USES OF CENSUS RESULTS, MAPS AND LISTS IN FOOD CONSUMPTION STUDIES IN THE PHILIPPINES

By
BURTON T. OÑATE**

Two of the most important components of the level of living are food and nutrition. There are many indicators of these components, some of which are national or regional averages or per capita food supplies in terms of calories and other nutrients at the retail level or national or regional averages or per capita food supplies in terms of total cereals and other food groups. Each indicator of a given component of the level of living is compared to the corresponding recommended allowance or "standard" of living. These concept were described recently by the author in connection with the uses of these indicators in the circular process of planning, implementation and evaluation of developmental growth [1, 1964]. Individually, the indicator will tell about the state of food and nutrition in the country while the direction of and the interactions between these indicators will inform us about the nature of food and nutrition of the country.

Since these components and indicators are generally products of the statistical system, it will be of interest to show how these statistical products are derived by the Philippine statistical system. Special emphasis will be given to the use of census information, maps and list in the derivation of these indicators as applied to the Food Balance Sheet (FBS) and to Household Food Consumption Surveys (HFCS) in the Philip-

*Paper presented to the participants of the Training Center on Food Consumption Surveys conducted in Manila under the auspices of the Republic of the Philippines and the Food and Agriculture Organization (FAO) of the United Nations. October 29, 1964.

**Statistician. The International Rice Research Instiute (IRRI) Los Baños, Laguna, Philippines.

pines. This set of information will be discussed from the point of view of their uses in the following areas:

1) preparaton of sampling frame for surveys and enumeration frame for censuses,
2) preparation and execution of the sampling procedures, and
3) changes in the estimation procedure in order to establish higher statistical precision.

It must be emphasized at this point that the availability of accurate maps and pertinent lists will also add to the overall efficiency of the survey or census operations. This paper will give emphasis on the uses of these information on the statistical framework.

Note that generally, the estimate of a food commodity obtained from the FBS emanates primarily from the farm. The path of a particular food commodity is traced as it leaves the farm. During this journey, it may be converted into a new product by industry. Losses are incurred until the product reaches the consuming household. The stock and trade also come into the picture. The final estimate appearing in the FBS consists of many components, many of which are obtained beyond the scope of probability sampling. In this context, the preparation of the FBS is similar to the preparation of the country's national income and product accounts. Because of many interactions in the flow of people, trade and funds within a country, it becomes extremely difficult to prepare regional or provincial FBS.

In the preparation of the FBS, the consuming household is generally not involved. However, in household food consumption surveys (HFCS), the household is the sampling and the responding unit. Thus, the technique of collection of the data and the preparation of the estimate will necessarily involve a statistical framework and this framework will require the use of probability sampling. Also, estimates can be made on the provincial or regional levels depending on the statistical procedures which are followed. The interactions of other indicators of the level of living are obtained from the consuming households. The FBS cannot give these important aspects of the household. These points are some of the major differences in the approaches used in the FBS and HFCS. The limitations of the data from the FBS and the advantages of the information derived from the HFCS are described elsewhere [2, 1962; 3, 1964]. In this paper, the FBS and the HFCS are considered as statistical frameworks for food consumption studies in the Philippines.

1. **The Food Balance Sheet** (FBS)

Conceptually, the available food supply (AFS) may be expressed as a simple relationship as follows:

$$(AFS) = P + S_t + I - E - S - W - A - M$$

where

AFS	is the available food supply for a crop year (July 1 to June 30),
P	is the production for the crop year,
S_t	is the available stock at the beginning of the period,
I	is the amount imported during the period,
E	is the amount exported during the period,
S	is the amount used for seed purpose,
W	is the amount that goes to waste or losses,
A	is the amount that goes to animal feeds,

and

M	is the amount allocated to manufacturing and other industries.

In the Philippines, annual estimates of P, S_t, S, W, and A are obtained from the Crop and Livestock Survey (CLS) conducted annually by the Bureau of Agricultural Economics (BAE) of the Department of Agriculture and Natural Resources (DANR). Imports and exports of food commodities are processed by the Bureau of the Census and Statistics (BCS) of the Department of Commerce and Industry (DCI). Other estimates like that on M are obtained from sources. The workings of the Philippine Statistical System were given by the author in a recent publication [4, 1962].

Since five of the most important components of (AFS) are the responsibility of the BAE, the design of the crop and livestock survey (CLS) will be given in some detail. It must be noted that basic data on the production of other food commodities are obtained from other sources. For example, the volume of production of fish is obtained from the Philippine Fisheries Commission. Also, non-farm production of food crops is adjusted using the 1948 and 1960 agriculture censuses and a production index is applied to this base line data. Production of crops not included in the 1948 census is estimated from indications of the nutrition surveys of the seven regions by the Food and Nutrition Research Center (FNRC), NIST, NSDB. Thus, the preparation of the FBS requires considerable use of imputations from various sources, one of the most important is from the results of agricultural censuses [5, 1963; 6, 1962; 7, 1963].

1.1. Crop and Livestock Survey (CLS)

The CLS was started in 1954 under the then Division of Agricultural Econmics, DANR. Since then, the design has undergone changes and this paper will describe the details of the present CLS design and its requirements.

1.1.1. The design of the survey

The 1964 (CLS) design was essentially the same design used in the 1961, 1962, and 1963 surveys. The provinces were treated as separate areas of study with the exception of Batanes, Palawan, and Sulu which were not included in the survey due to administrative difficulties.

Sampling was done in two stages: the survey barrio as the primary sampling unit (PSU) and the farm household as the secondary sampling unit (SSU). The survey barrio which may differ from the revenue barrio includes all farm households located within the confines of the revenue barrio regardless of whether their farms are inside or outside the revenue barrio boundaries. A farm is any parcel or parcels of land with a total area of at least 1,000 square meters devoted to agricultural crops, or any area regardless of size used for the raising of at least 20 heads of livestock or 100 units of poultry. When a landowner has one or more tenants, the land worked by each tenant is a separate farm. A household is a group of persons living together in the same dwelling unit with a common arrangement in the preparation and consumption of of food and whose members may not be necessarily related by ties of kinship. A farm household is one which derives income from operating a farm.

1.1.2. Stratification

Two stratification schemes were employed, namely:

1.1.2.1. Palay density (main strata). Barrios within a province were grouped into the following three strata on the basis of the ratio of palay area to farm area.

I - ratio of 1/2 or greater, or barrios with ratio less than 1/2 but with 400 or more hectares of palay area;
II - ratio of 1/4 or greater but less than 1/2, or where the ratio is less than 1/4 but with 100 or more hectares of palay area; and
III - ratio of less than 1/4 and less than 100 hectares of palay area.

1.1.2.2. Geographical criterion (sub-strata). For administrative convenience and in order to minimize yield differences due to geographical location, contiguous barrios within each main stratum were further grouped into sub-strata of approximately the same absolute area. Each sub-stratum contributed exactly three sample barrios such that the number of sub-strata in a stratum depended upon the number of sample barrios allocated for the province. The poblacion was considered as a barrio in the survey design. Note that the main criterion in the stratification was palay density.

There were about 1,200 sample barrios or a sampling fraction of around (1/22) and the samples were distributed among the provinces on the basis of an optimum allocation scheme wihch is given by

$$n_g = n(N_g\hat{\sigma}_g) \;/\; (\sum_{i=1}^{K} N_i\hat{\sigma}_i)$$

where

n_g = the total number of sample barrios for the g^{th} province;

N_g = the total number of barrios for the g^{th} province;

σ_g = the estimate of the standard deviation of palay production in the g^{th} province;

N_i = the total number of barrios in the i^{th} province (i = 1, 2, ..., K province);

and

σ_i = the estimate of the standard deviation of palay production in the i^{th} province.

Approximately 6,000 sample farm households were drawn as the secondary sampling units or a national sampling fraction of (1/350). A systematic sampling method was employed in identifying these sample farm households. To accomplish this, the enumerator was required to prepare a list of all farm households in the sample barrio during the actual survey operations. The field statistician then furnished him a random start (any number from one to fifteen). He was then allowed to interview the farm household corresponding to that number and every fifteenth farming household from thereon. At least two sample farm households are located per sample barrio [8, 1964; 9, 1964].

