the rejection list to validation prooflist then to the questionnaire itself.

To monitor possible inaccuracies at this point, refer to *FM-BLES 04-4.8 Monitoring of Accuracy in Data Processing*. This instrument, together with the other survey monitoring forms, will help survey managers to determine areas for improvement not only in the survey procedures but more importantly in enhancing the skills of data processors.