The Philippine Statistician

JAN. - AUG., 1969 • VOL. XVIII • NOS. 1-2

THE PHILIPPINE STATISTICIAN

(Entered as second class mail matter at the Manila Post Office on August 25, 1953)

Annual Subscription: Philippines — ₱6.00; Foreign — $4.00

NOTICE TO AUTHORS

The Editor welcomes the submission of manuscripts on theoretical and applied statistics.

Manuscripts should be clearly typewritten double spaced with ample spacing on all sides, top and bottom, left and right. Our conventions are: italics are underlined by a straight line, bold-faced letters by a wavy line, Greek by a red line, and script letters are encircled in blue. Graphs, figures and diagrams should be drawn expertly in India ink on plain white paper and roughly twice the size they are to appear in print. All formulas should be written in linear form: vertical forms should be avoided, e.g.

$$(a + b) / (c + d) \text{ instead of } \frac{a + b}{c + d}$$

Subscripts and superscripts should be properly and clearly placed below and above the line, and, if possible, should be written in India ink. All symbols should be automatically italicized.

Footnotes are to be enumerated successively on a separate sheet and should be minimized, or, if at all possible, completely avoided.

References should be given at the end of the article on a separate sheet. The following forms should be followed:

[2] M. A. NAIMARK, *Normed Complete Rings.* Groningen: P. S. Noordhoof, 1964.

[3] JAMES R. CLAY AND JOSEPH J. MALONE JR. "The near rings with identities on certain finite groups," *Mathematica Scandinavica,* Volume 19, Number 1 (1966) pp. 146-167.

The Philippine Statistical Association is not responsible for the theories, statements, or opinions expressed in the addresses delivered and papers read in its meetings, published in The Philippine Statistician. The authors of addresses and papers assume sole responsibility.

Office of Publication
RIZAL HALL — Padre Faura, Manila
P.O. Box 3223

The Philippine Statistician
Vol. XVIII, Nos. 1 and 2, 1969

DISTRIBUTION OF FAMILY INCOME IN THE PHILIPPINES[1]

CRISTINA P. PAREL[2]

1. *Introduction*

The wide gap between the very poor and the very rich has always been the concern of the government. Government leaders are aware that great inequalities of wealth are the roots of real social problems. Senator Manahan in one of his speeches has expressed such fears: "The Philippines faces a mass upheaval unless it narrows the gap between the rich and the poor.... There is a strong possibility that unless remedial measures are taken, there may be a mass upheaval.... The social unrest in Central Luzon may be traced to the poor lot of the peasants...."

The present study is an attempt:

i) to examine the techniques by which data on family household income may be most effectively presented and summarized;

ii) to ascertain how family household incomes are distributed and to make clear the implications of such a distribution.

The distribution of income has attracted the attention not only of economists but also of the statisticians since Pareto's time. It has been the subject of many theoretical and empirical investigations. A number of local economists have been making studies on personal and family incomes.

---

[1] Paper is based on the data from the BCSSSH, 1965, Bureau of Census and Statistics, Department of Commerce and Industry.
[2] Director of the U.P. Statistical Center

The information on the distribution of family household income is undoubtedly of far reaching importance to government leaders and economic planners. The distribution may suggest the level of living that is obtaining among the family households. With this knowledge, government administrators and socio-economic planners have a useful indicator by which they can determine which sectors merit attention and assistance.

The shape of the distribution of family households may call attention to the nature of the inequalities or disparities. With this information, efforts may be exerted by government leaders to discover solutions or means by which these disparities can be reduced or minimized.

2. *Source of Data*

The data on family household incomes are obtained from the third national survey on family income and expenditures of the Bureau of Census and Statistics Statistical Survey of Households (BCSSSH) conducted in May, 1966.

The survey used a two-stage sampling scheme in both the urban and rural areas. In the urban areas, in each region (I to X), the municipalities were ordered according to the number of urban electoral precincts (as of the 1963 elections) in them. These electoral precints were then grouped into 291 strata in such a manner that the strata contain approximately the same number of electoral precints. However, in no case was a municipality divided. From each stratum, two electoral precints were chosen at random, and from each chosen precincts, household were chosen. The over-all sampling fraction in each stratum was 1/400, except for urban Romblon where an overall sampling fraction of 1/200 was applied.

In the rural areas, the barrios were ordered according to population in each region and then grouped into 200 strata. Strata were so formed that each stratum consists of about the same population size, but in no way was a barrio divided.

Barrios were the primary sampling units, and households were the secondary sampling units. An overall sampling fraction of 1/1,200 was applied in each stratum, except for Batanes (all rural) where 1/200 was used.

Definition of Terms

The following definitions are in accordance with those of the BCSSSH:

i) Family Household — group of persons related by blood, marriage or adoption, and living together under the same roof, excluding boarders, guests, or domestic help. A person living alone is taken as a family household.

ii) Family Income — the aggregate income received or realized by family members during the calendar year, 1965. The income of the family is classified according to source into two general categories: (a) income from work or any activity of members of the family; and (b) income from sources other than work.

The Log-Normal Distribution

The univariate log-normal distribution is completely specified by two parameters μ and σ^2. Suppose that X is essentially a positive random variable such that $Y = \log_e X$. If Y is normally distributed with mean μ and variance σ^2, then it is said that is log-normally distributed. The cumulative distribution of is given by:

$$F(X) = \int_{o}^{x} \frac{1}{t\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma}(\log t-\mu)^2} dt, \; x > 0$$

$$= o, \; x \leqq O$$

The rth moment of X can be easily shown to be

$$E(X^r) = \int_0^\infty x^r \, dF(x)$$

$$\exp\{ r\mu + \frac{r^2\sigma^2}{2} \} ; \; r = 1,2,3,\ldots$$

In particular, if r = 1,

$$E(X) = \exp\{\mu + \frac{\sigma^2}{2}\}, \text{ the mean of the distribution.}$$

If r = 2,

$$E(X^2) = \exp\{2\mu + 2\sigma^2\}.$$

The variance is, therefore,

$$\sigma^2 = E(X^2) - [E(X)]^2$$

$$= \exp\{2\mu + \sigma^2\} [\exp\{\sigma^2\} - 1]$$

The coefficient of variation is

$$\text{c.v.} = \frac{\sigma}{\bar{X}} = [\exp\{\sigma^2\} - 1]^2$$

The measure of skewness is

$$\alpha_3 = \frac{\mu_3}{(\sigma^2)^{3/2}} = \frac{\exp\{3\sigma^2\} - 3\exp\{\sigma^2\} + 2}{[\exp\{\sigma^2\} - 1]^{3/2}}$$

In a distribution, commonly the arithmetic mean and the standard deviation are computed. In a log-normal distribution, the arithmetic mean and standard deviation are computed from the logarithms of the observations. The antilogarithm of the mean is the geometric mean, and the antilogarithm of the standard deviation is the so-called *standard deviation ratio* of the variable. It is a convenient index of the degree of skewness of the lognormal curve. The standard deviation ratio, σ_r, and Bowley's quartile measure of skewness (Sk_Q) were found to be related, and the relation is expressed by the equation:

$$\sigma_r = \left(\frac{1 + Sk_Q}{1 - Sk_Q}\right)^{1.4826}$$ [3]

Fitting A log-Normal Curve To Family Household Income Data

The family household income distributions for Greater Manila, the urban areas of Regions II — X[4] and the rural areas for the same regions exhibit the same well-known characteristics of income distributions: they are unimodal and are positively skewed.

Pearson Type II and the log-normal distribution were "fitted" to the observed distributions for the above-mentioned areas, but the observed distributions are "closest" to the log-normal distribution.

As can be noted in Table II, the mean (geometric) family household incomes range from ₱1,600 (Region VI) to ₱3,498 (Greater Manila Area) in the urban areas. For the rural areas, the mean (geometric) family household incomes range from ₱854 (Region III) to ₱1,465 (Region V). The least skewed of the distributions in the urban areas is that for Region X; the most skewed is that for Region IX. However, the differences are very slight; the difference between the skewness of the least skewed and the most skewed being only .0627. For the rural areas, the least skewed is that for rural Region VIII, and those for rural Regions II and III are the most skewed. The fitted curves are shown in Figures 1 - 3

---

[3] Davies, G.R. "The Logarithmic Curve of Distribution", JASA, Dec. 1925.

[4] The regions are as defined by the Bureau of Census and Statistics. See next page Table II.