1.1.3. The sampling frame

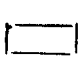
The development of the 1964 sampling frame dates back to as early as 1957. At the Conference of Asian Statisticians held in April, 1957 in Bankok, Thailand, the Philippines reported plans for the improvement of the CLS. Hence, a subproject entitled "Construction of Sample Frames for Future

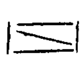
Surveys" under the project "Improvement of Agricultural Statistics" was included in the Philippine Statistical Program for Fiscal Year 1957-58 [10, 1958].

Consequently, maps of all municipalities within each province were secured from provincial and district engineers and from other sources. The barrio layout were sketched and estimates of farm population, farm area, primary and secondary crops were collected. In cases where municipality maps were not available, the Bureau of Agricultural Economics fieldmen together with the municipal agriculturist of the then Bureau of Agricultural Extension (now Commission on Agricultural Productivity, CAP) prepared the maps themselves. Thus, these efforts consisted of updating census results, maps, and lists for the improvement of the sampling frame.

From these information evolved the 1958-60 CLS design. Finally, with modifications in stratification and constant updating of lists of barrios and farm households, the 1961-64 design came into being.

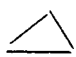
1.1.3.1. Barrio maps. The 1961 to the 1964 CLS's used the same set of sample barrios, but each year, updated sample barrio maps were used. For instance, with the aid of the old 1963 sample barrio map, the field statistician prepared a skeleton map for 1964 showing the boundaries and natural landmarks such as hills, rivers, roads, churches, etc. This was then passed on to the interviewer together with the 1963 map. Using the old map as a guide, the interviewer was required to complete and update the skeleton map by following the same course (indicated in the 1963 map) taken by the 1963 interviewer. The following symbols were used.

 — farm households

 — two farm households in the same dwelling unit.

 — one farm and one non-farm household in the same dwelling unit.

 — two non-farm households in the same dwelling units.

 — non- farm households.

In order to minimize listing errors and field biases, any interviewer who was hired in 1964 was assigned to sample barrios other than those given to him in 1963, if he was a part of the survey team in 1963.

1.1.3.2. Relisting of households. The interviewer, tracing the same course taken by the last year's interviewer, visited and listed all the farm and non-farm households. The field statistician checked this new list by comparing it with the 1963 list. This updated list would differ from last year's list only if there were changes in locations, number and/or status of households. Again, to avoid biases on the part of the interviewer, the old list of households was not shown to him. Any wide discrepancies between the two lists were checked with the interviewer and/or by actual visit of the supervisor to the sample barrio concerned. Emphasis was made on the completeness of the new list. Any under-estimation or over-estimation in the listing operation will substantially bias the estimation procedure

1.1.3.3. Selection of the sample farm households. The same random number assigned to the sample barrio in 1963 was used in identifying the samples in 1964. Thus, if there were no changes in the old and new lists of households, the 1963 and 1964 surveys would have the same sample farm households. Before starting field operations, interviewers usually asked permission from the municipal mayor and the barrio lieutenant and explained the nature and purpose of the survey. In many instances, the barrio lieutenant even acted as guide to the interviewer and introduced him to the houshold members. Information about official changes in the barrio boundaries are obtained from these municipal and barrio officials.

1.1.4. Estimation procedure

The estimation procedures for palay production and hectarage will be discussed in this section. For purpose of illustration, the 1962 CLS data for the province of Laguna will be used. The sample barrios were drawn with probability proportional to size and with complete replacement.

The unbiased estimator of total rice production (X_g) of the g^{th} province is

$$\hat{X}_{.g} = \sum_h \sum_i (1/3) \sum_j (1/P_{ghij}) \sum_k (X_{ghijk}/n_{ghij}) \quad [1.1]$$

$$= \sum_h \sum_i \hat{X}_{ghi}$$

where

X_{ghijk} = the production of the k^{th} sample farm household ($K = 1, 2, \ldots n_{ghij}$) of the j^{th} sample barrio ($j = 1, 2, 3$) of the i^{th} sub-stratum ($i = 1, 2, \ldots, s$) of the h^{th} stratum ($h = I, II, III$) of the g^{th} province (Laguna);

P_{ghij} = a_{ghij}/A_{ghi} = the probability of selection of the j^{th} sample barrio in the ghi^{th} sub-stratum;

a_{ghij} = total area (viz., palay area for stratum I and II and farm area for stratum III) of the $ghij^{th}$ sample barrio (PSU)

A_{ghi} = total area (vis., palay area for stratum I and II and farm area for Stratum III) of the ghi^{th} sub-stratum;

N_{ghij} = total number of farm households in the $ghij^{th}$ PSU;

and

n_{ghij} = total number of sample SSU's in the $ghij^{th}$ PSU.

From the requirements of this estimation procedure, P_{ghij}, a_{ghij}, and A_{ghij} are obtained from the census results and/or enumeration procedure. The value of N_{ghi} is obtained also by the use of an updated map and by following a rigid listing procedure.

If we let

$$\hat{x}_{ghij} = (1/P_{ghij})\,(N_{ghij}) \sum_k x_{ghijk}/n_{ghij} \qquad [1.2]$$

$$= (1/P_{ghij})\,(N_{ghij})\,\bar{x}_j$$

our estimator would simply be

$$\hat{X}_g = (1/3) \sum_h \sum_i \sum_j \hat{x}_{ghij} \qquad [1.3]$$

The estimate of variance of the provincial total ($\hat{X}_g$) on the barrio basis would be

$$\hat{\sigma}^2(\hat{X}_g) = \sum_h \sum_i (1/3) \sum_j \left[(\hat{x}_{ghij} - \hat{x}_{ghi})^2 / 2\right] \qquad [1.4]$$

$$= \sum_h \sum_i (1/3)\,\hat{\sigma}^2\,(\hat{x}_{ghij}$$

$$= \sum_h \sum_i \hat{\sigma}^2\,(\hat{x}_{ghi})$$

where

$\hat{x}_{ghij}$ is as defined in [1.2]

and

$$\hat{x}_{ghi} = (1/3) \sum_j^3 \hat{x}_{ghij}$$

Similar formulas are used for hectarage (Y). The estimation procedure for the province of Laguna is given in Table 1.1. The variance of estimates for each sub-stratum is given in Table 1.2.

TABLE 1.1.

ESTIMATION OF PRODUCTION (X) AND HECTARAGE OF PALAY. JULY 1961 TO JUNE 1962 LAGUNA, PHILIPPINES.

Sample barrio and municipality by sub-strata	1/P	Number of farm households	Average palay production per farm household (sacks of 44 kilos)	Average hectarage per farm household (hectares)	$\hat{X}_{ghij}$ (sacks of 44 kilos)	$\hat{Y}_{ghij}$ (hectares)
I-A					$\hat{X}_{ghi}$ = 215,136[a]	$\hat{Y}_{ghi}$ = 4,766[a]
1. Cuyab, San Pedro	112	76	35.8	0.8	304,730	6,810
2. Kalumpit, Sta. Maria	22	32	297.8	6.7	209,638	4,717
3. Poblacion, Pangil	20	126	52.0	1.1	131,040	2,772
I-B					205,013	4,961
1. Calabuso, Biñan	34	35	273.3	6.3	325,210	7,497
2. Manplasan, Biñan	31	46	145.3	3.8	207,173	5,419
3. Canado, Biñan	41	32	63.0	1.5	82,656	1,968
I-C					672,299	10,993
1. Bubucal, Sta Cruz	85	82	222.0	3.7	1,547,340	25,789
2. San Antonio, Bay	27	43	220.8	3.2	256,338	3,715
3. Dila, Bay	10	79	269.9	4.4	213,220	3,476
I-D					174,388	4,222
1. Limao, Calauan	34	35	45.2	3.0	53,778	3,570
2. Dayap, Calauan	13	89	139.1	3.1	160,927	3,587
3. Balayhangin, Calauan	18	60	285.6	5.1	308,448	5,508
II-A					109,690	3,092
1. Pulo, Cabuyao	16	60	127.7	3.0	122,592	2,880
2. Poblacion, Pakil	13	183	32.5	0.8	77,311	1,903
3. San Antonio, Los Baños	54	104	23.0	0.8	129,168	4,493
II-B					25,419	1,493
1. San Juan, San Pablo	26	90	12.6	0.9	29,484	2,106
2. Panglan, Majayjay	59	18	22.5	0.8	23,895	850
3. San Gabriel, San Pablo	31	83	8.9	0.6	22,878	1,525

Table 1.1. (Continued) . . .

