Redesigning of the Corn Production Survey

Abubakar S. Asaad[1,2], Sarah B. Balagbis[1], Jay M. Manlapaz[1]
Melanie C. Estrada[1], Divina Gracia L. del Prado[3], Erniel B. Barrios[4]

[1]Statistical Methodology Unit, Philippine Statistics Authority
a.asaad@psa.gov.ph

[2]College of Public Health, Department of Epidemiology and Biostatistics,
University of the Philippines Manila
asasaad@up.edu.ph

[3]Economic Sectoral Statistics Section, Philippine Statistics Authority
g.delprado@psa.gov.ph

[4]School of Statistics, University of the Philippines Diliman
ebbarrios@up.edu.ph

Abstract

Redesigning of the Philippine Statistics Authority's (PSA's) nationwide quarterly Corn Production Survey (CPS) was done to reflect the current behavior of corn production in each province. Simulation results show that stratified Lavalle-Hidiroglou (LH) algorithm is the best stratification with probability proportional to size systematic (PPS-SYS) and simple random sampling without replacement (SRSWOR), as the best selection methods for sample barangays and sample households, respectively. Further, design-based estimation is recommended over bootstrap method, in terms of precision and accuracy.

Keywords: sampling frame, sampling design, design-based, bootstrap method

1 Introduction

The Corn Production Survey (CPS) is a nationwide quarterly agricultural survey conducted by the Philippine Statistics Authority (PSA) during the months of April, July, October and December. This survey is conducted simultaneously with the Palay Production Survey (PPS), and together, the two surveys are referred to as the Palay and Corn Production Survey (PCPS). The CPS aims to generate provincial estimates and forecasts on corn production, area and yield as inputs for government policies and programs on corn[1].

The domain of the survey is the *province* and it is a two-stage stratified sampling design with *barangay*, selected using probability proportional to size sampling, as the primary sampling unit (PSU) and the farming household, which is systematically selected, as the secondary sampling unit (SSU). The CPS uses the 1991 Census of Agriculture and Fisheries (CAF) as a sampling

frame, with the total farm area based on the 1980 Census of Agriculture (CA) served as the stratification variable (i.e., certainty stratum and non-certainty stratum) and the number of strata is either 10 or 5 depending on domain size. The coverage of the survey include Isabela, Laguna and Bukidnon wherein all barangays in the municipalities were covered. For the National Capital Region (NCR), selection of sample barangays was done independently by district (for certainty barangay - barangay with the largest TFA in the district and for non-certainty barangay - 10% of the remaining barangays in the district) and for other provinces outside the NCR (for certainty barangay - barangay with the largest TFA in the muncipality and for non-certainty barangay - 50% of the remaining bangays in the municipality)[2]. The sample farming households is computed as follows:

$$n_{khi} = \frac{1}{b_h} \times \frac{P_{kh}}{P_{khi}} \times \frac{N_{khi}}{R_k}$$

where: $n_{khi} \equiv$ number of sample farming households in the ith sample barangay in the hth stratum in the kth province, $N_{khi} \equiv$ total number of farming households in the ith sample barangay in the hth stratum in the province, $R_k \equiv$ uniform raising (expansion) factor used for the kth province, $b_h \equiv$ number of sample barangays in the hth stratum (sample barangays), $P_{khi} \equiv$ corn area of the ith sample barangay in the hth stratum in the kth province P_{kh} aggregate corn area in the hth stratum in the kth province.

The major weaknesses of the current design are: obselete frame; lack of weight adjustment for non-responding units (if no, applied imputation); weights of uncertainty barangays in the 1991 CAF are not taken into account in the survey weights of CPS; and unclear procedure in determining the sample size.

The histogram of corn production and corn farm area (1st Semester) in the Philippines using the 2013 CPS[3] data and 2012 CAF[4] respectively, were skewed as seen in Figures 1 and 2. The correlation coefficient of corn production and corn farm area by quarter is from moderate to high is presented in Table 1.
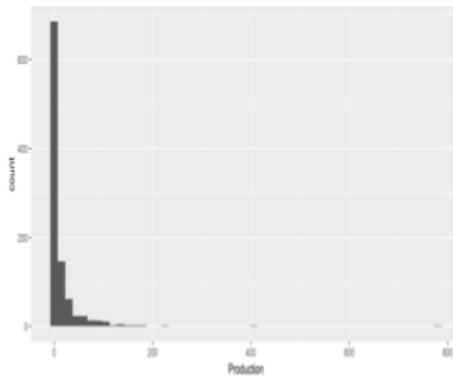

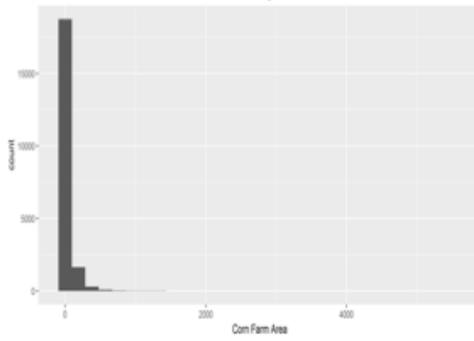Figure 1: Histogram of Corn Production, Philippines: 2013

Figure 2: Histogram of Corn Farm Area, Philippines: 2012

Table 1: Correlation of Corn Production and Corn Farm Area

Quarter	Correlation
Q1	0.82
Q2	0.67
Q3	0.72
Q4	0.89

The current sampling design was used for more than twenty years ago and has not been revisited/redesigned. The redesigning of the survey was done to reflect the current behavior of corn production in each province and is also intended to improve the CPS by taking into account the current behavior of corn production in the country, through the use of an updated sampling frame, enhancing the sampling design, and selecting appropriate estimation procedures.

2   Objectives

The redesigning of CPS aimed to come up with a better sampling designs that will generate reliable corn production estimates at the provincial level. Specifically, this paper aims to perform simulation to determine the best sample size (barangays and households), the best sampling design; and the best estimation procedure.

3   Terminology

1. **Household** - a person or a group of persons who sleep in the same dwelling unit and usually have a common arrangement in the preparation and consumption of food. The household members may not necessarily be related by ties of kinship, although they are usually relatives. In some instances, more than one household may occupy the same dwelling unit.

