Redesigning of the Commercial Livestock and Poultry Survey (CLPS) in the Philippines

Nikkin L. Beronilla[1], Quindale E. Caraos[1], WIllreign Zoren D. Dela Cruz[1], Divina Gracia L. Del Prado, Ph. D.[2], Jay R. Manlapaz[1]

[1]Statistical Methodology Unit, Philippine Statistics Authority
[2]Economic Sector Statistics Service, Philippine Statistics Authority

ABSTRACT

The Commercial Livestock and Poultry Survey (CLPS) is one of the major agricultural surveys conducted by the Philippine Statistics Authority (PSA). This survey aims to generate estimates on supply, disposition, average farmgate price, and other related data of livestock and poultry commodities from establishments. The old design of CLPS uses stratified random sampling with maximum farm/housing capacity as the stratification variable. In the old design, the number of strata vary per animal type per province which imposes difficulty in terms of field operations and since the stratum boundaries are not updated regularly, there is a possibility that some farms/establishments that have increased in terms of maximum farm/housing capacity are still on lower strata. Given the situation, a new sampling design is proposed in order to generate reliable and accurate estimates at the provincial level with the use of the regularly conducted Updated List of Establishments (ULE) as its new sampling frame.

After performing sampling experiments and simulations, the new sampling design for CLPS is a one-stage Stratified-Probability Proportional to Size (PPS) where the domain is the province, and the establishment is the primary sampling unit (PSU). The animal type serves as a stratification variable and the size measure is the maximum farm/housing capacity of the establishment. The new sampling design has better design effect (DEFF) and coefficient of variations (CVs) for indicators estimated in CLPS as compared to the old design.

***Keywords**: Agriculture, Livestock, Poultry, Establishment, Single-Stage Stratified-Probability Proportional to Size (PPS) Sampling Design*

1.0 Introduction

1.1 Background and Rationale

The Commercial Livestock and Poultry Survey (CLPS) of the Philippine Statistics Authority (PSA) is a major survey conducted by the Livestock and Poultry Statistics Division (LPSD) to measure the performance of livestock and poultry industry. Specifically, the survey aims to generate primary data on supply and disposition of animals from commercial farms. Its sampling frame is based on the 2006 Avian Population Survey and 2010 Livestock Population Survey. During survey rounds, Frame Maintenance Forms (FMFs) are used to update the CLPS frame. Currently, the unit of enumeration can be either commercial farms or households which tends to at least:

a. 21 heads of livestock, or
b. 1000 broiler birds, or
c. 500 layer birds, or
d. 100 ducks.

The current sampling design of CLPS is stratified random sampling with the provinces as the domain. Its strata boundaries were delineated using the Dalenius-Hodges (DH) method with maximum farm/housing capacity as the stratification variable. The samples are allocated across strata using Neyman allocation with a 5% target coefficient of variation (CV). However, a more recent stratification algorithm may be considered. Like DH, the Lavallee-Hidiroglou (LH) algorithm is an R function which can be found inside the package 'stratification' and is said to perform well even with skewed distribution of the stratification variable.

However, the stratum boundaries are not updated on a regular basis. In fact, the boundaries computed when the survey design was first implemented are the ones still being used. The only recent updating was done during the enhancement of the survey design in collaboration with the Statistical Methodology Unit (SMU) of PSA.

During a series of meetings and discussions conducted by LPSD and other units, it was decided to redesign the CLPS to align with the existing PSA concept on establishment-based surveys. Thus, the 2021 Updating of Listing of Livestock and Poultry Establishments (ULLAPE) was conducted in the provinces of Bukidnon and Laguna. Specifically, the special listing activity was done to come up with an updated register that will be used in simulation exercises for the redesigning of the survey. In addition, the frame may serve as the basis of sample selection for the pilot survey in the two provinces. In the long run, the goal is for the CLPS to base its frame completely on the regularly conducted Updated List of Establishments (ULE).

Given that there exists an updated establishment-based frame, it is of interest to perform sampling exercises to provide empirical basis on the redesigning of CLPS. Furthermore, the effectiveness of total employment variable, which is a common stratification variable of various establishment-based surveys, will be considered as well.

1.2 Objectives

This study aims to determine an updated sampling design for the Commercial Livestock and Poultry Survey. Specifically, it aims to:

1. identify appropriate auxiliary variable;
2. compute sample sizes that is minimal and is close to the current counts;
3. recommend a sampling design to be used in the pilot survey; and
4. assess the performance of estimates based on the pilot survey results.

1.3 Scope and Delimitation

This research study will provide the best sampling design based on the acceptable reliability measures and implementation considerations. However, only establishments that are reported to be in operation as of the 2021 ULE enumeration period are included in the simulation exercises.