III-A					$\hat{X}_{ghi}$ = 362,048	$\hat{Y}_{ghi}$ = 12,038
1. Segunda Polo, A Lumban	101	24	179.0	6.5	433,896	15,756
2. Segunda Polo, B Lumban	101	24	162.0	4.5	392,688	10,908
3. Masinao, Sta. Maria	126	25	82.4	3.0	259,560	9,450
III-B					135,390	5,602
1. San Antonio, Luisiano	33	178	32.7	1.2	192,060	7,049
2. Prenza, Calamba	90	37	22.5	1.0	74,880	3,330
3. Polo Alto, Calamba	90	119	13.0	0.6	139,230	6,426
III-C					110,692	3,826
1. Sabang, Magdalena	96	11	133.7	4.2	141,120	4,435
2. Dagatan, Lilio	143	14	76.9	3.0	153,868	6,006
3. Burol, Majayjay	61	17	35.8	1.0	37,088	1,037
III-D					32,167	1,089
1. San Cristobal, San Pablo	32	219	1.7	0.2	9,300	1,095
2. San Diego A, San Pablo	32	97	0	0	0	0
3. San Diego B, San Pablo	32	97	28.1	0.7	87,200	2,173
Estimate of provincial total					$\hat{X}_g$ = 6,126,726/3	$\hat{Y}_g$ = 156,250/3

[a] The underlined figures are the corresponding estimates of the sub-strata ($\hat{X}_{ghi}$; $\hat{Y}_{ghi}$).

TABLE 1.2.

VARIANCE OF ESTIMATES AND CORRELATION BETWEEN PRODUCTION AND HECTARAGE BY SUB-STRATA. LAGUNA PROVINCE, PHILIPPINES. 1962.

Sub-strata	$\hat{\sigma}^2(\hat{X}_{ghi})$	$\hat{\sigma}^2(\hat{Y}_{ghi})$	$\hat{\sigma}(\hat{X}_{ghi}, \hat{Y}_{ghi})$	$\hat{\rho}(\hat{X}_{ghij}, \hat{Y}_{ghij})$
I - A	7,563,990,000	4,078,186	175,542,460	0.999
B	14,711,404,900	7,799,555	336,011,686	0.992
C	574,735,572,000	164,198,095	9,712,683,820	0.999
D	16,349,344,200	1,241,063	129,790,048	0.911
II - A	797,301,140	1,710,733	31,526,042	0.854
B	12,651,741	395,121	1,695,232	0.758
III - A	8,298,384,900	10,899,084	248,764,224	0.827
B	3,441,816,450	3,967,385	111,322,185	0.953
C	4,103,789,480	6,450,901	158,967,944	0.977
D	2,293,123,333	1,180,507	47,274,218	0.908

TOTAL $\sum_i \hat{\sigma}^2(\hat{X}_{ghi}) = 634{,}307{,}378{,}144$

$\sum_i \hat{\sigma}^2(\hat{Y}_{ghi}) = 201{,}920{,}630$

The estimated variances of provincial totals $(\hat{X}_g)$ and $(\hat{Y}_g)$ are given below:

$$\hat{\sigma}^2(\hat{X}_g) = (1/3)\ (634{,}307{,}378{,}144) = 211{,}435{,}792{,}700,$$

and

$$\hat{\sigma}^2(Y_g) = (1/3)\ (201{,}920{,}630) = 67{,}306{,}877\ .$$

The pooled correlation coefficient between $\hat{X}_{ghij}$ and $\hat{Y}_{ghij}$ is estimated at $\hat{p} = 0.968$ and the cv($\hat{X}_{ghij}$) is 39 per cent and cv($\hat{Y}_{ghij}$) is 27 per cent. This level of correlation is high and should be utilized in the improvement of the estimaton procedure.

Note that $\hat{X}_g$ is an important component in the FBS since $\sum_g \hat{X}_g = \hat{X}$ is the estimated production of a given crop in the country.

1.2. Results from the 1960 Agricultural Census

1.2.1. Objectives and scope of the 1960 census of agriculture

In the seminar on the analysis, evaluation and uses of agricultural census results held in Manila in December 1960, the Philippines informed the group on the status of the censuses [11, 1961]. In the 1957 conference of Asian Statisticians held in Bangkok, Thailand, the delegates from the various countries of Asia and the Far East arrived at an understanding to convince their respective governments to take a census in or around 1960 in order to achieve and establish better international comparability. In line with this understanding and in view of the fact that the 1948 census figures were no longer adequate and satisfactory as basis for planning, implementation, and evaluation of agricultural growth, the government felt a great need for another census of agriculture. The following were the specific objectives of the April 1960 Census of Agriculture:

1. To ascertain the structure and characteristics of the farms as to variations in size, type, tenure, fragmentation, land use, area of important crops, number of livestock and poultry, farm equipment and machinery, irrigation, source of power, fertilizer and soil dressing, and farm value.

2. To provide a benchmark for improving the reliability of current forecasts and estimates of crop and livestock production.

3. To provide data on the geographic distribution of rates of production of important crops.

4. To provide a satisfactory sampling frame for designing future sample surveys.

5. To provide in published form the agricultural statistics for geographic and/or administrative areas in order to meet top priority needs for government planning and for other important purposes.

Regional, provincial, congressional district, and city maps were prepared. In addition, 1200 municipal maps and 21,300 barrio maps were also prepared and distributed for use in the sample census.

1.2.2. Sampling design

The Bureau of the Census and Statistics adopted a single-stage sampling design with some elements of stratification. The stratification consisted of dividing the farms within each enumeration district in two size groups as follows:

Stratum I - Those farms with an area operated for agriculture of at least 1,000 square meters but not exceeding 10 hectares. One-third sample of this group was obtained by systematic sampling with a random start.

Stratum II - Those farms with an area operated for agriculture greater than 10 hectares. Also included in this stratum are all farms having 20 or more livestock or 100 or more poultry on the farm. All the farms in this stratum were included in the sample.

The "frame" for the Census of Agriculture was obtained from the Housing census which was conducted in February 1960, two months earlier than the Census of Agriculture. Information for the sample farms was collected by interviewing the farm operator. The results of the 1960 Census of Agriculture are presented by municipalities, by province, and for the entire country.

1.2.3. Effects of listing errors on the census frame

During the enumeration, quality checks revealed that a few farms did not belong to the stratum in which they were listed. Also, some of the units listed did not qualify as farms, and some farms were not included in the list. The farms which were listed were completely enumerated. These changes must be reflected in the estimation procedures. Let us consider the following notations:

Items	Stratum I	Stratum II
Number in the population as listed	N_1	N_2
Number in the sample selected from the lists	$N_1/3$	N_2
Non-farm units found in the sample	m_1	m_2
Additional small farms reported but not included in the frame	n_1	
Additional large farms reported but not included in the frame		n_2
Characteristic for the j^{th} farm included in the frame	Y_{1j}	Y_{2j}
Characteristic for the j^{th} farm not included in the frame	Y^*_{1j}	Y^*_{2j}

For example, the total number of farms is estimated from the relationship

$$\hat{N} = (N_1 + n_1 - 3m_1) + (N_2 + n_2 - m_2).$$

Note that the sample size in Stratum I is

$$n = [(N_1/3) - m_1].$$

The estimate of the total of a given characteristic is

$$\hat{\tau} = \hat{\tau}_I + \tau_2$$

where $\hat{\tau}_1 = 3 \left(\sum^{n} Y_{1j} \right) + \sum^{n_1} Y^*_{1j}$,

where $\tau_2 = \sum^{N_2^*} Y_{2j} + \sum^{n_2} Y^*_{2j}$

$$n = \left[(N_1/3) - m_1 \right]$$

and $N_2^* = (N_2 - m_2)$.