2. **Farming household** - any household in which a member operates an agricultural land, either solely or jointly with other members and the aggregate area operated by the operator-members of such household qualifies to be a farm.

3. **Non-farming household** - any household in which a member operates an agricultural land, either solely or jointly with other members but the aggregate area operated by the operator-members of such household does not qualify as a farm.

4. **Non-agricultural household** - any household in which none of the members operates an agricultural land.

5. **Operator** - a person who takes the technical, financial and administrative responsibility in managing the farm, including the management and supervision of hired labor; he may work on the land himself or may employ others to work on the land. He may or may not be the owner of the land.

6. **Farm** - a parcel or parcels of land which has a total land area of at least 1,000 square meters (one-tenth of a hectare) used for agricultural purposes.

7. **Parcel** - a piece of agricultural land which meets any of the characteristics arranged in order of importance:

    (a) contiguous area with natural boundaries;

    (b) under only one tenurial status; and

    (c) regarded as such by the farmer.

4 Methodology

4.1 Data Preparation

The 2013 CPS and the 2012 CAF data are used in this study. For the 2013 CPS, only data from completed survey of corn farming households were used. The data were filtered based on two items in the questionnaire: (1) (Item B2) Status of respondents: 10 = Corn Household and (2) (Item B21) Result of Visit: 40 = Survey Completed. The data which satisfied these two conditions were included in the study. The data were then inspected for completeness of required variables (total farm area, corn production and harvesting status). For the 2012 CAF, only the first semester data with the total farm area information were used.

4.2 Phase 1: Determination of Initial Number of Sample Barangays

The first phase of this study involved determining the initial optimum number of sample barangays. It uses the Lavalle-Hidiroglou (LH) algorithm[5]. The LH algorithm is an iterative procedure that allocate samples into strata. It obtains stratification boundaries that will give minimum variance. A stratification variable that is highly correlated with the variable of interest, required CV and the number of strata have to be specified in the algorithm.

Different combinations of number of strata (3, 5 and 10 strata) and coefficient of variations (1%, 5% and 8%) were explored to compute for the initial number of sample barangays.

4.3 Phase 2: Determination of Best Sampling Design

Three sampling designs were considered for simulation: probability proportional to size-systematic (PPS-SYS), stratification using iterative method and stratification using LH algorithm. The actual number of barangays per province was first computed as follows:

1. Solve for n'

$$n' = \begin{cases} 2n_0 & \text{if } n_0 \text{ is even} \\ 2(n_0 + 1) & \text{if } n_0 \text{ is odd} \end{cases}$$

   where n_0 is the initial sample size chosen in Phase 1.

2. Divide the sample barangays into four (4) distinct groups to obtain the number of sample barangays per quarter (n). This will give result to independent sample for each quarter.

4.3.1 Probability Proportional to Size-Systematic (PPS-Systematic)

For this sampling design, corn farm area was used as the measure of size, hence, larger farm areas have bigger probability of being included. Barangays were selected using PPS-SYS and households were selected using either SRSWOR or systematic. Below is the procedure:

1. Assign 1 to X_i natural numbers to U_1, $X_1 + 1$ to $X_1 + X_2$ to U_2, etc.

2. Compute for the sampling interval, $K = \frac{X}{a}$, where X is the total corn farm area of the province and a is the number of sample barangay in a province.

3. Choose a random number between 1 and X, where X is the total corn farm area.

4. Unit whose range includes the random number is included in the sample.

5. Repeat the process until the required sample size is achieved.

Weight Estimation

The design-based weight was computed as:

$$W_i = W_{ij} = \frac{\sum_{i=1}^{A} X_i}{aX_i} \times \frac{N_i}{n_i}$$

where: w_i = weight for all household in barangay i, w_{ij} = weight for all household j in barangay i by quarter, X_i = corn farm area in barangay i by quarter, A = total number of barangays per quarter, a = number of sample barangays per quarter, N_i = number of corn farm households in barangay i by quarter, and n_i = number of sample corn farm households in barangay i by quarter.

Estimation of Total Production

The total corn production was computed as:

$$\hat{Y} = \sum_{i=1}^{a} \sum_{j=1}^{n_i} w_{ij} y_{ij}$$

where: $\hat{Y}$ = sum of weighted total corn production in the province, w_{ij} = weight for all household in barangay i by quarter, y_{ij} = corn production of household j in barangay i per quarter, a and n_i as defined above.

Estimation of Variance

The variance was computed as follows:

$$V(\hat{Y}) = \left(1 - \frac{a}{A}\right) a s_1^2 + \frac{a}{A} \sum_{i=1}^{n_i} \left(1 - \frac{n_i}{N_i}\right) n_i s_2^2$$

where: $s_1^2 = \frac{\sum_{i=1}^{a_h} (y_{hi} - \bar{y}_h)^2}{a_h - 1}$, $s_2^2 = \frac{\sum_{i=1}^{n_{hi}} (y_{hij} - \bar{y}_{hi})^2}{n_{hi} - 1}$, s_1^2 = first stage sample variance (provincial sample variance), s_2^2 = second stage sample variance (barangay sample variance), a_h = total number of stratum barangays in stratum h, y_{hi} = weighted total corn production of barangay i in stratum h, y_{hij} = weighted total corn production for household j, barangay i in stratum h, $\bar{y}_h$ = mean corn production in stratum h, and $\bar{y}_{hi}$ = mean corn production in barangay i, stratum h, n_{hi} = total number of sample corn farm households in stratum h, barangay i and a, A, n_i, N_i as defined previously.

4.3.2 Stratified using Iterative Method

In this design, corn farm area was used as the stratification variable. This method is also known as Dalenus & Hodges (1959) $cum\sqrt{f}$ rule which is based on constructing equal intervals on the cumulative of the square roots of the frequencies of the stratification variables[6]. Proportional allocation was employed

and stratum boundary was determined using the iterative method.

$$y_h = \frac{\bar{y}_h + \bar{y}_{h+1}}{2}$$

Weight Estimation

Case 1: PPS selection of barangays, and SRSWOR and systematic selection of households.

$$W_{hi} = W_{hij} = \frac{\sum_{i=1}^{A_h} X_{hi}}{a_h X_{hi}} \times \frac{N_{hi}}{n_{hi}}$$

where: w_{hi}= weight for all household in barangay i at stratum h, X_{hi} = corn farm area in barangay i in stratum h, A_h = total number of barangays in stratum h, a_h = total number of sample barangays in stratum h, N_{hi} = total number of corn farm households in stratum h, barangay i, n_{hi} = total number of sample corn farm households in stratum h and barangay i.