1.4 Significance of the Study

The results of the study will update the sampling design with consideration to the shift of survey concept bearing from the change of the sampling frame of the survey. This sampling design will ensure the reliability and accuracy of the Commercial Livestock and Poultry estimates at the provincial level.

2.0 Methodology

2.1 Survey Data Experiments

2.1.1 Sampling Frame

The ULE 2021 register served as the sampling frame of the simulation exercises. This frame contains information on the two possible auxiliary variables, namely, Maximum Farm/Housing Capacity (MAXCAP) and Total Employment (TE). In addition, the information on a farm's Inventory is also present in the frame. Since both the auxiliary variables and the target dependent variable are present in the frame, there is no need for data generating process.

2.1.2 Auxiliary Variable

The CLPS currently uses MAXCAP as its stratification variable. It can still be considered as the auxiliary variable for the proposed design, given its usually strong correlation with the target dependent variable. However, TE is a commonly used stratification variable for establishment-based surveys of PSA. Both the variables are known to follow positively skewed distributions. That is, majority of observations tend to concentrate on lower values, with some outliers that contain high values.

In this study, the correlation between an auxiliary variable and inventory was checked across all domains. Figures 1 shows the correlation of MAXCAP, Total Employment, coded employment size, which is based on the Employment Stratum by TE Size used in the Annual Survey of Philippine Business and Industry (ASPBI), with Inventory per animal type.

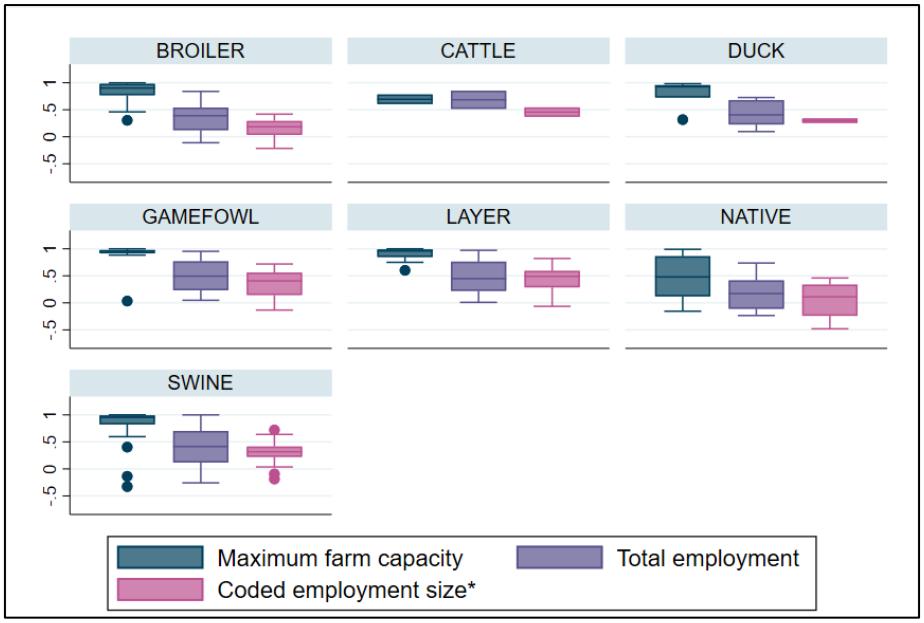

Figure 1. Distribution of Correlation Coefficients per Animal Type

Per animal type, it is observed that the correlation with inventory is strong for maximum farm/housing capacity. To further prove that MAXCAP is the best auxiliary variable to be used, different sampling scenarios using MAXCAP and TE as auxiliary variables will be tested in the simulation.

2.1.3 Number of Sample Establishments

For the animal type in a province that will proceed to the survey design, the next step would be the computation of sample size. The main considerations are sample size must be near to the old sample size, and it must be minimal. Below is the computed sample size per target percentage of coefficient of variation (CV%) and stratification method (DH and LH).

Table 1. Sample Size Computer per Target CV(%) and Stratification Method

Animal Type	Old Sample Size (4th Quarter of 2021)	Stratified DH			Stratified LH		
		Sample Size per Target CV (%)			Sample Size per Target CV (%)		
		1	3	5	1	3	5
Broiler	807	1655	900	660	1197	875	640
Carabao	127	53	53	53	53	53	53
Cattle	688	177	155	148	154	147	147
Duck	408	249	180	158	214	199	160
Gamefowl		362	299	280	307	280	276
Goat	414	121	121	121	121	121	121
Layer	725	1186	753	621	887	729	582
Native		436	368	347	406	365	368
Swine	620	1150	761	623	747	604	545
Total	**5095**	**5389**	**3590**	**3011**	**4086**	**3373**	**2892**

The computed sample size using CV = 3% for LH method in Table 1 was used to generate estimates for the evaluation of the different scenarios.