The variance for the estimated total $(\hat{\tau})$ is

$$\sigma^2 (\hat{\tau}) = \sigma^2 (\hat{\tau}_1)$$

where

$$\sigma^2 (\hat{\tau}_1) = 6\, n\, \hat{\sigma}^2 ,$$

$$\hat{\sigma}^2 = \left[\sum_{j}^{n} (Y_{1j} - \bar{y}_1)^2 \right] / (n-1)$$

$$\bar{y}_1 = \sum^{n} Y_{1j}/n .$$

Thus,

$$cv\ (\hat{\tau})\% = (\sqrt{6n}\, \hat{\sigma}) / \hat{\tau}_1$$

and

$$cv\ (\hat{\tau})\% = \left[(\sqrt{6n}\, \hat{\sigma}) / \hat{\tau} \right] = \left[(\sqrt{6n}\, \sigma) / (\hat{\tau}_1 + \tau_2) \right] .$$

This cv $(\hat{\tau})$ is given for each crop by municipality and by each province.

This section also illustrates the listing errors in the enumeration area and their effect on the frame and the estimation procedures. Considerable amount of time and labor was spent on these adjustments, causing considerable delay in the processing of census results.

Since the agricultural census is a sample census, then we can increase the precision of estimates by using relationships derived from the sample census results in ratio estimators. Some of these findings are given in the next section.

1.2.4. Ratio estimators to improve census results

Since imputations are obtained from census results, in the preparation of the FBS, then it is equally important to find out how the results of the sample census can be improved and who the results of these studies can be used in the improvement of the CLS. Some results on palay production will be given [12, 1964].

The Philippine 1960 Agricultural Census [13, 1963] contains data on palay production in cavans of 44 kilograms by municipality. The production data are given under five categories or strata, namely: (a) first crop, irrigated; (b) first crop, not irrigated; (c) second crop, irrigated; (d) second crop, not irrigated; and (e) kaingin and upland. In addition to palay production, the total number of farms reporting and total hectarage under palay also are tabulated by each type of strata. These two concomitant variables can be used to increase the precision of the estimate of palay production by province. Estimates of variances will be used to show the increase in precision of ratio estimators.

Five different estimators of palay production were studied and their efficiencies were compared. These estimators of total palay production of the province are as follows:

Simple random (X-only) = $\bar{T}_X = N\bar{x}$;

Stratified (optimum) = $T_X(oa) = N\bar{x}\ (oa)$;

Ratio of means = $\tilde{T}_Q = \bar{q}Z$ or $\bar{q}Y$;

Separate ratio (st) = $\ddot{T}_Q(st) = \sum_{i=1}^{L} \bar{q}_i Z_i$

or $\Sigma \bar{q}_i Y_i$;

Combined ratio (st) = $\tilde{T}_C(st) = q(st)\,Z$

or $= q(st)\,Y$;

and the

Unbiased separate ratio = $\tilde{T}_u$

The variances of each of these estimators were computed and compared.

The concomitant variables in the ratio estimators were the number of farms operating (Z_{ij}) or the hectarage (Y_{ij}) which were obtained from the same census. The results for the provinces of Abra and Pampanga as given in Table 1.3 and 1.4 show the ranges in the gain of statistical efficiency of ratio estimators. As shown in this paper, the correlation between production (X_{ij}) and number of farms reporting (Z_{ij}) is about 0.87 while the correlation between production (X_{ij}) and hectarage (Y_{ij}) averaged about 0.94. High levels of correlation exist in the 20 provinces where census results have been released by the Bureau of the Census and Statistics (see Table 1.5). Thus, we expect considerable gain in statistical efficiency in the use of ratio estimators in each of these provinces.

TABLE 1.3.

ESTIMATE OF VARIANCE OF ESTIMATOR AND RELATIVE EFFICIENCY OF ESTIMATES FOR PALAY PRODUCTION. ABRA PROVINCE. JULY 1959 TO JUNE 1960.

Type of estimator	Concomitant variable			
	Number of farms		Hectarage	
	Variance (000,000)	R.E.%*	Variance (000,000)	R.E.%*
Simple random = $\bar{T}_X$ (X-only)	20,809	100	20,809	100
Stratified (optimum) = $\bar{T}_X(oa)$	12,435	169	12,435	169
Ratio of means (r) = $\tilde{T}_Q$	10,153	205	5,160	403
Separate ratio (st) = $\tilde{T}_Q(st)$	4,980	426	3,854	682
Combined ratio (st) = $\tilde{T}_C(st)$	6,775	315	5,952	372

*R.E. is relative efficiency with the variance of $\bar{T}_X$ as numerator.

TABLE 1.4.

ESTIMATE OF VARIANCE OF ESTIMATOR AND RELATIVE EFFICIENCY OF ESTIMATES FOR PALAY PRODUCTION. PAMPANGA PROVINCE. JULY 1959 TO JUNE 1960.

Type of estimator	Concomitant variable			
	Number of farms		Hectarage	
	Variance (000,000)	R.E.%*	Variance (000,000)	R.E.%*
Stratified (optimum) = $\overline{T}_X$ (X-only)	402,376	100	402,376	100
Stratified (optimum) = $\overline{T}_X$(oa)	141,383	285	141,383	285
Ratio of means (r) = $\widetilde{T}_Q$	31,212	1,289	10,508	3,829
Separate ratio (st) = $\dot{T}_Q$(st)	14,306	2,813	6,512	6,179
Combined ratio (st) = $\widetilde{T}_C$(st)	19,834	2,029	7,463	5,392

*R.E. is relative efficiency with the variance of $\overline{T}_X$ as numerator.

TABLE 1.5.

CORRELATION COEFFICIENT BETWEEN PALAY PRODUCTION (X_{ij}) IN CAVANS, AND NUMBER OF FARMS REPORTING (Z_{ij}) AND HECTARAGE (Y_{ij}) BY STRATA AND BY PROVINCE.

Province	$\hat{\rho}(X_{ij}, Z_{ij})$						$\hat{\rho}(X_{ij}, Y_{ij})$					
	Type of strata					Whole	Type of strata					Whole
	(a)	(b)	(c)	(d)	(e)	province	(a)	(b)	(c)	(d)	(e)	province
Abra	0.78	0.87	0.82	0.85	0.59	0.90	0.91	0.94	0.91	0.89	0.26	0.87
Bulacan	.96	.97	.98	.99	.93	.96	.99	.98	.99	.99	.98	.97
Pampanga	.95	.95	.95	.99	.95	.95	.97	.99	.98	.99	.98	.98
Tarlac	.75	.90	.95	.85	.94	.81	.87	.98	.97	.99	.99	.95
Zambales	.99	.93	.99	.98	.96	.97	.99	.93	.95	.93	.94	.97
Aklan	.95	.90	.95	.91	.90	.90	.96	.91	.97	.98	.95	.96
Albay	.88	.93	.81	.92	.94	.74	.95	.99	.91	.99	.99	.92
Bataan	.77	.95	.79	.92	.92	.82	.92	.99	.92	.99	.98	.94
Cagayan	.87	.82	.96	.97	.91	.83	.95	.97	.99	.96	.96	.97
Camarines Norte	.95	.98	.95	.94	.99	.71	.99	.98	.98	.98	.99	.93
Catanduanes	.77	.95	.80	.97	.94	.82	.82	.95	.81	.92	.96	.87
Ilocos Norte	.92	.90	.99	.98	.58	.89	.88	.91	.97	.99	.99	.92
Misamis Occidental	.94	.73	.91	.56	.92	.85	.99	.91	.93	.96	.88	.81
Misamis Oriental	.93	.98	.94	.97	.95	.89	.99	.97	.99	.97	.99	.98
Nueva Ecija	.99	.94	.97	.98	.76	.83	.99	.99	.99	.99	.86	.84
Nueva Viscaya	.95	.97	.91	.74	.98	.93	.99	.97	.90	.95	.98	.97
Occidental Mindoro	.91	.96	*	.77	.84	.94	.96	.99	*	.99	.97	.99
Romblon	.95	.97	.93	.58	.99	.90	.95	.99	.99	.97	.99	.95
Sorsogon	.77	.95	.78	.98	.97	.80	.96	.95	.96	.97	.97	.93
Surigao del Sur	.99	.95	**	.43	.98	.92	.99	.98	**	.96	.99	.98

*Contains three pairs of observations.
**Contains two pairs of observations.

As the Philippine 1960 census was a sample census, then these statistical findings can be used to improve the results of the 1960 census. These findings also can be incorporated into the estimation procedure of the major crops covered in the current CLS of the Bureau of Agricultural Economics, DANR.