Case 2: Selection of Barangays - SYS and selection of Households - SRSWOR & SYS

$$W_{hi} = \frac{A_h}{a_h} \times \frac{N_{hi}}{n_{hi}}$$

where: A_h, a_h, N_{hi} and n_{hi} as defined above.

Estimation of Total Production

The total corn production for stratified using iterative method was computed as:

$$\hat{Y} = \sum_{h=1}^{L} \hat{Y}_h, \quad \hat{Y}_h = \sum_{i=1}^{a_h} \sum_{j=1}^{n_{hi}} w_{hij} y_{hij}$$

where: $\hat{Y}$= sum of weighted total corn production in all strata, $\hat{Y}_h$= weighted total corn production in stratum h, w_{hij}= weights for household j in barangay i, stratum h, and y_{hij}= corn production in household j, barangay i, stratum h.

Estimation of Variance

The variance is computed as:

$$V(\hat{Y}_h) = \left(1 - \frac{a_h}{A_h}\right) a_h s_{h1}^2 + \frac{a_h}{A_h} \sum_{i=1}^{n_{hi}} \left(1 - \frac{n_{hi}}{N_{hi}}\right) n_{hi} s_{h2}^2$$

where: $s_{h1}^2 = \frac{\sum_{i=1}^{a_{hi}} (y_{hi} - \bar{y}_h)^2}{a_h - 1}$, $s_{h2}^2 = \frac{\sum_{i=1}^{n_{hi}} (y_{hij} - \bar{y}_{hi})^2}{n_{hi} - 1}$, a_h = total number of sample barangays in stratum h, A_h= total number of barangays in stratum h, s_1^2= first stage sample variance (provincial sample variance), s_2^2= second stage sample variance (barangay sample variance), y_{hi}= weighted total corn production of

barangay i in stratum h, $y_{hij}=$ weighted total corn production for household j, barangay i in stratum h, $\bar{y}_h=$ mean corn production in stratum h, and $\bar{y}_{hi}=$ mean corn production in barangay i, stratum h.

4.3.3   Stratified using LH Algorithm

The third proposed sampling design is stratified random sampling using LH algorithm. As mentioned in Phase 1, this method is appropriate for skewed population, hence, appropriate for the skewed corn farm areas and corn production as shown in Figures 1 and 2. The LH algorithm provides optimum stratum boundaries for farm areas of barangays. The general idea of the algorithm is to minimize the sample size with respect to the boundary of each stratum given the CV. Five (5) percent CV and five (5) strata are considered since these are the values yielding optimal sample size in Phase 1.

Weight Estimation

Case 1: PPS selection of barangay, and SRSWOR and systematic selection of households.

$$W_{hi} = W_{hij} = \frac{\sum_{i=1}^{A_h} X_{hi}}{a_h X_{hi}} \times \frac{N_{hi}}{n_{hi}}, \quad h = 1, \cdots, L$$

where: $w_{hi}=$ weight for all household in barangay i at stratum h, $w_{hij}=$ weights for household j in barangay i, stratum h, $A_h=$ total number of barangays in stratum h, $a_h =$ total number of sample barangays in stratum h, $X_{hi}=$ corn farm area in barangay i, stratum h, $N_{hi}=$ total number of corn farm households in stratum h, barangay i and $n_{hi}=$ total number of sample corn farm households in barangay i, stratum h.

Case 2: Systematic selection of barangays, and SRSWOR and systematic selection of households

$$W_{hi} = \frac{A_h}{a_h} \times \frac{N_{hi}}{n_{hi}} \quad h = 1, \cdots, L$$

where: w_{hi}, A_h, a_h, N_{hi} and n_{hi} as defined above.

Estimation of Total Production

The total corn production is computed as:

$$\hat{Y} = \sum_{h=1}^{L} \hat{Y}_h + \hat{Y}_1, \quad \hat{Y}_h = \sum_{i=1}^{a_h} \sum_{j=1}^{n_{hi}} w_{hij} y_{hij}$$

where: $\hat{Y}_1 = \left(\dfrac{\% \ share \ of \ stratum \ 1 \ to \ X}{1 - \% \ share \ of \ stratum \ \% \ 1 \ to \ X} \right) \sum_{h=1}^{L} \hat{Y}_h$, $\hat{Y}=$ sum of total corn production in all strata, $\hat{Y}_1=$ total production for the take none stratum, $\hat{Y}_h=$

weighted total production of stratum h, $X=$ corn farm area, $w_{hij}=$ weight estimated for household j in barangay i stratum h, and $y_{hij}=$ total corn production for household j in barangay i, stratum h.

Estimation of Variance

$$\hat{V}(\hat{Y}) = \sum_{h=1}^{L} \hat{V}(\hat{Y}_h) + \hat{V}(\hat{Y}_0)$$

where: $\hat{Y}_0 = \left(\dfrac{\% \, share \, of \, stratum \, 1 \, to \, X}{1 - \% \, share \, of \, stratum \, \% \, 1 \, to \, X} \right)^2 \sum_{h=1}^{L} \hat{V}(\hat{Y}_h)$

$$\hat{V}(\hat{Y}_h) = \left(1 - \frac{a_h}{A_h} \right) a_h s_{h1}^2 + \frac{a_h}{A_h} \sum_{i=1}^{n_{hi}} \left(1 - \frac{n_{hi}}{N_{hi}} \right) n_{hi} s_{h2}^2, \quad h =, \cdots, L$$

4.3.4 Recreation of Frame for Simulation

Frame for simulation was recreated by initially building a regression model for the non-zero corn production data of 2013 CPS:

$$Y_{ij} = \beta_i X_{ij} + \epsilon_{ij}$$

where: $Y_{ij}=$ corn production of province i for quarter j and $X_{ij}=$ corn farm area of province i for quarter j.

The regression model was used to predict corn production of each province per quarter in the 2012 CAF. Also, the percentage of zero production (corn farming households not harvesting) in the 2013 CPS was computed by province and by quarter and the same percentage of households was set to zero in the recreated frame. This recreated frame served as the sampling frame for simulation.