TABLE II. ESTIMATES OF MEANS, STANDARD DEVIATION RATIOS AND SKEWNESS OF FAMILY HOUSEHOLD INCOME DISTRIBUTIONS BY AREAS AND REGIONS (1965)

REGIONS	URBAN					RURAL				
	$\overline{X}_{log}$	(Pesos) G.M.	σ_{log}	σ_r	SK_Q	$\overline{X}_{log}$	(Pesos) G.M.	σ_{log}	σ_r	SK_Q
II	7.5982	1,998	0.8281	2.2933	0.2722	6.8736	963	0.8080	2.2479	0.2659
III	7.3811	1,605	0.9048	2.4556	0.2957	6.7458	854	0.8065	2.2479	0.2654
IV	7.6906	2,186	0.8760	2.4109	0.2868	7.3154	1,097	0.7651	2.1598	0.2525
V	7.9368	2,807	0.7580	2.1383	0.2502	7.2921	1,465	0.7846	2.1815	0.2584
VI	7.3780	1,600	0.9084	2.4843	0.2970	7.1196	1,236	0.7779	2.1815	0.2562
VII	7.6181	2,038	0.8072	2.2479	0.2654	7.1016	1,212	0.7489	2.1170	0.2473
VIII	7.5027	1,808	0.8475	2.3396	0.2783	6.8531	944	0.3871	1.4770	0.1298
IX	7.6013	1,998	0.9130	2.4843	0.2982	6.9725	1,064	0.7427	2.0959	0.2452
X	7.7047	2,219	0.7073	2.0340	0.2343	7.1592	1,287	0.7807	2.1815	0.2574
Greater Manila	8.1601	3,498	0.7063	2.0340	0.2365	—	—	—	—	—
All Non-Manila, Urban	7.6649	2,143	0.8413	2.3164	0.2764	—	—	—	—	—
All Rural	—	—	—	—	—	7.0511	1,153	0.8243	2.2705	0.2710

Reg. II: Abra, Ilocos Norte, Ilocos Sur, La Union, Mt. Province
Reg. III: Batanes, Cagayan, Isabela and Nuva Vizcaya
Reg. IV: Bataan, Bulacan, Nueva Ecija, Pampanga, Pangasinan, Tarlac and Zambales
Reg. V: Batangas, Cavite, Laguna, Marinduque, Occidental Mindoro, Oriental Mindoro, Palawan, Quezon, Rizal
Reg. VI: Albay, Camarines Norte, Camarines Sur, Cantanduanes, Masbate and Sorsogon
Reg. VII: Aklan, Antique, Capiz, Iloilo, Negros Occidental and Romblon
Reg. VIII: Bohol, Cebu, Negros Oriental, Leyte, Samar
Reg. IX: Cotabato, Davao Zamboanga del Norte, Zamboanga del Sur and Sulu
Reg. X: Agusan, Bukidnon, Lanao del Norte, Lanao del Sur, Misamis Occidental, Misamis Oriental and Surigao.

TABLE III A. LOG NORMAL DISTRIBUTION FITTED TO INCOMES OF FAMILY HOUSEHOLDS IN GREATER MANILA AREA (1965)

CLASS LIMITS	Log of Limits		Xi log (Log of Limit —X)	Xlog / σ	Cum. Freq. As %	Theoretical Freq. As %	Theoretical Freq.	Actual Freq.
	Lower	Upper						
0 — 999	—	—			50.000	3.836	28.46	25
1,000 — 1,999	6.9078	—	— .2527	— 1.7733	46.164	17.640	130.89	114
2,000 — 2,999	8.0064	—	— .5597	— .7924	28.524	19.818	147.06	175
3,000 — 3,999	8.0064	8.2939	— .1542 + .1338	— .2183 + .1894	8.706 7.535	16.241	120.50	118
4,000 — 4,999	—	8.5170	.3569	.5053	19.497	11.962	88.76	81
5,000 — 5,999	—	8.6994	.5393	.7635	27.637	8.140	60.40	58
6,000 — 6,999	—	8.8536	.6935	.9818	33.646	6.009	44.59	38
7,000 — 7,999	—	8.9871	.8270	1.1708	37.900	4.254	31.54	23
8,000 — 8,999	—	9.1048	.9447	1.3375	40.988	3.088	22.93	27
9,000 — 9,999	—	9.2103	1.0502	1.4869	43.189	2.201	16.32	24
10,000 — 10,999	—	9.3056	1.1455	1.6218	44.738	1.549	11.50	26
11,000 — 11,999	—	9.3926	1.2325	1.7450	45.994	1.256	9.35	13
12,000 — 12,999	—	9.4727	1.3126	1.8584	46.856	.862	6.38	15
13,000 — 13,999	—	9.5468	1.3867	1.9633	47.500	.644	4.75	5
14,000 and above	—	—	—	—	50.000	2.500	18.55	—
							741.98	742

$\overline{X} = 8.1601$ $\sigma = .7063$

TABLE III B. LOG NORMAL DISTRIBUTION FITTED TO INCOMES OF FAMILY HOUSEHOLDS IN THE URBAN AREAS (1966)

CLASS LIMITS	Log of Limits		Xlog (Lot of Limit —X	Xi / σ	Cum. Freq. As %	Theoretical Freq. As %	Theoretical Freq.	Actual Freq.
	Lower	Upper						
0 — 999	—	—	—	—	50.000	18.406	331.38	293
1,000 — 1,999	6.9078	—	— .7571	— .8999	31.594	28.406	511.38	495
2,000 — 2,999	7.6009	8.0061	— .0640 .3412	— .0760 + .4055	3.188 15.910	19.098	343.80	396
3,000 — 3,999	—	8.2939	.6290	+ .7476	27.337	11.427	205.74	229
4,000 — 4,999	—	8.5170	.8521	1.0128	34.375	7.038	126.54	106
5,000 — 5,999	—	8.6994	1.0345	1.2296	39.065	4.690	84.42	92
6,000 — 6,999	—	8.8536	1.1887	1.4129	42.073	3.008	54.18	47
9,000 — 9,999	—	8.9871	1.3122	1.5597	44.062	1.989	35.82	34
8,000 — 8,999	—	9.1048	1.4399	1.7115	45.637	1.575	28.44	34
7,000 — 7,999	—	9.2103	1.5454	1.8369	46.712	1.075	19.44	22
10,000 — 10,999	—	9.3056	1.6407	1.9501	47.441	.729	13.14	20
11,000 — 11,999	—	9.3926	1.7277	2.0536	47.982	.541	9.72	13
12,000 — 12,999	—	9.4727	1.8078	2.1488	48.422	.440	7.92	13
13,000 — 13,999	—	9.5468	1.8819	2.2368	48.745	.323	5.76	7
14,000 — and above	—	—	—	—	50.000	1.255	22.68	—
							1,800.36	1,800

$\bar{X} = 7.6649$ $\sigma = .8413$

TABLE III C. LOG NORMAL DISTRIBUTION FITTED TO INCOMES OF FAMILY HOUSEHOLDS IN THE RURAL AREAS (1965)

CLASS LIMITS	Log of Limits		Xlog (log of Limit — X	Xi / σ_{log}	Cum. Freq. As %	Theoretical Freq. As %	Theoretical Freq.	Actual Freq.
	Lower	Upper						
0 — 504	—	—	—	—	50.000	15.866	317.16	306
505 — 1,004	6.2246	—	— .8265	— 1.0027	34.134	27.778	555.25	460
1,005 — 1,504	6.9126	7.3159	— .1385 + .2648	— .1680 .3212	6.356 12.552	18.908	377.97	395
1,505 — 2,004	—	7.6029	.5518	.6694	24.857	12.305	245.98	293
2,005 — 2,504	—	7.8257	.7746	.9397	32.639	7.782	155.56	185
2,505 — 3,004	—	8.0077	.9566	1.1605	37.698	5.059	101.13	146
3,005 — 3,504	—	8.1617	1.1106	1.3473	41.149	3.451	68.99	88
3,505 — 4,004	—	8.2951	1.2440	1.5092	43.448	2.299	45.96	44
4,005 — 4,504	—	8.4138	1.3627	1.6532	45.053	1.605	32.08	66
4,505 — 5,004	—	8.5180	1.4679	1.7808	46.246	1.192	23.85	26
5,005 and above	—	—	—	—	50.000	3.754	75.04	—
							1,998.95	1,999

$\overline{X} = 7.0511$ $\sigma = .8243$

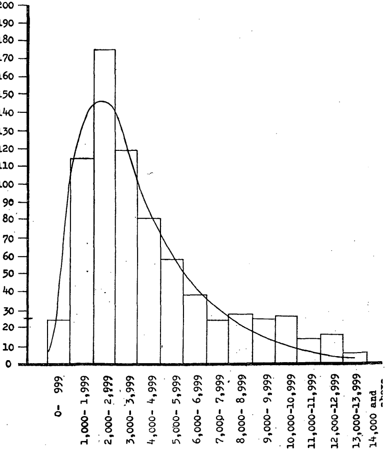
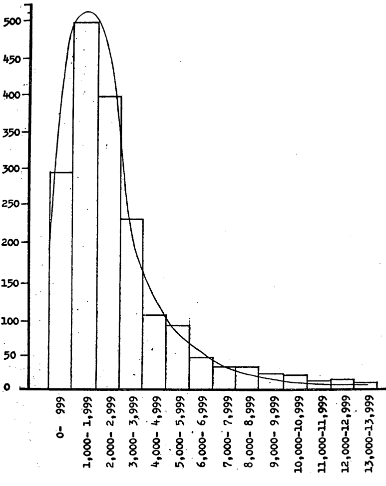


Figure I. Log Normal Distribution Fitted to Family Household Data (Greater Manila Area), 1965.