2. Household Food Consumption Surveys (HFCS)

The vehicle of HFCS in the Philippines is the national sample survey of households known as the Philippine Statistical Survey of Households (PSSH). Thus, it is pertinent to show how census results, maps, and lists are utilized for HFCS.

2.1. The Philippine Statistical Survey of Households (PSSH)

Briefly, PSSH divided the country into four sectors, namely: (a) Metropolitan Manila, (b) chartered cities and provincial capitals, (c) poblacions, and (d) barrios. Sectors (a) and (b) comprise the urban area, (c) the urban-rural, and (d) the purely rural area. There are a total of 10 regions. Attempts are being made to make sharper distinctions between rural and urban areas [14, 1963]. The urban area was stratified using geographical criteria, while the rural area was dissected through the use of paper strata. In the urban areas, equal probability was used, while probability proportional to size (pps)was used in the rural areas. In both areas, sampling was done with complete replacement of primaries. A description of the PSSH is given in Table 2.1. Figure 2.1 illustrates the basic estimation procedures used in the PSSH. In May, 1958, and May, 1959 consumption of rice and corn was obtained by the interview method with the use of one year recall period. In October 1960, the recall period was reduced to one week. Some pertinent results for the May 1958 surveys are as follows:

PSSH Results, May 1958

$$\frac{\text{Total population}}{\text{Rice eating population}} \qquad \frac{100\%}{76.8\%}$$

TABLE 2.1.

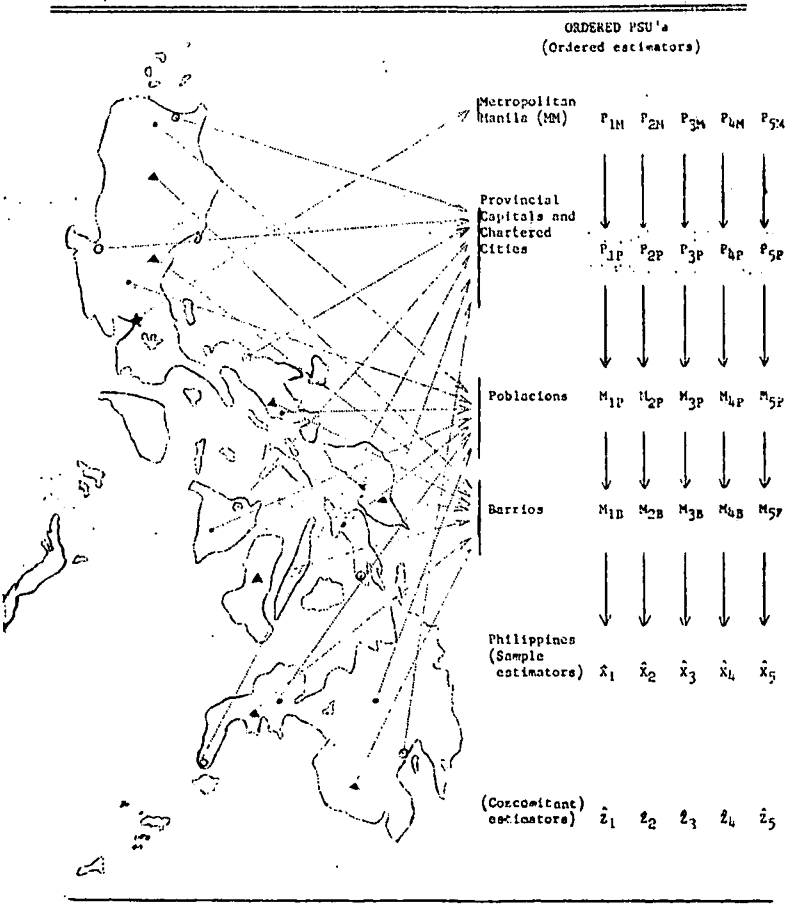
DESCRIPTION OF THE PHILIPPINE STATISTICAL SURVEY OF HOUSEHOLDS.

Sectors	Type of sampling	Primary sampling unit (PSU)	Number of strata	Number of precincts, municipalities or barrios	Number of sample households
Urban[a]					
Metropolitan Manila	Equal probability with complete replacement of PSU	precinct	32	160 (precincts)	800
Provincial capital and chartered cities	— ditto —	— ditto —	30	150 (precincts)	1050
Urban-rural[b]					
Poblacions	Unequal probality with complete replacement of PSU	municipality	30	150 (municipalities)	1500
Rural					
Barrios	— ditto —	— ditto —	30	300	3000

[a] two stages of sampling
[b] three stages of sampling

USES OF CENSUS RESULTS, MAPS AND LISTS IN FOOD CONSUMPTION STUDIES IN THE PHILIPPINES

Figure 1. SAMPLING AND ESTIMATION PROCEDURES IN THE PHILIPPINE STATISTICAL SURVEY OF HOUSEHOLDS.



Corn-eating population	20.8%
Total household consumption of rice	100%
Consumption for food	94%
For animal and poultry feed	2%
For commercial use	3%
For other purposes	1%
Average annual per capita rice consumption in sacks of 56 kilos milled rice	1.95

The Inter-Agency Committee on rice and corn production and consumption utilizes the results from the PSSH May 1959 survey [15, 1963].

Some of the most pertinent results are given below:

PSSH May 1959

Sex and age group	Annual average consumption	
	Rice[a] (sacks of 56 kilos)	Corn[b] (sacks of 57 kilos)
Males, 10 years old and over	2.299	2.761
Females, 10 years old and over	2.105	1.910
Children elow 10 years	1.138	1.440

[a] Rice-eating population, PSSH May 1958 survey

[b] Corn-eating population, PSSH May 1958 survey

The Food and Nutrition Research Center (FNRC) regional surveys use the PSSH as the sampling frame. Only three of the municipalities or PSU's were used in the rural area and only 75 per cent of the PSSH sample households were utilized in the dietary phase and also in the clinical phase. About 50 per cent of the dietary or clinical samples were utilized for the biochemical phase. Also, the FNRC nutrition surveys used the food weight method as purchased for three days, namely: Tuesday, Wednesday, and Thursday [5, 1963].

2.1.1. Uses of census results in the preparation of the sampling frame

Let us consider the rural design of the Philippine Statistical Survey of Households (PSSH) in the discussion of the uses of census results, lists, and maps in HFCS. The origin and development of the PSSH are given by Oñate (16, 1964). The design in the rural area is known as multi-stage sampling with the primary units (PSU's) drawn with probability proportional selecting PSU's with probability proportional to size (PPS) to size and with complete replacement. A simple method of selecting PSU's with probability proportional to size (PPS) which eliminates the computation of cumulative totals was utilized in drawing the PSU's in the rural areas.

It is required to select from the h^{th} stratum a PSU from list of M_h municipalities with PPS of the PSU. Let the measure of size of the i^{th} PSU be P_{hi}. The method is as follows [17, 1952]:

Choose a pair of random numbers, r and s, the first from the range 1 to M_h and the second in the range O to P'_{hi}, where P'_{hi} is the largest of the P_{hi}'s. If $P_{hi} < s$, reject the unit and if $P_{hi} \geq s$, accept the unit. If rejected, repeat the operation until a sample selection is made. Since the sampling is done with complete replacement of the PSU, this operation is repeated until m = 5 PSU's are selected. It is apparent that the list of PSU's must be accurate and that the measurement of size, P_{hi}, also must be accurate. The P_{hi} was derived from the 1948 population census results.

The estimator for the rural area is described below. The unbiased estimate of the th^{th} stratum total, X_{th} from the i^{th} sample municipaltiy (PSU) is given by

$$X^*_{hi} = (P_{hi}/p_{hi})\ (B_{hi}/b_{hi})\ \sum_j{}'\ (N_{hij}/n_{hij})\ \sum_k{}'\ x_{hijk} \qquad [2.1]$$

where

x_{hijk} is the characteristic of the k^{th} sample hh; in the jth sample barrio (or poblacion); in the i^{th} numbered PSU ($i = 1, 2, 3, 4, 5$); in the h^{th} stratum ($h = 1, 2, \ldots Lt$),

n_{hij} is the number of sample hh in the j^{th} sample barrio (or poblacion); in the ith numbered PSU; in the h^{th} stratum,

N_{hij} is the number of hhs listed in the j^{th} sample barrio; in the i^{th} numbered PSU; in the h^{th} stratum,

b_{hi} is the number of sample barrios in the ith numbered PSU; in the h^{th} stratum (this is equal to two for the barrio),

B_{hi} is the number of barrios (one for the poblacion) in the i^{th} numbered PSU; in the h^{th} stratum, barrio (or poblacion); in the i^{th} numbered PSU;

P_{hi} is the population of the i^{th} numbered PSU; in the h^{th} stratum,

and

P_h is the population in the h^{th} stratum.