Simulation Scenarios: A summary of sampling design/stratification, selection methods for barangay and households, and number of sample households considered for simulation is presented below. A total of 30 scenarios per province per quarter with 10,000 replicates (100 replicates for barangays and 100 replicates for households) were used for simulations.

1. Design/Stratification Method	PPS-SYS; Iterative; LH
2. Selection of Barangays	PPS; SYS (Except PPS-SYS)
3. Selection of Households	SRSWOR; SYS
	10
	15
	20
4. Stratum	5

9

4.3.5 Performance Measures

Results of the simulation are compared in terms of accuracy and precision.

Mean Absolute Percentage Error (MAPE)

$$MAPE = \frac{\sum_{i=1}^{S} \left| \frac{Y_i - \hat{Y}_i}{Y_i} \right|}{S} \times 100\%$$

where: Y_i= actual corn production value, $\hat{Y}_i$= predicted corn production and S= total number of simulation datasets.

Average Coefficient of Variation (CV)

$$CV = \frac{\sum_{i=1}^{S} \left| \frac{\sqrt{\hat{V}(\hat{Y}_i)}}{\hat{Y}_i} \right|}{S} \times 100\%$$

4.4 Phase 3: Selection of Best Estimation Procedure

Bootstrap method is a resampling method that provides approximate variance which is useful when exact variance is hard to derive due to design complexity[7]. The general procedure of bootstrap consists of taking large number of independent samples, estimating the parameter of interest, and estimating the variance of the estimate. Two modifications in the bootstrap method were explored for the selection of best estimation procedure. For both procedures, simulation was done using $K = 500$ bootstrap samples.

4.4.1 Modified Bootstrap 1st Stage

In this procedure, resampling is done at the first stage (selection of sample barangays).

Steps:

1. Draw sample barangays of size α from A for each province and for each quarter using the design chosen from Phase 2 of the simulation (e.g., PPS-SYS);

2. Draw sample households from each barangay of size $n = (n_1, n_2, \cdots, n_\alpha)$;

3. Compute for $\hat{\theta} = (\hat{\theta}_1, \hat{\theta}_2, \cdots, \hat{\theta}_\alpha)$ from the original sample where $\hat{\theta}_i = \alpha(\frac{T_i}{\phi_i})$, $T_i = \sum_{j=1}^{n_i} \frac{N_i}{n_i} y_{ij}$ and ϕ_i is the probability of selection of barangay i computed as the reciprocal of the design weight for barangay i. Each $\hat{\theta}_i$ will serve as an estimate of PSU total and $\hat{\theta}$ will serve as the pseudo-population.

4. Generate size K bootstrap sample from $\hat{\theta}$ using SRSWOR.

4.4.2 Modified Bootstrap 2nd Stage

For this method, resampling is done at the second stage (selection of sample corn farm households)

Steps:

1. Same steps as in 1 and 2 of Modified Bootstrap 1st Stage.

2. Let $N^* = (k_1^*)n_1, (k_2^*)n_2, \cdots, (k_\alpha^*)n_\alpha$ where k^* is the vector of rounded-up elements of $k = \frac{N_1}{n_1}, \frac{N_2}{n_2}, \cdots, \frac{N_\alpha}{n_\alpha}$, where $n_1, n_2, \cdots, n_a lpha$ is the number of sample corn farm households in barangays $1, 2, \cdots, \alpha$ while $N_1, N_2, \cdots, N_\alpha$ is the total number of corn farm households in barangays $1, 2, \cdots, \alpha$.

3. Generate U^* by copying each households from each sampled barangays k_α^* times and generate new households IDs for these.

4. Draw $K = 500$ samples of size n households using SRSWOR from each barangay from U^*.

4.4.3 Total Production (Bootstrap 2nd Stage)

The bootstrap estimator of total corn production is computed as:

$$\hat{\theta}_K = \frac{\sum_{k=1}^{K} \hat{\theta}_k}{K}, \quad k = 1, 2, 3, .., K$$

where: $\hat{\theta}_k = \sum_{i=1}^{a} w_i \sum_{j=1}^{n_i} y_{ij}$, w_i=weight of barangay i computed via PPS-SYS and y_{ij}= corn production for household j, barangay i.

4.4.4 Variance (Bootstrap 2nd Stage)

The estimator of variance is the Monte Carlo Variance computed as:

$$V(\hat{\theta}_K) = \frac{\sum_{k=1}^{K} (\hat{\theta}_k - \theta_K)^2}{K - 1}$$

5 Results

5.1 Determination of Initial No. of Sample Barangay

For the first phase of this research study, the initial sample barangays was determined. Corn farm area was used as the stratification variable when using the LH algorithm since it is highly correlated with corn production as presented in Table 1 of the Introduction. Among all combinations of CVs (1%, 5% and 8%) and strata (3, 5, and 10), five (5) strata and 5% CV yielded the optimal sample size. Results revealed that using 10 strata would result to too many zero cells while using three (3) strata on the other hand would result to too large sample size per strata. For the CVs, the sample size would be too small

if 8% CV was used while using 1% CV would result to too large sample size. The computed initial number of barangays is 1,952 as shown Table 2 of the Appendix.

5.2 Determination of Best Sampling Design

For the second phase of this study, the different scenarios for the determination of best sampling design was performed using 5% CV and five (5) strata. Results showed that stratified-LH is the best stratification in terms of precision and accuracy with PPS Systematic as the best sampling selection method for sample barangays, in general. Also, SRSWOR is the best selection method for sample households (Refer to Table 2 of the Appendix). Further, it was observed that accuracy and precision generally increase as the number of sample household increases. Hence, of the three (3) sample sizes considered (10, 15 and 20) households), a sample of 20 households is the best in terms of CVs and MAPEs. Also, the CVs and MAPEs are high for minor producing provinces/HUCs.