Figurt II. Log Normal Distribution Fitted to Family Household Income Data (Urban Areas, Reg. II-X), 1965.

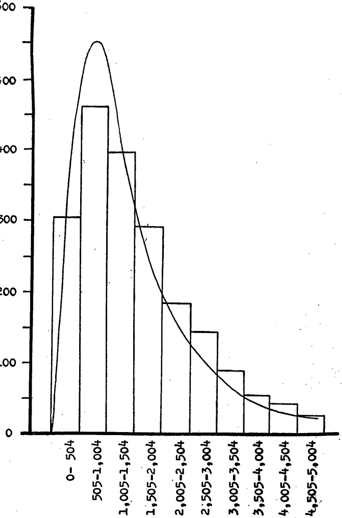


Figure III. Log Normal Distribution Fitted to Family Household Income Data (Rural Area, Reg. II-X), 1965

Analysis of Variance of the Family Household Income Data

In view of the inequality of the number of sample family households in the regions, a two-way analysis of variance was made on mean of the logarithms of family household incomes. The two variables used for classification are "regions" and "type of areas" (urban or rural).

TABLE OF MEANS BY REGION AND AREA

REGIONS	Areas				Total	
	Urban		Rural			
II	7.5982	(70)	6.8736	(149)	14.4718	(219)
III	7.3811	(42)	6.7458	(96)	14.1269	(138)
IV	7.6906	(367)	7.3154	(280)	15.0060	(650)
V	7.9368	(360)	7.2961	(235)	15.2369	(595)
VI	7.3780	(122)	7.1196	(184)	14.4976	(306)
VII	7.6181	(247)	7.1061	(242)	14.7197	(489)
VIII	7.5027	(269)	6.8531	(401)	14.3558	(670)
IX	7.6013	(130)	6.9725	(259)	14.5738	(389)
X	7.7047	(193)	7.1592	(153)	14.8639	(346)
Total	68.4115	(1800)	63.4369	(1999)	131.8484	(3799)

Note: The numbers in parenthesis are the numbers of sample family households.

ANALYSIS OF VARIANCE TABLE

SOURCE OF VARIATION	d.f.	S. of S.	M.S.	F
Regions (A)	8	0.4627	.0578 (8.6700)	26.76*
Areas (Rural-Urban) (B)	1	1.3748	1.37488 (206.22)	641.1**
Interaction (A x B)	1	.1597	.1996 (2.994)(1)	9.24*
Error	3,781	—	.00216 (.3251)	—

** Significant at 1% significance level.

(1) $S^2 = \dfrac{S.S.}{3781}$ (within cells); $S^{12} = \dfrac{S^2}{\bar{n}_L}$ where $\bar{n}_L = \dfrac{pq}{\sum_{i=1} \frac{1}{n_i}}$ 20

From the analysis of variance table, it may be inferred that geographical location and urbanization affect family household income. The level of development of the region is an important consideration in this case. It is to be noted that region as a factor in family income is very highly significant. The "type of area" (urban or rural) as a factor can be explained by the fact that in cities, there are more activities, social and economics; hence, more money in circulation. There is also a significant interaction between "regions" and "type of areas" (rural and urban), although it is not as highly significant as the two factors.

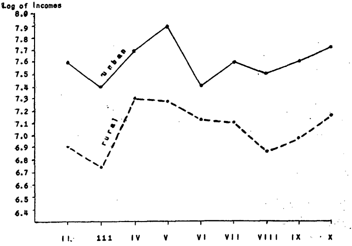


Fig. 4. Logarithm of Family Household Incomes for Rural and Urban Areas

The Concentration of Family Household Incomes

In measuring the concentration of family household income, the standard deviation of the log-normal distribution is used.

Let $F(x) \equiv$ the proportion of family household incomes not exceeding x pesos; and

$F_1(x) \equiv$ the proportion of total income accruing to family households not exceeding x pesos; that is,

$$F_1(x) = \int_0^x t\, dF(t) \Big/ \int_0^\infty t\, dF(t)$$

In this case, F(x) is the log-normal distribution; that is,

$$F(x) = \int_0^x \frac{1}{t\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2\sigma^2}(\log t - \mu)^2\right\} dt,$$

$$\text{for } x > 0$$

$$= 0, \text{ for } x \leqq 0.$$

Notationally, if $F(x) \equiv \Lambda(x)|\mu,\sigma^2)$, then

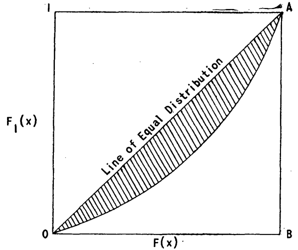
$$F_1(x) = \int_0^x td\, \Lambda(t\ |\mu,\ \sigma^2) \Big/ \int_0^\infty td\ (t\ |\mu,\ \sigma^2)$$

$$= \int_0^x \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2\sigma^2}(\log t - \mu)\right\} dt \ /\ \mu + \sigma^2/2$$

$$= \int_0^x \frac{1}{t\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2\sigma^2}(\log t - \mu - \sigma^2)^2\right\} dt$$

$$= \Lambda\ (x\ |\ \mu + \sigma^2,\ \sigma^2).$$

From the Lorenz diagram, the measure of concentration, L, is defined as:



$$L = \frac{\text{Area of the shaded portion}}{\text{Area of } \triangle \text{ AOB}}$$

It can be noted that $0 \leq L \leq 1$. The two limits $L = 0$ and $L = 1$ indicate a completely egalitarian community (that is, family household incomes are equally distributed), and a community wherein all the income goes to only one individual.

From the definition of L above, since the area of the $\triangle$ AOB = ½ sq. units, L = 2 (area of the shaded portion).

But the area of the shaded portion is $\{ \frac{1}{2} - \int_0^1 F_1(x)\, dF(x) \}$.

Hence,

$$L = 2 \{ \tfrac{1}{2} - \int_0^1 F_1(x)\, dF(x) \}$$

Or, $L = 1 - 2 \int_{0}^{1} F_1(x)\, dF(x)$

With the assumption that $F(x) = \Lambda(x \mid \mu, \sigma^2)$,

$$L = 1 - 2 \int_{0}^{1} \Lambda \mid (x \mid u + \sigma^2, \sigma^2) \mid d\Lambda(x \mid \mu, \sigma^2).$$

Using the transformation $Y = (\log x - \mu)/\sigma$, the preceding equation becomes,

$$L = 1 - 2 \int_{0}^{1} N(y - \sigma \mid 0, 1)\, dN(y \mid 0, 1)$$

$$= 1 - 2N(y - \sigma \mid 0, 2)$$

Or,

$$L = 2N\left(\frac{\sigma}{\sqrt{2}} \mid 0, 1\right) - 1,$$

where $N(\omega \mid 0, 1) = \int_{\infty}^{\omega} dN(t \mid 0, 1)$.

The following table gives the values of L for the different distributions of the family households by region and by type of area. As it was pointed out earlier, in view of the wide scatter of the households in the higher income levels, these households were not included in the distribution analysis.

LOCATION		σ log		L	
		Rural	Urban	Urban	Rural
Region:	II	0.8281	0.8080	0.4420	0.4320
	III	0.9048	0.8065	0.4778	0.4314
	IV	0.8760	0.7651	0.4648	0.4114
	V	0.7580	0.7846	0.4080	0.4212
	VI	0.9084	0.7779	0.4792	0.4176
	VII	0.8072	0.7489	0.4315	0.4038
	VIII	0.8475	0.3871	0.4520	0.2152
	IX	0.9130	0.7427	0.4818	0.4004
	X	0.7073	0.7807	0.3830	0.4190
Greater Manila		0.7063	—	0.3826	—
Overall Rural		—	0.8413	—	0.4481
Overall Urban (except Manila)		0.8243	—	0.4401	—

It can be noted from the table that in the urban areas, Greater Manila and Region X (urban area) have their measures concentration of family household incomes closest to zero — the measure for equal distribution. The highest measure of L is that one for the urban area in Region IX. In the rural areas, L is least in Region VIII; in fact among all the rural and urban areas, Region VIII (rural) has L closest to zero.

High Income Family Households

In the rural areas, about one percent of those family households receiving an annual income of more than ₱5,000, get an annnual income of more than ₱20,000; about 9 percent receive more than ₱10,000. In Region VIII (rural) where the measure of concentration I is closest to zero, no family household seems to have annual income of more than ₱7,000.

In the urban areas, out of the total family income households receiving an annual income of ₱14,000 and above, about one per cent of the family households obtain more than ₱20,000 per year; and about 8 per cent receive ₱60,000 or more per annum

In the Greater Manila Area, among those family households receiving an annual income of ₱14,000 and above, about 17 per cent get annual income of ₱50,000 or more; and about 10 per cent receive an annual income of ₱100,00 or more. The wide gap between rural poor and the urban rich can thus be noted.