Population is defined as the 1948 census population. The uses of census results in the construction of the design of the PSSH are now apparent. An example of the probabilities for 1948 and 1960 are shown in Table 2.2 for Region VII or Western Visayas. The author [18, 1960] has recommended the updating of the probabilities with the use of the 1960 census results using techniques developed by Kish and Hess [19, 1959].

We can show that X^*_{hi} is unbiased i.e. the overall mean of the X^*_{hi}'s is equal to the parameter in the th stratum, X_h and the best linear unbiased estimate of X_h is the mean of the five independent estimate which is

TABLE 2.2.

POPULATION AND PROBABILITY OF MUNICIPALITIES FOR THE 1948 AND 1960 CENSUSES FOR REGION VII — WESTERN VISAYAS. USES OF CENSUS RESULTS IN THE PSSH FRAME. PHILIPPINES.*

Stratum number	Sample municipality	1948 Population	1948 Probability (percent)	1960 Population	1960 Probabilty (percent)
16	Dauin	11,208	2.015	13,176	1.645
	Jimalalud	14,639	2.632	16,529	2.063
	Sipalay	7,270	1.307	20,313	2.536
	Cajidiocon	8,966	1.612	9,706	1.212
	San Agustin	14,425	2.593	18,372	2.294
	TOTAL	556,287		801,024	
17	Isabela	21,778	3.898	29,517	4.318
	Anilao	8,518	1.525	10,158	1.486
	Culasi	20,601	3.687	23,992	3.510
	Passi	30,918	5.534	37,807	5.531
	Buenavista	15,548	2.783	22,527	3.296
	TOTAL	558,716		683,503	
18	Ayungon	14,357	2.556	21,544	3.285
	Ayungon	14,357	2.556	21,544	3.285
	Bago	56,693	10.093	58,878	8.976
	Bago	56,693	10.093	58,878	8.976
	Murcia	28,243	5.028	23,146	3.592
	TOTAL	561,733		655,923	
19	Sagay	67,152	11,844	71,046	10.820
	Dueñas	17,842	3,147	19,000	2.894
	Hamtic	18,127	3.197	18.488	2.816
	Pontivedra	18,060	3.185	22,768	3,468
	Silay	35,570	6.274	60,133	9.158
	TOTAL	566,988		656,605	
20	Pulupandan	11.726	2.168	15.358	2.462
	Cabatuan	24.743	4.574	26,414	4.234
	Pototan	34,717	6.418	37,340	5.986
	Dumangas	29,336	5.423	30.252	4.850
	Dumangas	29.336	5.423	30,252	4.850
	TOTAL	540,935		623,802	

*Source of basic data: Bureau of the Census and Statistics. Population Census of 1948 and 1960.

independent estimate which is

$$X_h^* = (1/5)\ \Sigma'\ X_{hi}^* \qquad [2.2]$$

with variance equal to

$$\sigma_h^2 / 5 \qquad [2.3]$$

where

σ_h^2 is the variance of X_{hi}^*

For processing mass data, (2.1) and (2.2) are very inconvenient and laborious to use. The following two relations were used to simplify the computational requirements:

1) $$R = (1/m)\ (\bar{P}/\bar{p})\ .\ (\bar{B}_t/\bar{b}_t)\ .\ (\bar{N}_t/\bar{n}_t) \qquad [2.4]$$

where

R is the overall raising factor for the barrio sector (equals 850),

$\bar{n}$ is the expected average number of sample hhs per barric (equals to 10),

$\bar{N} = \sum_{hij} N_{hij} / \sum_{hi} B_{hi} = N/B$ is the average number of hhs listed per barrio,

$\bar{b}$ is the expected average number of sample barrio. This is equal to two for the barrio,

$\bar{B} = \sum_{hi} B_{hi} / \sum_{h} M_h = B/M$ is the average number of barrio per PSU or municipality

$\bar{p} = \sum_{hi} P_{hi} / \sum_{h} M_h = P/M$ is the average population per PSU,

$\bar{P} = \sum_{h} P_h / L = P/L$ is the average population per stratum,

and $m =$ is the number of sample PSU in the h^{th} stratum and is constant for all h ;

USES OF CENSUS RESULTS, MAPS AND LISTS IN FOOD CONSUMPTION STUDIES IN THE PHILIPPINES

2) $$_{hi} = n_{hij}/N_{hij} = (1/m)\ (P_h/P_{hi})\ (B_{hi}/b_{hi})\ (1/R) \qquad [2.5]$$

where

f_{hi} is the hi^{th} PSU constant and determines n_{hij} from $f_{hi} \cdot N_{hij}$ for all $j=1,2,\ldots B_{hi}$ in the hi^{th} PSU. The other symbols are as defined in [2.1].

$$\hat{X}_{hi} = (\bar{P}/\bar{p})\ (\bar{B}/\bar{b})\ (\bar{N}/\bar{n})\ \sum_j{}' \sum_k{}' x_{hijk} = m \times x_{hi}\ . \qquad [2.6]$$

The mean of the five (m) independent estimates, $\hat{X}_{hi}$'s, will give

$$\hat{X}_h = (1/m)\ \Sigma'\ \hat{X}_{hi} = R\ x_h\ . \qquad [2.7]$$

This estimator is very simple to compute.

Since $\hat{X}_h$ is unbiased, then the estimate of the sector total, X_t, is

$$\hat{X}_t = \sum_h \hat{X}_{th} = R\ x \qquad [2.8]$$

and its variance is

$$\sigma^2\ (\hat{X}_t) = \sum_h (\sigma_h^2/5) \qquad [2.9]$$

The variance of $\hat{X}_h$ is

$$(\sigma_h^2/5) .$$

and an unbiased estimate of σ_h^2 is

$$s_h^2 = \left[\Sigma' (\hat{X}_{hi} - \hat{X}_h)^2 \right] / (m-1) . \qquad [2.10]$$

Using sample totals, [2.10] becomes

$$s_h^2 = \left[(m\ R)^2/(m-1) \right] (\Sigma\ x_{hi}^2 - x_h^2/m) \qquad [2.11]$$

$$= \left[(5.850)^2/4 \right] (\Sigma'\ x_{hi}^2 - x_h^2/5)$$

where

R is the overall raising factor,
x_{hi} is the corresponding h^{th} PSU sample total,
x is the corresponding h^{th} stratum sample total,

and

m = 5.

The overall raising factor, R = 850, is obtained by the use of census data as indicated in the requirements of equation (2.4). Similarly, the PSU constant given in equation (2.5),

$$f_{hi} = \left[n_{hij}/N_{hij} \right]$$

$$= (1/m)\ (P_h/P_{hi})\ (B_{hi}/b_{hi})\ (1/850)$$

also is obtained by the use of census data P_h, P_{hi}, B_{hi} and R .

is also obtained by the use of census data P_h P_{hi}, B_{hi} and R.
The computation of these values for stratum 16 of Western Visayas is given in Table 2.3.

TABLE 2.3.

VALUES OF f^{hi} FOR THE FIVE PSU's OF STRATUM 16 OF WESTERN VISAYAS.*

Stratum (h)	Sample municipality (i)	(1/m)	(P_h/P_{hi})	(B_{hi}/b_h)	1/R	f_{hi}
16	Dauin	(1/5)	(556,287/11,208)	(16/2)	(1/850)	0.093
	Jimalalud	(1/5)	(556,287/14,639)	(14/2)	(1/850)	.063
	Sipalay	(1/5)	(556,287/ 7,270)	(23/3)	(1/850)	.207
	Cajidiocan	(1/5)	(556,287/ 8,966)	(9/2)	(1/850)	.066
	San Agustin	(1/5)	(556,287/14,425)	(12/2)	(1/850)	.054

*Source of basic data: Bureau of the Census and Statistics. See Table 2.2, stratum number 16 and year 1948.