5.3 Determination of Best Estimation Procedure

Results for the third phase showed that modified bootstrap 2nd stage has the lowest trimed CVs in all quarters with CVs ranging from 4.8% to 14.3%, while CVs of the modified bootstrap 1st stage and design-based ranged from 5.4% to 24.2% (Figures 3 to 6).
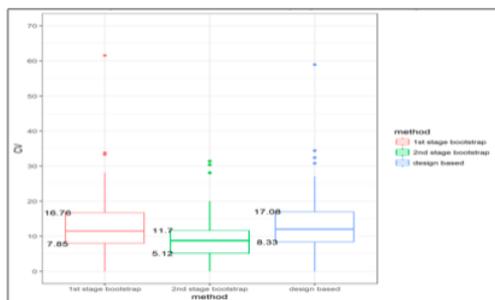

Figure 3. Comparison of CV Distribution for Quarter 1
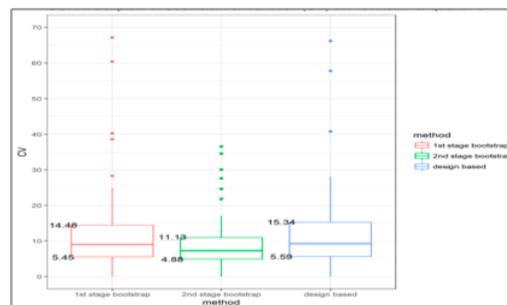
Figure 5. Comparison of CV Distribution for Quarter 3
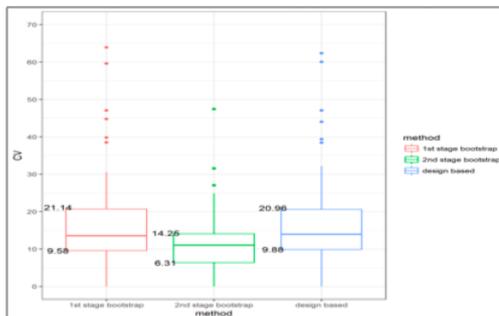
Figure 4. Comparison of CV Distribution for Quarter 2
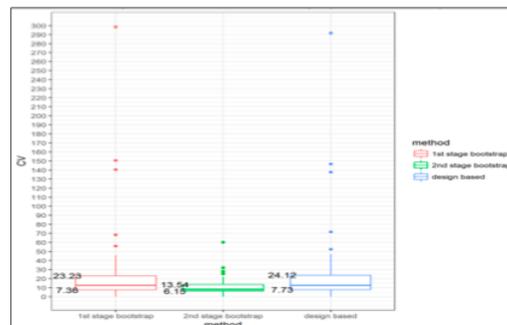
Figure 6. Comparison of CV Distribution for Quarter 4

12

Although, modified bootstrap 2nd stage is the most precise method, it is, however, the least accurate method, registering the highest MAPEs in all quarters. Design-based and modified bootstrap 1st stage are the most accurate methods in all quarters with trimmed MAPEs ranging from 3.4% to 31.0%, while MAPEs of modified bootstrap 2nd stage ranged from 5.5% to 35.8% (See Figures 7 to 10).
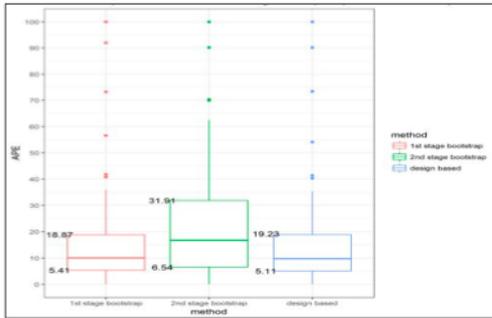
Figure 7. Comparison of MAPE Distribution for Quarter 1
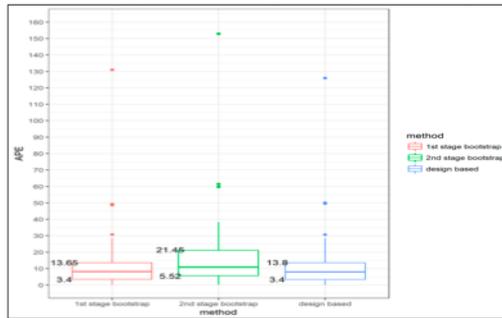
Figure 9. Comparison of MAPE Distribution for Quarter 3
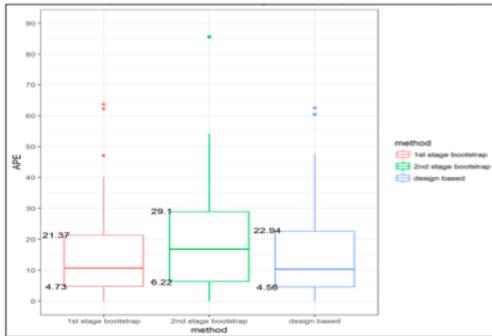
Figure 8. Comparison of MAPE Distribution for Quarter 2
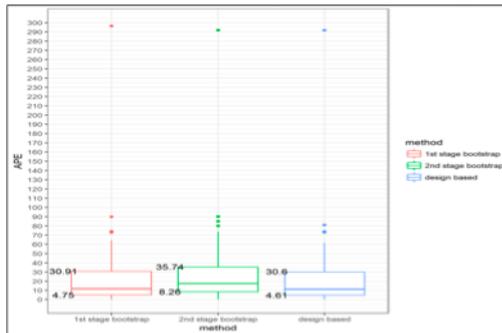
Figure 10. Comparison of MAPE Distribution for Quarter 4

6 Conclusions

Stratified-LH is the best stratification in terms of precision and accuracy with PPS-SYS and SRSWOR as the best selection method for sample barangays and sample households, respectively. Twenty (20) sample households seem to be optimal. Design-based estimation and modified bootstrap 1st stage are comparable in terms of accuracy and are both better than modified bootstrap 2nd stage in terms of accuracy.

7 Recommendations

Although, the best sampling design based on simulation results, is stratified-LH with PPS-SYS as the barangay selection method and SRSWOR as the sample household selection method, PPS-SYS with corn farm area as the measure of size is recommended over LH since the CVs of PPS-SYS and LH are very close and the estimation procedure for PPS-SYS is easier to implement. It

follows that the selection method for barangays is systematic. SRSWOR is the recommended selection method for sample households.

Design-based estimation is recommended over bootstrap since it is simpler and yields more accurate results. With this design, some provinces will be surveyed in some quarters while those provinces not surveyed for a certain quarter will be modeled for that particular quarter.

8

References

[1] countrystat.psa.gov.ph.

[2] *Compilation of Sampling Methodologies used in Agriculture and Fisheries Survey.* Survey Designs Research and Development Section, Philippine Statistics Authority, March 2015

[3] 2013 CPS, Philippine Statistics Authority.