2.1.4 Sampling Scenarios

Since the estimates are desired at the provincial level, the provinces serve as the domain. In addition, establishments are selected independently by animal type to ensure the presence of samples per category.

The sampling experiment begin with the tabulation of farm frequencies to determine whether to proceed to complete enumeration or to proceed with sampling. Moreover, if the number of establishments in a province per animal type is less than 25, complete enumeration is done. Otherwise, proceed to sampling.

One of the scenarios tried follows the old design, that is stratified random sampling using Dalenious-Hodges method of stratification. For other scenarios, stratification method is changed into either Lavallee-Hidiroglou (LH) algorithm, or the Adjacent Means Iterative method. The LH algorithm gives the minimal sample size and performs well with skewed distribution. On the other hand, the Adjacent Means Iterative method is an algorithm developed by the consultant of Backyard and Livestock Poultry Survey redesigning, Dr. Erniel Barrios. Also, two other scenarios where the selection of establishments are done using Probability Proportional to Size (PPS) per animal type. PPS ensures that establishments with high maximum farm capacity will be given priority in the selection for every survey round. The sampling designs considered are summarized in Table 2.

Table 2. Sampling designs for simulation

Scenario	Selection	Auxiliary variable	Stratification method for Sample Size Determination
1	Stratified-SRS	MAXCAP	LH
2	Stratified-SRS	MAXCAP	DH
3	Stratified-SRS	MAXCAP	Adjacent Means
4	Stratified-PPS	MAXCAP	
5	Stratified-SRS	TE	LH
6	Stratified-SRS	TE	DH
7	Stratified-SRS	TE	Adjacent Means
8	Stratified-PPS	TE	

2.2 New Sampling Design

The approved new sampling design and estimation procedure for CLPS are discussed in this section.

2.2.1 Sampling Frame

The CLPS sampling frame is based on the 2021 List of Establishments (LE) engaged in livestock and poultry raising. The PSA maintains a frame management system through which updating and maintaining of sampling frames are done. For CLPS, the sampling frame is updated and maintained at

LPSD using the structured Frame Maintenance Form 1 and 2 (FMF 1 and FMF 2).

2.2.2 Sampling Design

The sampling design for CLPS is single-stage stratified probability proportional to size, where the domain is the province, and the establishment is the primary sampling unit (PSU). The animal type serves as a stratification variable and the size measure is the maximum farm/housing capacity of the establishment. However, complete enumeration is done for provinces with less than 25 establishments per animal type. A panel sample establishment will be maintained for a year, and a new set of samples will be selected during the first quarter of the next year.

2.2.3 Estimation Procedure

The estimation procedure for provinces with completely enumerated establishments is the same with those provinces where sampling was applied.

The base weights per animal type is given as follows:

$$w_i = \frac{X}{aX_i}$$

Where:

$w_i =$ base weight of establishment i
$a =$ number of sample establishments in the province
$X_i =$ maximum farm/housing capacity of establishment *i*
$X =$ total maximum farm/housing capacity of the province

The adjustment factor formula is given as follows:

$$AF_p = \frac{\sum_{i=1}^{a} w_i E_i}{\sum_{i=1}^{a} w_i R_i}$$

Where:

$AF_p =$ adjustment factor for province *p*
$E_i =$ eligibility status of establishment *i* (1 if eligible, 0 otherwise)
$R_i =$ responding status of establishment *i* (1 if responding, 0 otherwise)

The final weight formula is given as follows:

$$w_i' = w_i \times AF_p$$

Estimation of province total for period t is done per animal type and the formula is given as follows:

$$\hat{Y}_{tp} = \sum_{i=1}^{a} w_i' y_i$$

Where:

$\hat{Y}_{tp} =$ estimated total for domain *p* at period *t*
$y_i =$ survey data (inventory, production, etc.) for establishment *i*

The formula for estimated variance of $\hat{Y}_{tp}$is given as follows:

$$\hat{V}(\hat{Y}_{tp}) = \left(1 - \frac{a}{A}\right) as^2$$

Where:
$\hat{V}(\hat{Y}_{tp})$ = estimated variance of $\hat{Y}_{tp}$

$$s^2 = \frac{\sum_{i=1}^{a}\left(y_{tp} - \bar{y}_{tp}\right)^2}{a-1}$$

For any estimate of total, the formula for the coefficient of variation (CV) is given as follows:

$$CV(\hat{Y}) = \frac{\sqrt{\hat{V}(\hat{Y})}}{\hat{Y}} \times 100\%$$

Where:
$\hat{Y}$ = any estimate of interest (e.g. $\hat{Y}_{tp}$ or $\hat{Y}_{p}$)

3.0 Results and Discussion

The results of the survey data experiment as well as the pilot survey results that led to the approval of the new sampling design of CLPS are discussed in this section.