2.1.2. Uses of maps and lists in the execution of the sampling frame

Since we have two sample barrios (b_{hi}), the solution to the problem of number of sample households to be drawn from each of the sampling area is obtained from the relation,

$$n_{hij} = f_{hi} N_{hij}.$$

For each municipality (i), we solve f_{hi} as indicated in Table 2.3. We need to know the total number of households in the sample barrio area. An accurate map with the required natural and other boundaries is essential to this operation. With this map, the field interviewer must list all the households in the barrio area to give us, N_{hij}. The results of these important requirements of the survey operations are given in Table 2.4 for stratum 16 of Western Visayas for the October 1963 and May 1964 survey rounds.

TABLE 2.4.

SAMPLING FRACTION, NUMBER OF LISTED HOUSEHOLDS AND NUMBER OF SAMPLE HOUSEHOLDS IN STRATUM 16 OF WESTERN VISAYAS.*

Stratum h	Sample municipality i	Barrio j	f_{hi}	May 1964 N_{hij}	May 1964 n_{hij}	October 1963 N_{hij}	October 1963 n_{hij}
16	Dauin**						
		1. Malongcay Dacu	.093	146	14	146	14
		2. Maayong Tubig	.093	183	17	183	17
	Jimalalud**						
		1. Bangcal	.063	154	10	154	10
		2. Owacan	.063	132	8	132	8
	Sipalay						
		1. Canturay	.207	191	40	183	38
		2. Cabadiangan	.207	198	41	188	39
	Cajidiocan						
		1. Gutivan	.066	174	11	174	11
		2. Lumbang	.066	210	14	210	14
	San Agustin**						
		1. Binonguan	.054	167	9	167	9
		2. Cagbo-Aya Cabolutan	.054	247	13	247	13

*Source of basic data: Bureau of the Census and Statistics.
**Sample PSU's for HFCS.

The list of N_{hij} is used as frame for drawing at random the sample households. The sample is drawn in a systematic manner with a random start. These techniques were described earlier in considering the CLS of the BAE, DANR. As mentioned earlier, the frame or vehicle of HFCS in the Philippines is a sub-sample of the PSSH. In the barrio sector of Western Visayas, 3 PSU's were drawn out of five and 252 households out of 439 households were selected for food weighing. Thus, adjustments must be made on the raising factor, R = 850. Other adjustments must be made to obtain the mean from the 3 PSU's and the estimated variance of this mean.

2.1.3. Uses of lists and census results in the improvement of the estimation procedures

Any set of data relating to food consumption is closely associated with size of household and to total population count. Thus, any improvement in the estimation procedure for total population count of the stratum, province, region or the nation will have direct bearing on the improvement of estimates of indicators on food and nutrition. The author's findings with the use of ratio estimators will be described in this section [20, 1963].

2.1.3.1. Population count.

The national significance of the use of ratio estimators in the design of the PSSH will be given for Metropolitan Manila. Population count is usually highly or moderately correlated with other important socio-econmic variables in a given area. Thus, the population in listed households will be used as the sample variable and this variate will generate the X-component ($\hat{X}_i$) while the number of registered voters by precinct is the Z-component ($\hat{Z}_i$). The number of registered voters is available every two years by precinct, by stratum and by the whole universe from the usual administrative channels with little or no cost to the survey. Empirical results for November, 1957 and November, 1959 are shown in Table 2.5 for the six types of estimators studied.

TABLE 2.5.

STATISTICAL EFFICIENCY OF SIX TYPES OF ESTIMATORS FOR POPULATION IN LISTED HOUSEHOLDS: METROPOLITAN MANILA, 1957 AND 1959.*

Type of estimator	Relative efficiency, percentage			
	1957a		1959b	
	separate	combined	separate	combined
Regular unbiased (X-only)				
Separate	100	—	100	—
Combined	—	100	—	100
Biased ratio of means				
Separate	127	—	105	—
Combined	—	135	—	148
Unbiased ratio				
Separate	125	—	105	—
Combined	—	136	—	148

a Seven strata were used.
b Ten strata were used.
* Source of basic data: Philippine Statistical Survey of Households and Philippine Electoral Commission.

Estimate of correlation, $\hat{p}(\hat{X}_i \hat{Z}_i)$ ranged from 4.0 to 0.5 and these correlations established gains in precision from 5 to 27 per cent for the separate ratio estimators and from 36 to 48 per cent for the combined ratio estimators as compared to the regular unbiased (X-only) estimates.

2.1.3.2 Forms of estimators

Three forms of ratio estimators may be used, one of which is applicable to any given stratum. These forms are as follows:

(a) Complete ratio estimation. The precincts are divided into segments which are known, recognizable, and identical for both electoral list and survey list. Street names or city blocks may be used to identify segments in precincts with electoral list for the precinct. Consider as estimate of X (the stratum total)

$$\tilde{X}^{**} = \tilde{q}\, Z$$

where

$$q = \left[(M/m) \sum_i{}' (S_i/s) \sum_j{}' X_{ij} \right] / \left[(M/m) \sum_i{}' (S_i/s) \sum_j{}' Z_{ij} \right]$$

X_{ij} is the listed segment total,

Z_{ij} is the corresponding electoral list total for the segment,

Z is the electoral list total for the stratum,

and

s s the number of segments (segmentation of precinct is done for possible rotation of sample).

(b) Ratio estimation in the primary stage. In this case, the segment totals $X_{ij}(j=1,2,3)$ are available for the survey data only, while for the electoral counts Z only precinct (PSU) totals, $Z_i(i=1,2,\ldots,M)$, are available. The ratio estimator has the form

$$\hat{X}^{**} = \hat{q}\, Z$$

where

$$\hat{q} = \left[(M/m) \sum_i{}' (S_i/s) \sum_j{}' X_{ij} \right] / \left[(M/m) \sum_i{}' Z_i \right] .$$

(c) No segmentation. The precinct is completely listed by the survey. We have available at the precinct or PSU level, the PSU totals, $X_i(i=1,2,3,4,5)$ and the PSU electoral counts, $Z_i(i=1,2,\ldots,M)$. Let the estimator be

$$\bar{X}^{**} = q\,Z$$

where

$$q = \left[(M/m) \sum_i{}' X_i \right] / \left[(M/m) \sum_i{}' Z_i \right] = \hat{X}^{*}/\hat{Z}^{*} .$$

This is the theoretical framework which was used to obtain the statistical efficiency of the six types of estimators given in Table 2.5. Note that the precision of these estimators hinges on the availability of accurate maps, lists, and census results specifically the value of Z.

One may attempt to estimate from the sample household stage. It should be pointed out that this technique is too difficult to implement in the field. This technique will be workable if we are estimating characteristics other than population. Estimation at the sample household stage will be incated for food group (X) and size of households (N).

On a national or regional level, the signifance of the use of simple ratio estimators as applied to the urban area of the Philippine Statistical Survey of Households becomes more apparent in view of the following:

a) Sampling was with complete replacement of PSU, and the reduction in variance of the X-only estimate due to the application of the finite population correction (1-f) is lost where f is the sampling fraction.

b) The PSU's were drawn with equal probability and not with probability proportional to size (PPS).

While conditions a) and b) will result in rather simplified estmating procedures which are of advantage when applied in developing economies, considerable losses in precision may exist if the X-only estimate is used. The use of ratio estimators generally will recover these losses in statistical efficiency with little or no cost to the survey proper. On the average, there is a gain of about 28 per cent precision in the use of ratio estimators. This implies that we can reduce the number of PSU's by 28 per cent and still maintain the level of precision of the X-only estimate. Peso-wise, this reduction in the number of PSU will also result in the reduction of cost of the survey by about 28 per cent.

2.1.3.3. Indicators of food consumption.

Let us consider the use of an auxillary variable in the estimation of the total cereal intake (X) of an area or region. Such an estimate will require the use of a concomitant variable, say, the size of the household (N) which is correlated to (X). The ratio estimator of total cereal intake in the hth stratum is

$$\hat{X}'_h = q_h N_h$$

where

$$q_h = \hat{X}_h / \hat{N}_h$$

= sample total of the X_{ij}'s/sample total of th N_{ij}'s,

and

N^h is the total population of the area or stratum and is obtained primarily from the list of census results.