[4] 2012 CAF, Philippine Statistics Authority.

[5] Lavallee, P. and Hidiroglou, M.A. (1988). On the stratification of skewed populations. Survey Methodology, 14, 33-43

[6] Dalenius, T. and Hodges, J. L. (1959), Minimum Variance Stratification, Journal of the American Statistical Association, 54, 88-101.

[7] Chernick, M.R. (1999). *Bootstrap Methods: A Practitioner's Guide.* New York: John Wiley & Sons, Inc.

9 Appendix

Table 2: Number of Farming Barangays and Initial Sample Barangays (CV=5%, No. of strata=5)

	n	a			n	a
REGION 1				**REGION 8**		
Ilocos Norte	418	31		Biliran	72	9
Ilocos Sur	615	36		Eastern Samar	132	14
La Union	408	34		Leyte	769	39
Pangasinan	936	40		Northern Samar	188	15
REGION2				Western Samar	620	40
Cagayan	615	37		Southern Leyte	71	11
Isabela	961	39		Tacloban City	37	7
Nueva Vizcaya	211	20		**REGION 9**		
Quirino	129	15		Zambonga City	76	12
REGION 3				Zamboanga del Norte	625	37
Angeles City	7	NA		Zamboanga del Sur	652	30
Aurora	100	10		Zamboanga Sibugay	300	29
Bataan	83	8		**REGION 10**		
Bulacan	180	15		Bukidnon	464	33
Nueva Ecija	353	28		Cagayan de Oro City	44	10
Olongapo City	6	NA		Camiguin	42	8
Pampanga	229	21		Iligan City	40	8
Tarlac	292	26		Lanao del Norte	425	32
Zambales	114	15		Misamis Occidental	343	29
REGION 4A				Misamis Oriental	364	29
Batangas	463	30		**REGION 11**		
Cavite	214	20		Compostela Valley	222	25
Laguna	283	23		Davao City	137	12
Lucena City	21	5		Daval del Norte	182	16
Quezon	575	28		Davao del Sur	190	15
Rizal	116	18		Davao Occidental	95	13
REGION 4B	961	39		Davao Oriental	127	18
Marinduque	440	33		**REGION 12**		
Occidental Mindoro	123	12		General Santos City	26	8
Oriental Mindoro	687	37		North Cotabato	514	35
Palawan	142	17		Sarangani	135	22
Puerto Princesa City	516	36		South Cotabato	194	18
Romblon	168	16		Sultan Kudarat	237	28

	n	a		n	a
REGION 5			**CAR**		
Albay	123	16	Abra	201	22
Camarines Norte	127	20	Apayao	119	18
Camarines Sur	180	18	Baguio City	35	8
Catanduanes	200	20	Benguet	97	14
Masbate	9	6	Ifugao	130	11
Sorsogon	143	15	Kalinga	111	16
REGION 6			Mountain Province	121	10
Aklan	147	9	**ARMM**		
Antique	365	30	Basilan	65	8
Bacolod City	20	6	Lanao del Sur	995	45
Capiz	259	26	Maguindanao	468	33
Guimaras	76	14	Sulu	300	26
Iloilo	1070	51	Tawi-Tawi	70	10
Iloilo City	28	8	**CARAGA**		
Negros Occidental	395	28	Agusan del Norte	114	16
REGION 7			Agusan del Sur	283	27
Bohol	722	43	Butuan City	61	11
Cebu	987	39	Dinagat Islands	36	8
Cebu City	38	7	Surigao del Norte	56	6
Lapu-Lapu City	8	NA	Surigao del Sur	161	23
Negros Oriental	3	NA			
Siquijor	494	32	Total	25,545	1,952
Cebu City	38	7	Surigao del Norte	56	6
Lapu-Lapu City	8	NA	Surigao del Sur	161	23
Negros Oriental	3	NA			

Table 3: Best Sampling Design Per Province

PROVICE	BEST DESIGN	BEST SELECTION METHOD (BRGY)	BEST SELECTION METHOD (HH)	BEST SAMPLE SIZE (HH)	CV				APE			
					Qtr 1	Qtr 2	Qtr 3	Qtr 4	Qtr 1	Qtr 2	Qtr 3	Qtr 4
ABRA	LH	PPS	SRSWOR	20	12.06	5.96	7.30	5.38	10.87	8.36	7.83	6.81
AGUSAN DEL NORTE	LH	PPS	SRSWOR	20	15.98	17.19	26.73	5.91	17.08	14.85	29.85	7.52
AGUSAN DEL SUR	PPS-Sys	PPS	Systematic	20	16.59	17.44	11.67	6.59	12.78	13.86	10.56	8.79
AKLAN	LH	PPS	SRSWOR	20	5.04	20.45	9.87	15.66	6.33	18.42	9.08	14.98
ALBAY	LH	PPS	SRSWOR	20	19.16	5.94	6.25	16.83	17.70	6.24	6.24	13.56
ANGELES CITY	Iterative	PPS	SRSWOR	10	6.71	38.97	1.04	-	4.84	59.50	5.32	-
ANTIQUE	LH	PPS	SRSWOR	20	1.56	20.82	8.10	15.33	2.02	20.38	6.69	13.28
APAYAO	PPS-Sys	PPS	SRSWOR	20	15.67	6.09	8.33	5.31	15.43	8.89	11.54	8.31
AURORA	LH	PPS	Systematic	20	7.03	24.61	5.64	57.66	10.27	25.89	9.85	77.06
BAGUIO CITY	LH	PPS	SRSWOR	10	8.02	12.33	4.00	59.31	4.61	15.38	3.85	71.45
BASILAN	LH	PPS	SRSWOR	20	38.12	9.98	10.97	10.17	32.18	9.59	9.60	8.21
BATAAN	LH	PPS	SRSWOR	20	7.06	28.97	3.94	68.15	8.75	30.25	3.98	59.73
BATANGAS	LH	PPS	SRSWOR	20	9.62	21.13	3.99	20.09	9.60	16.17	6.86	17.78
BENGUET	LH	PPS	SRSWOR	20	15.30	15.70	11.51	9.34	12.06	12.93	8.17	24.66
BILIRAN	LH	PPS	Systematic	20	16.48	17.57	12.09	6.94	15.48	16.99	12.08	10.13
BOHOL	LH	PPS	SRSWOR	20	12.36	10.27	3.76	13.02	11.66	10.14	5.74	11.89
BUKIDNON	PPS-Sys	PPS	Systematic	20	12.03	16.87	7.25	11.93	10.57	14.04	7.82	10.63
BULACAN	LH	PPS	SRSWOR	20	5.15	12.21	4.11	48.58	5.01	10.71	4.30	44.66
BUTUAN CITY	LH	PPS	SRSWOR	20	10.37	19.74	11.36	5.46	11.98	19.18	11.06	6.02
CAGAYAN DE ORO	LH	PPS	SRSWOR	20	14.16	8.42	7.23	8.99	15.28	11.14	9.17	11.88
CAGAYAN	PPS-Sys	PPS	Systematic	20	6.74	7.80	6.67	14.56	6.12	7.33	6.57	12.68
CAMARINES NORTE	LH	PPS	SRSWOR	20	19.67	6.68	3.50	17.66	15.10	3.91	5.39	14.29