Age of Heads and Income of Family Households

Table IV A, B, show the proportion of heads of family households classified by age groups and income levels. In the Greater Manila Area, the family households with young heads (Less than 40 years old) receiving ₱6,000 and above per year constitute about 6% of the total family households. In the non-Manila areas (Regions II — X), the family households with the same characteristics constitute only about 2-7 per cent of the corresponding total. It is interesting to note that the family households with "old" heads (60 years and above) receiving ₱6,000 and above constitute only about 3.4 of the total family households. In the rural areas, the family households with "young" heads receiving ₱4,005 per annum and above, constitute about 2.5 per cent of the total family households; those with "old" heads receiving the same amount per year constitute only about 0.9 per cent.

TABLE IV A. AGE OF HEADS AND INCOME OF FAMILY HOUSEHOLDS
(Entries in Per cents)

Income (Pesos)	Area	Age (Years) Below 40	40-49	50-59	60 up	Total
Below 3,000	Greater Manila	23.9	8.3	3 5	3.8	39.5
	Urban (II-X)	29.4	15.0	11.6	8.2	64.2
3,000 to below 6.000	Greater Manila	13.1	8.9	6.9	3.4	32.4
	Urban (II-X)	6.8	6.8	5.5	4.0	23.2
6,000 to below 10,000	Greater Manila	4.2	3.8	4.4	1.8	14.1
	Urban (II-X)	1.8	2.0	1.9	1.6	7.4
10,000 & above	Greater Manila	1.9	4.5	4.9	2.6	14.0
	Urban (II-X)	0.9	0.7	1.7	1.8	5.2
Total	Greater Manila	43.1	25.1	19.8	11.6	100.0
	Urban (II-X)	38.9	24.6	20.8	15.7	100.0

TABLE IV B. AGE OF HEADS AND INCOME OF FAMILY HOUSEHOLDS
(Entries in Per cents)
(Rural — Region II — X)

Income (Pesos)	Age Group Below 40	40 - 49	50 - 59	60 up	Total
Below 2,505	37.3	18.2	12.5	10.9	78.9
2,505 to below 4,005	5.3	4.1	3.1	1.5	14.0
4,005 to below 5,005	1.1	1.2	0.6	0.5	3.4
5,005 and over	1.4	0.9	1.0	0.4	3.7
Total	45.1	24.4	17.2	13.3	100.0

1. Aitchison, J. & Brown, J.A.C.: "On Criteria For Descriptions of Income Distribution", *Metroeconomics*, Vol. VI, Dec. 1954.
2. Bowman, Mary Jean: "A Graphical Analysis of Personal Income Distribution in the United States" *Readings in the Theory of Income Distribution*, 1946.
3. Davies, G.R. and Smith, R.H.: "*Probabilities in Logarithmic Skewed Distributions*", JASA, Vol. 36, 1941.
4. Davies, G.R.: "*The Logarithmic Curve of Distribution*", JASA, Dec. 1925.

The Philippine Statistician
Vol. XVIII, Nos. 1 and 2, 1969

LABOUR FORCE PROJECTIONS FOR THE PHILIPPINES AND SOME METHODOLOGICAL ISSUES INVOLVED

By M. L. GUPTA*

I. GENERAL

Reasonably accurate estimates of labour force at certain points of time are some of the basic economic data required when a developing country embraces planning as the principal instrument of development. These are especially necessary if promotion of employment or creation of job opportunities of certain order is one of the major objectives of planning, as is implicity the case in the Philippines under the Four-Year Economic Program, FY 1967-70.

2. In the Philippines, some attempts in this sphere have been made in the recent past. The labour force projections made by Gerardo Sicat and Rosa Linda Tidalgo, School of Economics, University of the Philippines were obtained on the assumption of constant participation rates by sex and age as in 1960 according to the Census results at that time.[1] In a more recent exercise undertaken at the Population Institute the above-mentioned approach of constancy of participation rates up to 1980 has been maintained.[2] A serious limitation of these estimates is that they embody a static view of the participation rates, for neither the Philippine labour force data available since 1956 support that view, nor it can be corroborated by labour force participation rates observed in other countries.

* Manpower Assessment and Planning Expert, ILO, United Nations Development Programme, Manila.

1 Population Institute, University of the Philippines, *First Conference on Population 1965, published in 1966, pages 86 and 366.*

2 This was ascertained in the writer's discussion with Director of the Population Institute.

Many significant changes taking place in the socio-economic conditions of the Philippines are bound to change, to some extent, the country's labour force structure and participation rates.

3. In the paper "Current Employment and Labour Supply in the Philippines",[3] the labour force estimates were based on the population projections adopted by the Inter-Agency Committee on Demography and on trends of participation rates derived from the results of the BCSSH[4] October surveys for 1956-1965, and the extrapolated trends of participation rates provided the basis for projections up to 1970. This obviously and more sound approach than the ones described above, but it does not bring into sharp focus the bearings on the labour force of visible socio-economic phenomena relating to special groups such as workers in their early 'teens, women, aged persons etc. whose participation in economic activity has been observed to vary considerably with educational goals and policy, social attitudes, social security measures—public as well as private, and the level of economic development in a country.

4. *The purpose of this paper is*:

(a) to examine possibilities of establishing a link between the benchmark of labour force data provided by the 1960 Census on the one hand and the corresponding PSSH data on the other,

(b) to provide an objective view of the age groups whose inclusion in the labour force is relevant in the context of manpower planning, with a consideration of the changing place

---

[3] Paper included in the volume *Asian Institute/ILO Short-Term Course on Human Resources Development and Manpower Planning in the Philippines*, January 23-February 17, 1967, as document No. 67/Philippines B. 13. This paper embodied the work of two "Technical Units" set up by the National Economic Council as part of its manpower project with which A. B. Reiz, I LO Labour Statistician, was associated in an advisory capacity.

[4] Bureau of the Census and Statistics Survey of Households (BCSSH) which prior to May, 1965 was known as the Philippine Statistical Survey of Households (PSSH).

of young workers, females and aged persons in the Philippine labour force structure, and

(c) to review the existing labour force projections up to 1970 in the light of essential variables.

II. A LINK BETWEEN THE 1960 CENSUS DATA AND THE CORRESPONDING PSSH DATA

5. It may be helpful to examine the possibilities of establishing a link between the labour force data available from the February-March 1960 Census of Population and Housing on the one hand, similar data available from the PSSH for the nearest preceding and the nearest following points of time—October 1959 and October 1960[5] on the other. Essential figures in terms of the employment status of the Philippine population are presented in Table 1.

6. The PSSH October 1960 took place seven months after the 1960 Census. However, it would be seen in the preceding Table that the population 10 years old and over was about 900 thousand *smaller* and the number of persons in the labour force was 580 thousand *larger* in October 1960 in comparison with the corresponding figures for February-March 1960. This resulted in the both sexes labour force participation rate of 53.8% in October 1960 compared ith 47.8% in March of the same year as revealed by the Census. Conversely, the proportion of persons not in the labour force was 46.2% according to the October 1960 Survey vis-a-vis 51.0% observed in the March Census. Since employed persons account for an overwhelmingly large proportion of the labour force at any time, it is only natural that the differences of the above magnitude are correspondingly reflected in the differences between the numbers of the employed estimated by these two sources. More or less, the same significant magnitude of differences is observed in the participation rates revealed by the Census and October 1959 PSSH Survey which took place only five months prior to the Census.

[5] It is October 1960 because no PSSH survey round seems to have taken place in May of that year.

TABLE I

POPULATION 10 YEARS OLD AND OVER BY EMPLOYMENT STATUS AND SEX, 1959 AND 1960

	Census, Feb. — March 1960			PSSH October 1959			PSSH October 1960		
	Both Sexes	Males	Females	Both Sexes	Males	Females	Both Sexes	Males	Females
				(IN THOUSANDS)					
Population 10 years old and over	17,873	8,931	8,942	16,463	8,114	8,349	16,957	8,383	8,574
	(100.0)	(100.0)	(100.0)	(100.0)	(100.0)	(100.0)	(100.0)	(100.0)	(100.0)
In Labour Force	8,536	6,352	2,184	9,116	5,914	3,201	9,116	6,026	3,089
	(47.8)	(71.1)	(24.4)	(55.4)	(72.9)	(38.3)	(53.8)	(71.9)	(36.0)
Employed	7,944	5,990		8,575	5,656	2,920	8,539	5,721	2,818
	(44.4)		(21.9)	(52.1)	(69.7)	(35.0)	(50.4)	(68.2)	(32.9)
Unemployed	592	362	229	540	259	282	577	306	271
	(3.3)	(4.0)	(2.6)	(3.3)	(3.2)	(3.4)	(3.4)	(3.6)	(3.2)
Not in Labour Force	9,124	2,467	6,657	7,335	2,193	5,142	7,830	2,351	5,480
	(51.0)	(27.6)	(74.5)	(44.6)	(27.0)	(61.6)	(46.2)	(28.0)	(63.9)
Labour Force Status not reported	213	112	101	12	6	6	11	(b)	(b)
	(1.2)	(1.3)	(1.1)	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)

N.B. Figures within brakets are for percentages (b) denotes less than 9,500.