3.1 Survey Data Experiments

As seen in Table 3, higher frequencies with less than 10% CVs are observed using PPS sampling design with MAXCAP as auxiliary variable.

Table 3. Frequency of Provinces and Animals by CV (%) of Inventory Estimates and Sampling Design with MAXCAP as Auxiliary Variable

Sampling Design	Range of CVs (%)					Total
	[0,10)	[10,20)	[20,30)	≥30	NA	
PPS	70	4	0	0	0	**74**
StRS - LH	69	5	0	0	0	**74**
StRS - DH	56	11	2	3	2	**74**
StRS - AM	33	34	5	2	0	**74**

While in Table 4, higher frequencies with less than 10% CVs are observed using PPS sampling design with MAXCAP as auxiliary variable.

Table 4. Frequency of Provinces and Animals by CV (%) of Inventory Estimates and Sampling Design with TE as Auxiliary Variable

Sampling Design	Range of CVs (%)					Total
	[0,10)	[10,20)	[20,30)	≥30	NA	
PPS	16	24	19	15	0	**74**
StRS - LH	14	21	15	18	6	**74**
StRS - DH	11	18	17	19	9	**74**
StRS - AM	4	27	16	26	1	**74**

Comparing the two auxiliary variables, CVs using the PPS design using MAXCAP outperformed the PPS design with TE as the auxiliary variable.

Considering the CVs per animal type, PPS and LH with MAXCAP as auxiliary variable are two competing designs (see Figure 2). However, PPS performed better than LH.