PROVICE	BEST DESIGN	BEST SELECTION METHOD (BRGY)	BEST SELECTION METHOD (HH)	BEST SAMPLE SIZE (HH)	CV				APE			
					Qtr 1	Qtr 2	Qtr 3	Qtr 4	Qtr 1	Qtr 2	Qtr 3	Qtr 4
CAMARINES NORTE	LH	PPS	SRSWOR	20	19.67	6.68	3.50	17.66	15.10	3.91	5.39	14.29
CAMARINES SUR	LH	PPS	Systematic	20	13.04	13.12	7.14	24.21	14.59	13.74	13.96	20.96
CAMIGUIN	LH	PPS	SRSWOR	20	11.72	10.15	11.95	12.18	11.13	9.70	10.75	10.63
CAPIZ	Iterative	PPS	SRSWOR	20	3.63	17.53	4.44	12.44	7.38	16.59	8.23	12.03
CATANDUANES	Iterative	PPS	SRSWOR	15	27.29	11.27	5.50	28.26	25.22	5.46	6.08	19.27
CAVITE	LH	PPS	SRSWOR	20	10.96	35.50	3.74	23.81	9.80	33.17	4.06	19.36
CEBU CITY	LH	PPS	SRSWOR	20	19.03	19.28	19.37	20.60	17.79	19.41	19.57	19.40
CEBU	PPS-Sys	PPS	Systematic	20	10.92	11.74	3.90	11.43	9.40	11.39	6.37	10.89
COMPOSTELA VALLEY	PPS-Sys	PPS	Systematic	20	15.29	13.31	10.25	8.56	14.02	11.90	10.18	8.45
DAVAO CITY	LH	PPS	SRSWOR	20	11.13	9.45	14.05	14.10	12.62	11.09	14.36	14.32
DAVAO DEL NORTE	LH	PPS	Systematic	20	9.04	12.81	7.41	5.56	9.78	11.56	6.89	7.56
DAVAO DEL SUR	LH	PPS	SRSWOR	20	13.93	7.81	11.33	8.32	14.22	10.64	12.48	10.80
DAVAO OCCIDENTAL	LH	PPS	Systematic	20	11.50	7.25	12.12	6.77	10.66	6.72	8.73	5.72
DAVAO ORIENTAL	LH	PPS	Systematic	20	11.12	10.56	7.37	8.59	10.43	9.35	8.03	9.02
DINAGAT ISLANDS	LH	PPS	SRSWOR	10	6.81	17.45	58.30	7.02	45.81	16.32	75.14	0.02
EASTERN SAMAR	LH	PPS	SRSWOR	20	24.61	27.20	0.06	27.92	22.55	25.45	0.58	26.84
GENERAL SANTOS	LH	PPS	SRSWOR	20	12.19	16.09	13.41	9.63	15.53	18.30	16.94	12.47
GUIMARAS	LH	PPS	SRSWOR	20	4.36	17.51	8.56	17.41	4.88	24.64	7.04	12.74
IFUGAO	PPS-Sys	PPS	Systematic	20	3.21	3.48	3.33	3.51	2.72	3.22	2.92	3.12
ILIGAN CITY	LH	PPS	SRSWOR	20	10.01	15.41	10.66	10.53	12.72	16.76	13.15	12.58
ILOCOS NORTE	LH	PPS	Systematic	20	6.94	12.71	6.75	6.65	9.98	13.46	10.69	10.75
ILOCOS SUR	LH	PPS	SRSWOR	20	7.35	9.96	7.37	43.81	8.49	9.78	8.49	36.07
ILOILO CITY	Iterative	PPS	SRSWOR	10	4.95	10.89	11.14	6.23	5.31	17.99	10.22	9.37