Source: (i) Bureau of the Census and Statistics, Department of Commerce and Industry, Manila, Census of the Philippines, 1960, Population and Housing: Fertility and Labour Force Characteristics (Special Report) June 1965, page 2, and

(ii) Bureau of the Census and Statistics. Manila. The Philippine Statistical Survey of Household Bulletin, Series No. 7 Labour Force and Socio-Economic Data, May and October 1959, page 10, and Series No. 8, October 1960 page 8.

7. Notwithstanding the differences cited above, it is remarkable to notice a close similarity between the numbers as well as proportions of the males (from these two sources of information) in regard to characteristics such as population, labour force employed and persons not in the labour force. On the other hand, differences in the corresponding numbers of females from the two sources appear to be sizeable. For instance, according to the Census, only 24.4% females 10 years old and over were in the labour force. The comparable proportion shown by the PSSH was about one and one half times of the Census result (36.0% in October 1960 and 38.3% in October 1969). It will be interesting to see if age-specific participation rates, especially for the females, throw light on the significant differences just mentioned. The age-specific participation rates by sex are presented in Table 2.

It is evident from the preceding Table that the age-specific labour force participation rates revealed by the Census and PSSH for males are very close to one another in the age groups (65 years and over) males, the participation rates from the two sources have some difference, but this age group is numerically unimportant in the labour force. However, in respect of females the participation rates appear to be wide apart—for the age group of 10-24 years: 21.9% according to the 1960 Census and 32.6% in the October 1959 Survey findings; for the next age group 25-44 years: 26.6% according to the Census and 43.6% according to the Survey, and so on for the remaining age groups.

8. It is worth noting that there was a similarity in concepts and definitions used in these two major sources of labour force data, and a common agency (the Bureau of the Census and Statistics) conducted the Census and Survey operations. Nevertheless, such differences as have been noticed may probably be ascribed to (i) the sizes of the two samples[6] and (ii)

[6] As indicated earlier in this Paper, the Census of 1960 was reduced to its one-tenth size to constitute a sample for purposes of gathering data on labour force and fertility.

TABLE 2

AGE-SPECIFIC LABOUR FORCE PARTICIPATION RATES BY SEX, ACCORDING TO THE 1960 CENSUS AND OCTOBER 1959 PSSH

	February-March, 1960 Census			October 1959 PSSH		
	Population	Labour Force	Participation Rate	Population	Labour Force	Participation Rate
	(In thousands)		%	(In thousands)		%
Both Sexes						
10-24 years	8516	3036	35.6	7687	3221	41.9
25-44 years	6006	3572	59.5	5345	3672	68.7
45-64 years	2636	1616	61.3	2689	1931	71.8
65 years and over	715	312	43.6	742	294	39.6
Total	17873	8536	47.8	16463	9121	55.4
Males						
10-24 years	4292	2110	49.1	3859	1976	51.2
25-44 years	2934	2755	93.9	2499	2429	97.2
45-64 years	1353	1253	92.6	1377	1296	94.1
65 years and over	352	235	66.9	379	215	56.8
Total	8931	6352	71.1	8114	5915	72.8
Females						
10-24 years	4224	926	21.9	3828	1248	32.6
25-44 years	3072	818	26.6	2846	1241	43.6
45-64 years	1283	363	28.3	1312	634	48.3
65 years and over	363	77	21.2	363	78	21.6
Total	8942	2184*	24.4	8349	3198*	38.6

* The magnitude of difference between these two estimates is 1 million, surprisingly denoting a decrease of 1 million in female labour force, by Feb.-March 1960 Census only five months after the October 1959 Survey.

N.B. Due to rounding, figures and percentages do not always add up to the independent totals.

Source: Bureau of the Census and Statistics, op. cit.

the possibility of inadequate care in gathering economic data in the large operations of the Census through less-trained enumerators than might be the case for the PSSH.

9. The Director, Bureau of the Census and Statistics, in a recent technical note[7] states that "although estimated sampling variances were observed to be small in the usual surveys, they do not really indicate the total variations since the component of these due to the primary sampling unit (that is, the precincts or barrios) is excluded". It is further stated that the "frame, which is really a sample of the 'population' is fixed from one survey to another. To include the variation due to the primary sampling unit would mean to draw a new frame in every survey which would be prohibitive from the point of costs and time[8]

TABLE 3
UNPAID FAMILY WORKERS BY INDUSTRY, 1959 & 1960
(In thousands)

	Feb.-March 1960 Census[9]			October 1959 Survey[10]		
	Both Sexes	Males	Females	Both Sexes	Males	Females
1. Agriculture, livestock hunting, forestry & fashing	1764	1192	572	1844	1029	815
2. Mining & quarrying	—	—	—	—	—	—
3. Manufacturing .	113	33	80	120	16	104
4. Construction ...	3	3	—	1	1	—
5. Electricity, gas, water & sanitary services	—	—	—	—	—	—
6. Commerce	65	21	45	125	30	95
7. Others*	13	6	7	21	7	14
Total	1960	1255	705	2111	1083	1028

* Includes Transport, Storage and Communication Services and Activities not adequately described.

[7] Tito Mijares: Comments on the Paper: *"Labour Force Trends and Employment Target"* by M. L. Gupta and B.A. Perez, contributed to the Second Population Conference held at Manila during November 27-29, 1967.

[8] *Ibid.*

[9] Bureau of the Census and Statistics, *Fertility and Labour Force Characteristics (Special Report)*, *op. cit.* pp. 15-16.

[10] Bureau of the Census and Statistics, Data derived from Table 14 in *PSSH Bulletin Series No. 7, Labour Force and Socio-Economic Data, May and October 1959.*

It is his view that the interpretation of the concepts by the Census enumerators introduced some sizeable bias into the Census figures or resulted in a wrong classification of person enumerated. This can be illustrated with the help of the figures in the Table below:

10. A glance at the figures of the preceding Table indicates that in the 1960 Census, the numbers of unpaid family workers were smaller than those observed according to the October 1959 Survey held five months earlier. One might suspect that the enumerators classified many unpaid family workers as persons not in the labour force. Otherwise, there could not be that order of dimunition in the numbers of unpaid family workers, especially in manufacturing, commerce and 'others' within such a short duration of a few months. Some evidence of this discrepancy can be had by a comparison in the February-March 1969 Census and the October 1959 Survey for person not in the labour force.

TABLE 4

PERSONS NOT IN THE LABOUR FORCE BY SEX, 1959 & 1960
(In thousands)

Persons not in the Labour Force	1960 Census[11]			October PSSH 1959 Survey[12]		
	Both Sexes	Males	Females	Both Sexes	Males	Females
Housekeeping	5007	152	4855	3160	86	3074
Studying	3249	1733	1516	3132	1630	1502
Retired	79	49	31	382	160	224
Disabled	142	76	66	152	92	59
Others	646	457	189	368	197	171
Labour Force Status not reported ...	213	112	101	(b)	(b)	(b)
Total	9124	2467	6657	7200	2166	5033

"(b)" represents figure less than 9,500.

11. It is apparent from Table 4 that the total number of persons not in the labour force, shown by the 1960 Census increased

[11] Bureau of the Census and Statistics, *Fertility and Labour Force Characteristics (Special Report)*, *op. cit.* pp. 1-2.
[12] Bureau of the Census and Statistics, *The PSSH Bulletin Series No. 7, October 1959*, *op. cit.* Table 42, page 52.

by 1.9 million, representing more than one-fourth of the total persons recorded in this category only five months earlier in October 1959 according to the Survey. In the major constituent category "housekeeping" alone, the increase was of the order of about 1.8 million or the number of persons in this category increased by about 60% within the five months. The Director, Bureau of the Census and Statistics, is of the view that in the Census enumeration, there was a bias towards women having been accepted as dependents. Thus many of them who would get included in the labour force in the October 1959 Survey were not counted as such in the 1960 Census and swelled the numbers of persons not in the labour force. "Housekeeping activity, in particularly, during the Census enumeration removed from the labour force many women who otherwise were classified to be in the labour force by the survey method of questioning."[13]

12. The above analysis provides an explanation of the major discrepancies between the Census and Survey regarding the labour force estimates and participation rates. "The need for linking the sample survey results with the 1960 Census totals and for re-designing the sample with the corresponding 1960 enumeration of household population as frame was suggested as early as the time of planning of the said national census. The revised sample design was not effected until the May 1965 survey, and since the desirable linking of the existing survey figures with the Census population series was never carried out, the available data do not provide a reliable measure of trend in the labour force estimates since 1956, and over the new set of estimates in 1965."[14]

---

[13] Tito A. Mijares, *op. cit.*

[14] M.L. Gupta and B.A. Perez, *"Labour Force Trends and Employment Targets," op. cit.*, mimeographed, page 4.