Note that in q_h, the raising factors R will cancel out leaving only the sample totals. Also, N_h may be an estimate N^*_h which is derived from the listed population in each barrio (or precinct in the urban area). Thus, in the rural area,

$$N_h^* = (1/5) \sum_i{}' (P_h/p_{hi}) (B_{hi}/2) \sum_j{}' N_{hij}$$

$$= (P_h/10) \sum_i{}' (B_{hi}/P_{hi}) \sum_j{}' N_{hij}$$

where

N_{hij} is the number of listed households in the (ij)th sample barrio

and

P_h, P_{hi}, and B_{hi} are as defined before.

In the case of HFCS, the food intake (X_{hijk}) and the number or size of the households (N_{hijk}) are obtained at the household level. Our concomitant variable may be N_h^* as described above or by N_h^{**} which is estimated with the use of number of registered voters in the PSU or in the stratum (section 2.1.3.1). N_h^{**} will be comparable to $\tilde{X}^{**}$, $\hat{X}^{**}$ or $\bar{X}^{**}$.

The level of correlation between (X) and (N) is important in order that a gain in precision is obtained. The results shown in Table 2.6 will indicate the levels of correlation for some food groups.

TABLE 2.6.

CORRELATION BETWEEN FOOD GROUP AND SIZE OF HOUSEHOLDS.

Region	Stratum	Food Group	correlation, coefficient p (Food group, size of hhs.)
Resurvey of Greater Manila — September to December 1959	14	Total cereals	0.87[a]
Southern Luzon and Island — January to April 1962	9	Rice consumption	0.80[b]
Western Visayas · Region VII — February to May 1964	16	Corn and Rice consumption	0.85[b]

a Between precincts. Covariance analysis on pp. 80 of M.A. Thesis "On Ratio Method of Estimation in a Hierarchical Sample Survey Design" by M. Aslam. U.P. Statistical Center Library, 1962.
b Between barrios. The author acknowledges the help of the FNRC staff for the preparation and the use of the original data.

In all cases, there wll be a gain in precision in the use of ratio estimators since the relation

$$\hat{p} > 1/2 \text{ [cv (size of household)/cv (food group)]}$$

was found to hold for the three food groups in Table 2.6. This analytical work will be continued with the cooperation fo the FNRC technical staff.

3. Recent Efforts in Sampling for Population Count

The design of the PSSH will soon be changed in order to estimate population count and its components at the provincial level. The results of these findings for each province are given in a manuscript prepared recently by the author [21, 1964]. These findings were obtained with the use of the 1960 Population Census. The results of this study and the contemplated move to adopt these results in the redesign of the PSSH will have direct bearings on the current and future HFCS of the FNRC.

4. Summary and Recommendations

Undoubtedly, there are other uses of census results, maps and lists in food consumption studies. This has shown that if the boundaries or borders of the sampling area (SA) or enumeration district (ED) are not not clearly indicated in the map or the mapping material is not complete and lacks in the desired requirements, then this situation will lead into listing errors. This listing errors will definitely affect the census results or the precision of the estimation procedures of the survey. The Philippine 1960 Agricultural Census, the Crop and Livestock Survey, and the Philippine Statistical Survey of Households (PSSH) illustrate the interactions of quality of maps, accuracy of lists and correctness of census results on the quality of results on food consumption studies. In both censuses and sampling surveys, any bias in the total number of units under discourse will automatically bias the total for each of the characteristics under study. Even if maps are accurate, listing errors will still occur due to various causes. The accuracy of maps and lists must be subjected to rigid quality control.

Analytical studies are necessary to show the level of variability of each indicator in food consumption studies. These studies should include the use of ratio estimators with recognized auxilliary variables.

The design of the PSSH will undergo some changes. Some of the most important are (a) the use of provinces as strata, (b) the use of the barrios as PSU in the rural area and the precincts in the urban area, and (c) application of the new definitions of urban and rural areas. These innovations will create some disturbances in the original HFCS series and these changes will definitely affect the administrative aspects and field operations of the HFCS of the FNRC. The effects of these changes on the conduct of the nutrition surveys of the FNRC must be critically studied. Also, the results of the 1960 agricultural census are now ready for use in the preparation of the Philippine Food Balance Sheets.

The improvement of these basic sources of information in food consumption studies in the Philippines will provide the necessary statistical framework in the planning, implementation, and evaluation of food and nutrition programs in the country. The availability of census results, accurate maps, and lists will play an important role in the improvement of these sources for food consumption studies.

5. Literature Cited

1. OÑATE, B. T. The Role of Statistics in Philippine Development. Proceedings of the Seminar on the Social and Economic Aspects of Agricultural Development. College of Agriculture, University of the Philippines. May 1964.

2. REH, E. Manual on Household Food Consumption Surveys. Food and Agriculture Organization of the United Nations. Rome. 1962

3. FOOD AND AGRICULTURE ORGANIZATION of the United Nations. Program of Food Consumption Surveys. Rome. 1964.

4. OÑATE, B. T. The Statistical System in Philippine Development. The Economic Bulletin 1 (1). National Economic Council. Manila. November 1962.

5. QUIOGUE, E. S., E. M. Palma and G. M. Villavieja (Dietary Phase); ROXAS, B. V., L. G. Alejo, L. A. Salamat, M. D. Kuizon, C. R. Facundo, R. D. Salud and E. M. Tinio (Biochemical Phase); PERALTA, F. L., J. B. Jayme, F. F. Piedad, A. B. de Gala and L. C. Hortelano, (Clinical Phase). "Nutrition Survey of Ilocos-Mountain Province Region". Food and Nutrition Research Center, NIST, NSDB, Manila. NSDB, Printing Press. 1963

6. NATIONAL ECONOMIC COUNCIL. The Food Balance Sheet Manuel. Philippines. October 1962.

7. ——————————. The Food Balance Sheets of the Philippines for 1953 to 1962. The Statistical Reporter 7 (4). December 1963.

8. BUREAU OF AGRICULTURAL ECONOMICS. Manual of Instructions for Field Statisticians. 1964 Crop and Livestock Survey. BAE, DANR. 1964.

9. BUREAU OF AGRICULTURAL ECONOMICS. Reference Manual for Interviewers. 1964 Crop and Livestock Survey. BAE, DANR. 1964.

10. CASTILLO, J. V. Improvement of Frame for the Crop and Livestock Survey. The Statistical Reporter 2 (1). National Economic Council. Manila, 1958.

11. BUREAU OF THE CENSUS AND STATISTICS. Informative Report of the Philippines on the 1960 Census of Agriculture. The Statistical Reporter 5 (1). Manila. January 1961.

12. OÑATE, B. T. Estimation of Palay Production With the 1960 Agricultural Census as the Sampling frame. Proceedings issue of the Anual Conference of the Philippine Statistical Association. July 1964.

13. BUREAU OF THE CENSUS AND STATISTICS. 1960 Agriculture Census Bulletin. Vol. 1. Report by Province. April 1963.

14. NAZARET, F. V. and F. R. Barreto. Concepts and Definitions of Urban-Rural Areas in the Philippines. Read at the 11th annual conference. Philippine Statistical Association. 1963.

15. NATIONAL ECONOMIC COUNCIL. Inter-Agency Report on Rice and Corn Production and Consumption. October 30, 1963.

16. OÑATE, B. T. Statistics in Rice Research. Part III. Absolute Experiments. Bound manuscript. The International Rice Research Institute. Los Banños, Laguna, Philippines. 1964. p. 16-28; Chapter XII.

17. LAHIRI, D. B. Report to the Statistical Commission on the Fifth Session of the Sub-Commission on Statistical Sampling held from 19 to 31 December 1951. Sankya 12:202. 1952.

18. OÑATE, B. T. Development of Multi-Stage Designs for Statistical Surveys in the Philippines. Statistical Reporter 4 (4). National Economic Council. Manila. October 1960.

19. KISH, L. and I. Hess. Some Sampling Techniques for Continuing Survey Operations. Mimeo paper. Survey Research Center. University of Michigan. December 1959.

20. OÑATE, B. T. Ratio Estimators in Field Research. Paper presented to the Philippine Statistical Association on December 6, 1963 at the U.P. Statistical Center. (In Press).

21. ——————. Estimation of Population Count by Province With the 1960 Philippine Population Census as the Sampling Frame. The International Rice Research Institute. Los Baños, Laguna. 1964.