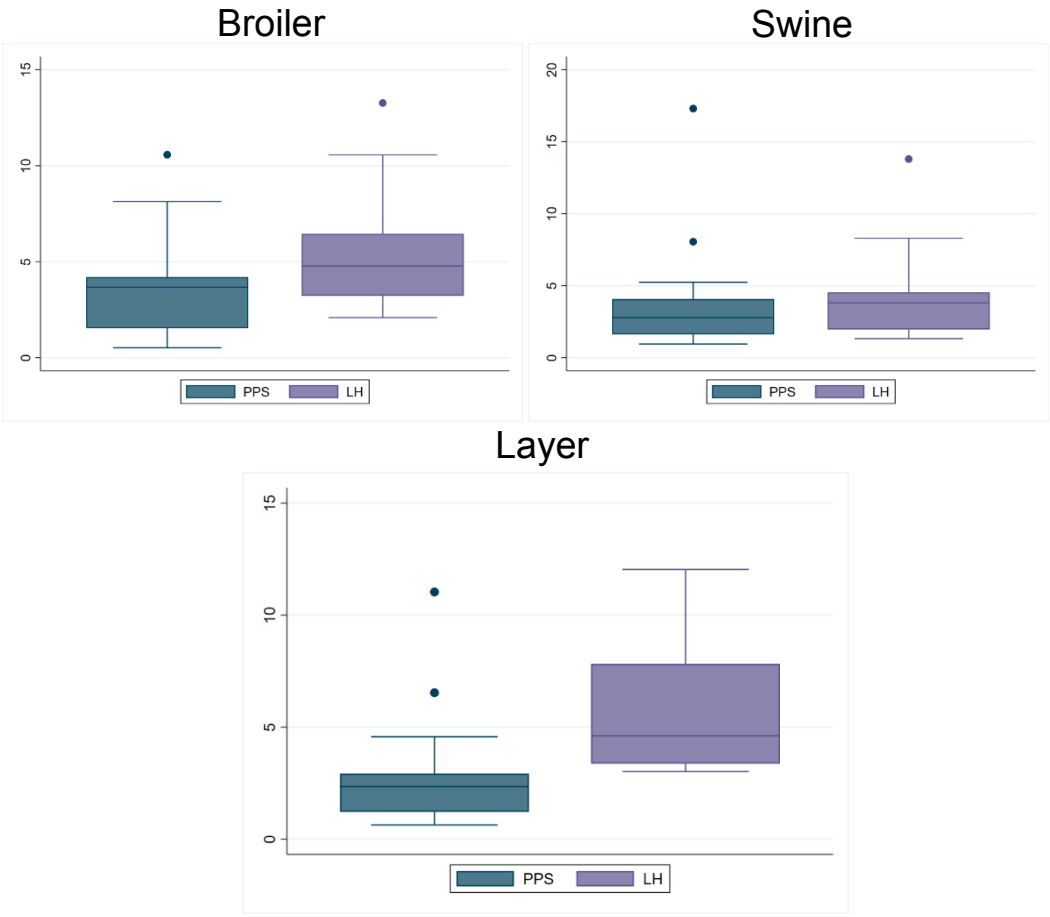

Figure 2. Distribution of CVs per Animal Type and Sampling Design

3.2 Parallel Pilot Survey

Laguna and Bukidnon are the two pilot provinces to evaluate the current and proposed sampling designs of CLPS. The pilot survey was conducted on 21 March to 08 April 2022.

The generated total from the proposed design is generally higher compared with the results of the current design as shown in Table 5. In Bukidnon, total from the proposed design are consistently higher than the results of the current design except for the production of swine. In Laguna, computed total using the proposed design is lower than the generated total using the current design.

Table 5. CLPS Current vs Proposed Design: Total Ending Inventory, Total Production (in MT), and Total Egg Production in Bukidnon and Laguna

Indicators		Bukidnon	Laguna
Total	Inventory (End)	**Proposed** > Current	Proposed < **Current**
	Production (MT)	**Proposed** > Current Except in Swine	Proposed < **Current**
	Egg Production	**Proposed** > Current	Proposed < **Current**

Moreover, Table 6 illustrates that generally the computed design effect (DEFF) for the proposed design is lower compared with the DEFF of current design. In Bukidnon, the DEFF for the proposed design is lower than the DEFF of the current design except for the total ending inventory. In Laguna, the DEFF of proposed design is consistently lower than the computed DEFF of the current design.

Table 6. CLPS Current vs Proposed Design: Design Effect (DEFF) of the Total Ending Inventory, Total Production (in MT), and Total Egg Production in Bukidnon and Laguna

Indicators		Bukidnon	Laguna
DEFF	Inventory (End)	Proposed > **Current** Except in Layer	**Proposed** < Current Except in Broiler and Duck
	Production (MT)	**Proposed** < Current	**Proposed** < Current Except in Duck
	Egg Production	**Proposed** < Current	**Proposed** < Current Except in Duck

Lastly, in Table 7, the computed coefficient of variations (CVs) for the proposed design are consistently lower than the CVs generated using the current design. In Bukidnon, the CVs computed using the proposed design are lower than the CVs from the current design except for the ending inventory. In Laguna, the generated CVs from the proposed design are steadily lower than the CVs from the current design.

Table 7. CLPS Current vs Proposed Design: Coefficient of Variation (CVs) of the Total Ending Inventory, Total Production (in MT), and Total Egg Production in Bukidnon and Laguna

	Indicators	Bukidnon	Laguna
CV(%)	Inventory End	Proposed > Current Except in Layer	**Proposed** < Current Except in Broiler
	Production (MT)	**Proposed** < Current Except in Broiler	**Proposed** < Current Except in Swine
	Egg Production	**Proposed** < Current	**Proposed** < Current

4.0 Summary and Conclusion

The proposed sampling design, which is a single-stage stratified-PPS sampling design, is indeed the best sampling design for the Commercial Livestock and Poultry Survey (CLPS) because the design effect (DEFF) and coefficients of variations (CVs) of the proposed design are often smaller than those of the current design, as observed in the pilot survey.

In addition, sampling of establishments will be done every first quarter of the year, and all establishments sampled will also be interviewed for the succeeding quarters.

5.0 References

Baillargeon, S. and Rivest L.P. (2011). The construction of stratified designs in R with the package stratification. Survey Methodology, 37(1), 53-65

Gunning, P., Horgan, J.M., & Yancey, W. (2004). Geometric Stratification of Accounting Data. Revista Contaduría y Administración, No. 214, September-December.

Philippine Statistics Authority. Commercial Livestock and Poultry Survey. [Online] [Cited: 4 June 2021.] https://psa.gov.ph/surveys-agriculture/clps.

Philippine Statistics Authority. 2020 Annual Survey of Philippine Business and Industry: Field Operations Manual

Slanta, J. & Krenzke, T. (1994). Applying the Lavallee and Hidiroglou Method to Obtain Stratification Boundaries for the Census Bureau's Annual Capital Expenditures Survey. Conference Paper American Statistical Association 1994, Proceedings of the Section on Survey Research Methods: Volume 1

Statistics Canada. Biannual Livestock Survey. [Online] [Cited: 12 November 2021] https://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=3460.