III. CHANGING ROLE OF YOUNG WORKERS, FEMALES AND AGED PERSONS IN THE PHILIPPINE LABOUR FORCE STRUCTURE

Young Workers:

13. The Philippine economic data including manpower statistics collected primarily by the Bureau of the Census and Statistics are for persons 10 years old and over for the age groups 10-24, 25-44, etc. The age group 10-24 may be considered as a heterogeneous age group in so far as it includes bulk of the students at various levels of education and persons entering the labour market, especially in their later teens and early 'twenties.' In many countries, bifurcation of this broad age group into 10-14 years, 15-24 years, or the like helps to analyze the labour force attachment of persons in these years of age. In the BCSSH, information for the age group 10-14 years is limited to employed persons at work in agriculture and non-agricultural industries by number of hours worked during the Survey week and by sex. For this age group additional information in terms of certain other essential labour force characteristics is required along with calculations of 'Standard Error' before very definite views can be expressed regarding the nature of employment and labour force disposition of these youngsters. These detailed data relating to the young workers are not available at present. However, several definite factors can still be stated which may set us thinking about the advisability of inclusion or non-inclusion of this group of persons in the labour force. These factors are presented below:

a) Persons in the age group 10-14 should be considered too young to work, at least too young for commitment to the labour force. Of the persons not in the labour force who gave the reason of being 'too young' for not wanting work in the October 1966[15] round of the BCSSH, 71% (as would be expected) were in this age group. The next higher age group, 15-19 years, accounted for 29% of such persons, denoting that even in the

[15] Data for the October 1966 Survey are under print. These (figures stated here and further in this Paper) were made available to the writer by the courtesy of the Bureau of the Census and Statistics.

next five-year higher age bracket many consider themselves premature for entrance in the labour market.

b) In the interest of the community, there are severe legal restrictions on the employment of children under the Philippine Republic Act 679 (as amended by Rep. Act No. 1131) commonly known as the Woman and Child Labor Law. This law lays down that no child below 14 years of age shall be employed or permitted or suffered to work on school days in any shop, factory, commercial, industrial or agricultural establishment or any other place of labour unless such child knows how to read and write. There are minor exceptions to this legal provision in terms of light work which is not harmful to the children's health or normal development and which is not such as to prejudice their attendance in schools.

c) In the Philippines, most children in the age group 10-14 are studying in any case. Referring again to the October 1966 results of the BCSSH, of the person not in the labour force who stated 'going to school' as the reason for not wanting work, about 69% were in this age group, and nearly 25% were even in the next age group 15-19 years.

c) In the Philippines, most children in the age group 10-14 are studying in any case. Referring again to the October 1966 results of the BCSSH, of the person not in the labour force who stated 'going to school' as the reason for not wanting work, about 69% were in this age group, and nearly 25% were even in the next age group 15-19 years.

d) The Philippine law for compulsory education applies to the children of ages 7 to 11 years; but educational policy of the country is even more progressive in the sense that it envisages every citizen to have at least an elementary education six years which normally can be completed around the age of 13 years.

e) The inclusion of age group 10-14 in the labour force gives an exaggerated view of unemployment in the country. For instance, according to the 1960 Census, of the total 591-550 unemployed, 84,750 unemployed were in this age group alone. This represented about 14% of the unemployment aggregate at that time. In October 1966, the unemployed in this age group were 77 thousand out of the all ages unemployed total of 821 thousand. When the attachment of these youngsters to labour market is rather nominal, there is not much point in counting those of them 'looking for work' as unemployed.

f) One gets an underrated view of the educational attainment of the labour force in the country by including the persons of ages 10-14 most of whom would be still in schools. To say, for instance, that in October 1965 (when special data were collected in this respect), 12.7%[16] of the total population 10 years old and over had at least a secondary education amounts to belittling the educational attainment of the working-age population. What can be the point in relating secondary education attainment to persons right from the age of 10, when no one can complete secondary education at that age?

g) Most of the persons in the labour force of ages 10-14 are unpaid family helpers in agricultural and allied pursuits in the rural areas. Unpublished data of the BCSSH for October 1966 show that over 70% (437 thousand out of 627 thousand) of employed persons of 10-14 years' age were unpaid family workers. Of the remaining less than 30%, most would be working as errand boys, housekeepers or maids, etc. Such economic activities are motivated to obtain some supplements to low earnings of their elders in poor families.

h) Most of the persons enumerated in the labour force from the age group 10-14 are part-time or nominal workers. According to the results of the October 1966 BCSSH, about 38% of the employed persons in this age bracket were working less than 20 hours a week. This point may derive strength by comparison of the position of this group with that of the persons of ages 15-19 years as shown in Table 5 below:

[16] This is obtained by totalling 5.8% for those who completed at least High School fourth year and 6.9% who went beyond the High School stage to different college years completed. vide Table 34 in the *BCS Survey of Households, Series No. 19, Labour Force including Educational Attainment Data, October 1965, page 35.*

TABLE 5

EMPLOYED PERSONS OF THE AGE GROUP 10-14 AND 15-19 YEARS AT WORK BY SEX, OCTOBER 1966

(In thousands)

	Both Sexes		Males		Females	
	10-14 years	15-19 years	10-14 years	15-19 years	10-04 years	15-15 years
(A) Total number of employed persons	627	1,594	386	952	241	642
(B) Of whom working less than 20 hours a week	238	198	133	97	106	102
(B) as % of (A)	37.9%	12.4%	34.5%	10.2%	43.9%	15.9%

(*Source*: Bureau of the Census and Statistics; data worked back from the Bureau's relevant figures).

It can be seen in the foregoing Table that only about 12% of the persons of the ages 15-19 years had limited work of less than 20 hours a week as against about 38% (which is about trice) doing such limited work among the employed persons of 10-14 years. Distributed by sex, this magnitude of difference is well reflected in the Table.

i) Introducing a seasonal factor in this reasoning, reference may now be made to the labour force attachment of the persons 10-14 years in the months of May and October of the same year 1966. This will be clear from Table 6.

TABLE 6

HOUSEHOLD POPULATION 10-14 YEARS OLD BY EMPLOYMENT STATUS AND SEX: MAY 1966 AND OCTOBER 1966

(In thousands)

Employment Status	October 1966			May 1966		
	Both Sexes	Male	Female	Both Sexes	Male	Female
10-14 years old						
Total population ..	4469	2316	2153	4294	2152	2142
Total labour force	704	422	282	986	643	343
Employed (fully or partially) ..	627	386	241	831	580	300
Totally unemployed	77	36	41	106	63	43

(*Source*: Bureau of the Census and Statistics: data compiled and supplied on special request of the writer).

It would be seen in the preceding Table that the number of persons of 10-14 years in the labour force swells in the month of May (986 thousand as against 704 thousand in October) when they get school vacations and when many of them are obliged to look for some pretence of work to augment low earnings of their poor parents. A glaring example of distortion in the labour force estimates, due to the inclusion of persons 10-14 years of age, is provided by the May 1967 labour force estimate of 13.2 million (representing an increase of 11.2% in one year alone) compared with 11.9 million in May 1966. The Bureau of the Census and Statistics has the explanation that in the May 1966 Survey, the school year 1965-66 was up to May 7, 1966 when an estimated 1.2 million persons were reported not in the labour force because they were then attending school. However, the school year 1966-67 ended on April 29, 1967, and the interview for the May 1967 Survey began on May 15. As such, persons 10 years old and over were, more or less, free to join (and, in fact, would have joined) the work by participating in actual economic activities or by looking for work during the long school vacation.[17]

14. The sum total of the reasoning advanced above may be that the labour force attachment of the early 'teenagers is nominal or very weak most of them are just unpaid family workers and further a very high proportion of them is in the category of part-time or nominal workers, working less than 20 hours a week. Their employment is not substantial. They happen to be in the labour force (and away from schools) when school facilities are inadequate on the one hand and their parents are hard-pressed for money on the other. When there is a marginal improvement in poor parents' incomes and when schooling facilities are well within reach, there is a substantial diversion of these children to schools and this is what seems to be happening at present in the Philippines. *Even granting that in the present socio-economic situation of the country many of these youngsters will continue to do this type of nominal work, from the point of view of public policy, for promotion of*

[17] Bureau of the Census and Statistics, (mimeographed) *Special Release No. 47*, pp. ii-iii.

employment objective and from the standpoint of manpower planning and training, on the whole, it seems appropriate to exclude persons of ages 10-14 from computations of labour force estimates and projections.