PROVICE	BEST DESIGN	BEST SELECTION METHOD (BRGY)	BEST SELECTION METHOD (HH)	BEST SAMPLE SIZE (HH)	CV				APE			
					Qtr 1	Qtr 2	Qtr 3	Qtr 4	Qtr 1	Qtr 2	Qtr 3	Qtr 4
ILOILO	LH	PPS	SRSWOR	20	4.40	4.40	5.37	4.20	5.70	5.78	6.62	5.80
ISABELA	PPS-Sys	PPS	Systematic	20	3.90	19.91	21.02	23.42	6.40	16.96	18.05	21.21
KALINGA	LH	PPS	Systematic	20	5.96	6.98	7.12	6.74	9.58	13.05	9.90	10.77
LA UNION	LH	PPS	SRSWOR	20	7.45	7.36	7.75	7.44	7.92	8.21	7.74	8.32
LAGUNA	Iterative	PPS	SRSWOR	20	13.07	71.93	4.01	28.65	10.71	70.56	3.82	26.41
LANAO DEL NORTE	Iterative	PPS	SRSWOR	20	6.07	10.18	6.53	6.52	7.28	9.98	7.52	7.35
LANAO DEL SUR	PPS-Sys	PPS	Systematic	20	11.29	7.48	7.86	8.84	9.00	6.05	6.92	7.75
LEYTE	LH	PPS	SRSWOR	20	10.35	10.82	3.77	11.23	10.26	10.96	6.18	11.26
LUCENA CITY	LH	PPS	SRSWOR	15	27.70	17.64	18.36	6.87	3.20	23.45	11.76	3.75
MAGUINDANAO	PPS-Sys	PPS	Systematic	20	10.48	12.38	9.33	10.28	9.10	10.75	7.60	9.00
MARINDUQUE	LH	PPS	SRSWOR	20	13.56	27.10	7.23	19.70	12.23	24.27	7.92	17.58
MASBATE	Iterative	PPS	Systematic	20	6.21	16.35	4.89	21.31	9.53	17.19	7.45	18.61
MINDORO OCCIDENTAL	LH	PPS	Systematic	20	10.48	38.23	7.48	21.69	14.06	36.17	12.45	17.89
MINDORO ORIENTAL	LH	PPS	SRSWOR	20	21.49	31.61	10.26	27.15	17.78	26.68	12.91	19.13
MISAMIS OCCIDENTAL	LH	PPS	SRSWOR	20	10.16	12.16	4.42	6.04	10.35	11.44	6.71	7.32
MISAMIS ORIENTAL	PPS-Sys	PPS	SRSWOR	20	14.35	7.47	7.77	8.59	15.13	12.22	11.63	11.43
MT. PROVINCE	LH	PPS	SRSWOR	20	21.15	22.56	9.40	23.60	19.72	20.16	12.44	22.31
NEGROS OCCIDENTAL	PPS-Sys	PPS	SRSWOR	10	23.50	37.98	0.00	26.84	21.17	29.04	3.17	22.42
NEGROS ORIENTAL	PPS-Sys	PPS	Systematic	20	8.70	16.91	4.53	15.19	7.40	15.30	4.41	12.92
NORTH COTABATO	PPS-Sys	PPS	Systematic	20	10.55	7.56	8.61	6.89	9.53	6.63	7.81	6.28
NORTHERN SAMAR	Iterative	PPS	SRSWOR	20	65.07	18.78	3.58	15.34	90.69	14.28	2.88	17.49
NUEVA ECIJA	LH	PPS	SRSWOR	20	5.50	21.46	3.30	50.23	5.90	18.69	4.86	58.30
NUEVA VIZCAYA	LH	PPS	Systematic	20	5.00	4.59	5.59	4.71	7.83	8.05	7.88	8.83
PALAWAN	LH	PPS	SRSWOR	20	14.19	28.08	17.09	19.78	10.85	31.42	15.07	17.24

PROVICE	BEST DESIGN	BEST SELECTION METHOD (BRGY)	BEST SELECTION METHOD (HH)	BEST SAMPLE SIZE (HH)	CV				APE			
					Qtr 1	Qtr 2	Qtr 3	Qtr 4	Qtr 1	Qtr 2	Qtr 3	Qtr 4
PAMPANGA	PPS-Sys	PPS	SRSWOR	10	11.37	49.70	10.62	55.16	9.35	42.56	9.47	78.32
PANGASINAN	PPS-Sys	PPS	Systematic	20	3.73	35.65	21.23	3.57	7.60	31.51	19.94	7.23
PUERTO PRINCESA CITY	LH	PPS	SRSWOR	20	7.25	15.18	6.64	6.17	22.00	44.10	33.27	14.85
QUEZON	LH	PPS	SRSWOR	20	10.82	14.05	10.89	13.17	12.29	13.04	11.20	13.01
QUIRINO	PPS-Sys	PPS	Systematic	20	6.94	12.30	7.50	12.02	8.01	11.17	6.87	10.42
RIZAL	LH	PPS	SRSWOR	20	16.75	49.18	5.95	28.59	15.58	45.26	6.47	24.80
ROMBLON	LH	PPS	SRSWOR	20	16.02	35.16	9.11	21.05	17.20	36.13	11.81	21.32
SARANGANI	LH	PPS	Systematic	20	10.74	8.78	9.71	10.00	9.32	8.18	8.83	8.75
SIQUIJOR	LH	PPS	Systematic	20	10.08	21.13	5.61	6.88	9.88	17.34	9.21	9.79
SORSOGON	LH	PPS	SRSWOR	20	26.02	18.25	6.72	21.28	24.81	18.35	6.04	20.76
SOUTH COTABATO	LH	PPS	SRSWOR	20	9.95	12.81	10.22	7.18	12.37	14.44	12.49	10.98
SOUTHERN LEYTE	LH	PPS	SRSWOR	10	28.74	39.37	2.41	29.26	23.07	21.88	1.02	29.45
SULTAN KUDARAT	PPS-Sys	PPS	Systematic	20	4.96	11.51	9.66	10.10	4.66	9.64	7.96	8.27
SULU	Iterative	PPS	Systematic	20	28.38	23.24	11.44	12.50	28.74	22.98	10.13	12.46
SURIGAO DEL NORTE	LH	PPS	SRSWOR	20	19.90	21.87	31.23	3.77	20.77	21.59	28.97	3.97
SURIGAO DEL SUR	LH	PPS	Systematic	20	16.62	17.35	25.97	9.21	15.03	15.72	25.26	9.69
TACLOBAN CITY	LH	PPS	SRSWOR	20	10.46	10.60	25.39	23.44	4.26	9.02	28.17	31.90
TARLAC	LH	PPS	Systematic	20	5.06	10.62	11.31	4.90	9.71	12.46	14.48	10.43
TAWI-TAWI	LH	PPS	SRSWOR	20	31.10	10.72	13.76	13.49	32.11	9.71	11.31	13.01
WESTERN SAMAR	LH	PPS	SRSWOR	20	27.88	12.40	3.43	36.39	26.03	11.62	4.37	33.16
ZAMBALES	LH	PPS	SRSWOR	20	6.89	25.21	2.59	44.62	6.21	23.19	3.39	69.38
ZAMBOANGA CITY	LH	PPS	SRSWOR	20	13.65	13.56	7.90	13.85	14.17	13.51	9.58	13.63
ZAMBOANGA DEL NORTE	LH	PPS	Systematic	20	11.70	9.67	4.66	6.28	10.69	9.54	7.68	8.49
ZAMBOANGA DEL SUR	Iterative	PPS	SRSWOR	20	11.17	19.74	5.44	6.46	11.79	19.68	8.74	9.11
ZAMBOANGA SIBUGAY	LH	PPS	Systematic	20	11.99	12.82	5.55	16.00	12.32	12.16	9.79	14.35