Aged Persons:

15. The Philippine labour force estimate include persons of 65 years ad over, besides person of the ages 60-64 included in the broad age group 45-64. In general, the old workers may decide to withdraw from labour force earlier than usual depending upon the economic assistance they expect from their younger relations and/or the institutional assistance provided by the community in the form of old-age pensions and the like. Such a development in the Philippines can be envisioned by reference to the Social Security Act 1966.[18] Section 18 of the Act provides retirement benefits stating that "upon reaching the age of 60 years and after having paid at least one hundred twenty monthly contributions to the System —a covered employee shall have the option to retire and shall be entitled, for as long as he lives but in no case for less than five years, to a monthly basic pension amount to be computed. . . ." In addition, "the suspension of retirement pension upon reemployment is limited only to those below 65. . . ."[19] It will be seen that, in the interest of leisurely life for aged persons, the suspension of pension benefits (if reemployment is secured between the ages of 60 and 65) will dissuade aged persons to reenter or remain in the labour market which they had to do until recently. The 1967 Annual Report of the Social Security System states that the benefit payments amounted to ₱28.3 million as compared with ₱18.3 million in 1966, and that the upsurge in benefits is attributable, among other things, to the steady increase

[18] The Social Security law was enacted in the Philippines in 1954 and has been amended since then in 1957, 1960, 1963, 1965 and 1966. The amendment of 1966 has enlarged the retirement benefits by 40%. The retirement benefits in 1966 were ₱3.9 million (about 21% of ₱18.3 million total benefits for death, disability, sickness and retirement). In 1967, these benefits were still higher at ₱6.8 million (about 24% of 28.3 million total benefits for death, disability, sickness and retirement).

[19] Social Security System, Annual Report, 1966, page 29.

in retirement beneficiaries. The same Report further states that the biggest portion of all social security payments will eventually go to members qualified for retirement benefits. Furthermore, the Act's coverage is not negligible so as to render this effect insignificant in practical life. The Social Security System is compulsory upon all employees not over sixty years of age and their employers except those in the Government Service,[20] domestic service, casual employment, and service or undertaking where there is no employer-worker relationship. The system covers all private employers, whether in business as single proprietorship, partnership or corporation, and all wage earners not otherwise excluded by law, be they engaged in agriculture, commerce or industry. (Major groups that are at present not yet covered are agricultural workers who are not wage-earners, the self-employed and the domestic help). The mass of SSS members consists of non-agricultural wage-earners between ages 15 and 64 years.[21] Already by the end of 1967, about 1.9 million employees (representing about one-sixth of the total labour force in the Philippines) were covered and entitled among others, to receive the retirement benefits. *The implication of this factor in labour force projections should be kept in view and a lower than the present proportion of older workers should be expected in the labour force of years to come.*

Females' Changing Role in the Labour Market:

16. The labour force structure and participation rates are influenced by the proportions of males and females of working ages in the labour force. Inter-country comparisons show that the males' participation in economic activities is remarkably similar. However, the females' joining the labour force is governed by a host of socio-economic factors. It may be stated in passing that some features of female labour force are univer-

[20] But for the Philippine Government employees, already another system of social security including retirement benefits is available.
[21] *Social Security System, Annual Report 1967, pages unnumbered in the Report.*

sal. For instance, broadly speaking, their attachment to labour force is weak almost everywhere. Females enter and reenter the labour force in the course of working age, conditioned as they are by marital status and child-bearing responsibilities. Another factor of widespread nature is the urbanization trend. Affected by the mass media of advertisement, the urge for conspicuous consumption and strong desire for improvement in the family living standards are widespread because of which many women enter the labour market. Apart from these factors of, more or less, universal validity which have their bearing on the extent of females' participation in economic pursuits, there are some factors special to the Philippines which may be mentioned below.

17. On an average, a married Philippine female has seven children during the child-bearing age. Among work-oriented females, ordinarily, this demographic feature should have the effect of their leaving (and rejoining) the labour market several times to take care of babies until they are put to school. In reality, this often does not take place in the Philippines. Among families of low stratum, if a mother works and she already has a few children, the possibility is that the eldest daughter would be withdrawn from school to take care of younger children. In well-to-do families where a mother may be a white-collar worker, the child care is often entrusted to maids or domestic helpers whose number can always be increased without much financial burden because of cheap wages. Women's participation in outdoor activities, especially wage employment, depends upon social attitudes and general acceptance for outdoor work. *It seems that the Philippine society has already developed a system in which females' participation in all spheres of work where they can be suitable is commonly accepted.*

18. Another factor which affects females' participation rates is the extent of education in the country. It is remarkable that the proportions of literate and educated males and females in the Philippines are very close to another. For instance, according to the October 1965 BCSSH, 63.4% all males 10 years old and over had at least Grade one education. The corresponding

percentage for females was 64.4. Further 14% of the males, 10 years old and over, in comparison with 11.4% females of the same age group had finished at least high school education. Over the last few years, the educational attainments of both sexes have advanced, and in many families educated girls and women work with the growing understanding that they should utilize their education. The growth of the tertiary sector of the economy (suitable for females' employment) absorbs increasing number of women, and it may be stated that more and more females are being enticed to work. The work in the tertiary sector is, generally speaking, light and less ardous and many married women are attracted to it, especially on part-time basis. Often the impression is held that women are, by and large, supplementary earners. While the fact of a much larger proportion of females than males being part-time (meaning below 40 hours a week) workers cannot be denied (there were some 42.8% of female workers as part-time workers against 71.4% of such male workers according to the October 1966 BCSSH), during 1962-65 a significant increase has taken place in the number of females engaged in full-time work. During October 1965-66, this enhanced rate of full-time work on the part of female labour remained stable. The relevant figures may be seen in Table 7 which follows.

TABLE 7

FEMALES WORKING 40 HOURS AND OVER DURING THE SURVEY WEEK, OCTOBER 1962-1966

	% of total employed females
October 1962	46.9
October 1963	48.9
October 1965	57.8
October 1966	57.3

(*Source*: Bureau of the Census and Statistics Bulletin Nos. 13, 16 and 19, and underprint report for October 1966)
N.B. No Labour Force Survey results are available for October 1964.

19. The general economic conditions in a country as well as income of husbands or other male members in "extended

TABLE 8

FEMALES' TOTAL EMPLOYMENT AND EMPLOYMENT IN SELECT INDUSTRIAL GROUPS, 1960-1966

	Oct. 1960	Oct. 1961	Oct. 1962	Oct. 1963	Oct.[12] 1965	Oct. 1966	Increase in October 1966 over Oct. 1960 Nos.	%	Average Annual increment
1. Agriculture	1198	1382	1590	1533	1232	1523	325	27.1	4.5
2. Manufacturing	617	614	606	653	566	645	28	4.5	0.8
3. Commerce	448	515	546	621	654	677	229	50.0	8.3
4. Government, Community, Business and Recreational Services	149	214	214	226	260	308	159	106.7	17.8
5. Domestic Services	273	315	312	293	431	420	147	53.8	9.0
Total Employed Females	2818	3163	3384	3492	3296	3742	924	32.8	5.5

(**Source:** Bureau of the Census & Statistics, Labour Force Bulletin Nos. 8, 10, 16, 19 and underprint report for October 1966).

[12] The October 1965 results are not strictly comparable with the results of the earlier years for the same month because of changes in the sample design introduced in May 1965. The BCSSH has reported smaller absolute number of female labour force (3,608 thousand in October 1965 as against 3,710 in October 1963) which is highly improbable. Even then, the percentage changes relating to employment for select industrial segments are very considerable. But for the apparent lower estimates of BCSSH in October 1965, the real employment increases in these industrial segments would have been even higher and reinforce the point made above.

families" also have some bearing on the extent of females' participation in economic activities. Over a period of five or ten years in the Philippines, these factors do not appear to be significant.

20. It is generally held that the country has had favourable economic situation during the last few years and this factor might have been responsible for not too stiff a competition for jobs between males and females with the result that females in large numbers entered the labour market and, in fact, improved their position in the tertiary sector. Over the years 1961 to 1966, substantial increases in females' employment in *select* segments of the tertiary sector can be noticed as shown in Table 8.

21. It would be seen in the preceding Table that during the six-year period, October 1960-66, females' employment increased by 50% in Commerce, about 107% in government, community, business and recreational services (thus more than doubling itself in six years), and about 54% in domestic service. These are some of the leading wage employment areas (except that part of "Commerce" which is organised as family retail business). Females' employment grew in agriculture, although modestly by 27% and in manufacturing although nominally by 4.5% over the six-year period. With reference to the spectacular rise of their employment in Commerce: Government, Community, Business and Recreational Services; and Domestic Services, it seems to be a fact that an increasing number of females are looking for and will seek wage employment and enter the labour market in the Philippines.

22. Ignoring the changes in the sample design of the BCSSH which were introduced in May 1965, sometimes comparisons are made in retrospect. In doing so, the picture presented is as shown in Table 9.

TABLE 9

TOTAL AND FEMALE LABOUR FORCE, 1960-66

	Oct. 1960	Oct. 1961	Oct. 1962	Oct. 1963	Oct. 1965	Oct. 1966	Increase during 1960-66	Average annual increase
	(IN	THOUSANDS)					%	%
Labour force								
Females	3089	3479	3740	3710	3608	4149	34.3	5.7
Males	6027	6234	6525	6523	7156	7608	26.2	4.7
Total	9116	9713	10266	10233	10764	11757	28.9	4.8

(Source: Bureau of the Census and Statistics, op. cit.)

If the Survey results over this period of six years are taken without a critical view and year-to-year fluctuations of total numbers in the labour force are not given much credence (part of the fluctuations are due to sampling errors and reporting biases), we get a broad picture of total female labour force rising from 3,089 thousand in October 1960 to 4,149 thousand in October 1966—a rise of 34.3% over the six-year period or an annual increase of 5.7% compared with the corresponding increase of 4.7% in the male labour force. The year 1965-1966 witnessed an increase of about 1 million in the total labour force, of which females accounted for about half million which is considerably more than their usual one-third share in the total labour force. This is an unprecedented and record in increase in the female labour force for one year.

23. This analysis may now be seen in terms of the overall participation rates of females since the beginning of this decade. Essential figures in this respect are presented in Table 10.

TABLE 10

FEMALE POPULATION 10 YEARS OLD AND OVER BY EMPLOYMENT STATUS, 1960-1966

	Oct. 1960	Oct. 1961	Oct. 1962	Oct. 1963	Oct. 1965	Oct. 1966
1. Total females 10 years old and over	8574	8816	9078	9359	10219	11757
2. In the Labour force ...	3089	3479	3740	3710	3608	4149
3. Participation rate	36.0%	39.5%	41.2%	39.6%	35.5%	38.9%

(Source: Bureau of the Census and Statistics, op. cit.)

It is evident from the above Table that the female participation rate steadily increased up to 41.2% in October 1962 from 36.0% in the same month of 1960. In October 1963, it slightly decreased, and by October 1965 it dropped to 35.3%. One may surmise that this decline in the females' participation rate is not real but probably caused by changes in the design and associate factors in the sample Survey, especially after the changes of May 1965. In any case, the 38.9% participation rate observed for October 1966 denotes a rise again in the female labour force participation. Despite this erratic picture of the females' participation rates from year-to-year in the recent past, one cannot miss the fact that an increase took place in most of the year under consideration, and there appear unmistakable signs of an overall increase in the females' participation in economic activity during the six-year period.

24. The patterns of economic activity are known to differ between urban and rural areas, especially among females. Limitation of space does not allow a presentation of relevant statistical details and analysis thereof. However, a *major conclusion that emerges after an examination of the pertinent figures is that the female participation rates in the Philippines are not lower in the urban areas compared with the rural areas, as seem to have been observed in some other Asian countries. As such, it may not hold good in the case of the Philippines, for some time at least, that with the rising tempo of urbanisation the participation rates of females for the country as a whole will have a downtrend.*

25. An interesting analysis of the females[23] participation rates can be carried in terms of their grouping by educational attainments. Recently, the October 1965 BCSSH undertook an additional investigation of educational attainments of the population and labour force. The results of this Survey can be utilised

[23] Since almost universally, men are the prime bread winners, differences in their educational attainments make virtually no difference

TABLE 11

LABOUR FORCE PARTICIPATION RATES BY SEX AND EDUCATIONAL ATTAINMENT: URBAN AND RURAL: OCTOBER 1965

	TOTAL (PHILIPPINES)			URBAN			RURAL		
	All persons 10 yrs. old and over	In the labour force	Labour force partici-pation rates	All persons 10 yrs. old and over	In the labour force	Labour force partici-pation rates	All persons 10 yrs. old and over	In the labour force	Labour force partici-pation rates
FEMALES	(In thousands)		%	(In thousands)		%	(In thousands)		%
1. No grade completed	1419	625	44.0	264	92	34.8	1156	533	46.1
2. Up to Elementary grade 5 only ..	4614	1391	30.1	1113	346	31.1	3500	1045	29.9
3. Elementary (grade 6) completed & those up to third year High School	3025	1050	34.7	1229	446	36.3	1796	605	33.7
4. High School completed & those up to third year college completed ..	846	292	34.5	575	219	38.1	272	73	26.8
5. Fourth year college completed and higher	313	250	79.9	229	177	77.3	85	73	85.9
TOTAL FEMALES	10217	3608	35.3	3409	1280	37.5	6809	2328	34.2
MALES									
1. No grade completed	1081	897	83.0	142	101	71.1	940	796	84.7
2. Up to Elementary grade 5 only ..	4585	3011	65.7	955	494	51.7	3630	2517	69.3
3. Elementary (grade 6) completed & those up to third year High School	2970	2139	72.0	1056	691	65.4	1914	1448	75.1
4. High School completed & those up to third year college completed ..	1090	825	75.7	684	521	76.2	406	304	74.9
5. Fourth year college completed and higher	315	285	90.5	249	223	89.6	66	59	89.4
TOTAL MALES	10043	7156	71.3	3087	2033	65.9	6956	5123	73.7

N.B. At some places figures do not add up to their totals due to the rounding.
(Source: BCSSH, Oct. 1965, op. cit.)

by grouping the school grades into such educational levels which correspond to the employers' prevalent consideration of educational qualifications for determination of job levels in various occupations. As such, for the present analysis, these groups are made as:

a. No grade completed.
b. Up to Elementary grade 5 only.
c. Elementary (grade 6) completed and those educated up to third year High School.
d. High School completed and those educated up to third college, and
e. Fourth year college completed and higher.

In terms of these groups, the next Table presents the participation rates by sex and educational attainment for the country as a whole as well as for the urban and rural areas.

26. It is evident from Table 11 that for the males the participation rates for the three lower educational levels (1 to 3) are somewhat lower in the urban areas compared with the rural areas. At the educational levels 4 and 5, the male participation rates are close to one another in the urban and rural areas. However, the situation is different in respect of females whose participation rates at educational levels 2, 3 and 4 happen to be higher in the urban than in the rural areas. *The main inference from these figures in the present context is that with the growing urbanisation and even with increasing number of schooling years, there is hardly any possibility of the female participation rates decreasing in the country.*

27. In the earlier mentioned paper "Current Employment and Labour Supply in the Philippines", on the basis of trends and projections of females' participation rates a conclusion has been drawn that their overall participation rate will decline

TABLE 12

AGE—SPECIFIC LABOUR FORCE PARTICIPATION RATES OBSERVED IN THE BCSSH, AS ASSUMED BY THE NEC TECHNICAL UNIT & AS AMENDED HERE

(IN PERCENTAGES)

	Age Groups	Participation rate estimates of the Technical Unit of NEC advised by A.B. Reisz, for 1970	BCSSH participation rates		Range of participation rates observed during October 1956-1966	Participation rates most likely to prevail in 1970 in the writer's opinion
			Oct. 1965	Oct. 1966	by PSSH (BCSSH)	
Males						
	10-24 years	41.7	47.9	48.0	41.9 to 56.1	45.0
	10-24 years	98.7	97.3	97.6	95.9 to 97.9	98.7*
	25-44 yeears	98.2	94.9	94.9	90.1 to 95.1	98.2*
	25-44 yeears	58.2	56.5	57.3	47.4 to 57.3	55.0
Females						
	45-64 years	28.3	29.2	32.3	29.2 to 35.9	30.0
	45-64 years	43.7	42.7	47.4	41.4 to 48.9	43.7*
	65 years and over	44.3	43.2	45.7	43.2 to 50.2	44.3*
	65 years and over	22.7	20.9	22.6	16.6 to 22.6	20.0

* These participation rates based on trend values are being used as arrived at by the NEC Technical Unit, without any alternation because the labour force disposition of persons in these age groups, more especially among males, has been observed to be rather stable universally over a medium period of time.

to 35.1% by the year 1970. This has been attributed to "the growth of the number of children cared for by mothers and female relatives". On the basis of the reasoning just concluded regarding the expected pattern of the women's participation rates, it may appear that the downtrend worked out in the projections of that paper may not be borne out by facts.

IV. CONCLUSION AND REVISED LABOUR FORCE PROJECTION FOR 1970

28. Except exclusion of the age-group 10-14 years from the labour force computation which can be done only after an official decision of the Philippine authorities, the implication of the analysis in this Paper is intended to be translated into statistical terms of the labour force projections, by making suitable adjustments in the anticipated rates of participation in economic activity as presented in Table 12.

29. In Table 12, it can be seen that for males of 10-24 years, the participation rate has been raised to 45.0% in comparison with the NEC Technical Unit figure of 41.7%. During the next few years up to 1970, the participation rate of males of this group will undoubtedly decline because of longer average schooling possibilities as well as increasing proportion of children at schools. However, this decline is expected not to exceed 3 points, so that compared with the 48% participation rate (in two consecutive years 1965 and 1966 according to BCSSH) it may decrease to 45%. For females in this age group, whereas the above-mentioned tendency applicable to males will prevail, there will be the added factor of increasing number of girls entering the labour market before marriage. Hence, their participation rate may be about 30% in 1970 as against 32.3% observed by BCSSH for October 1966.

3. For old persons (65 years and over), there are good indicators (such as the increasing impact of the Social Security System, moderately rising levels of family incomes, etc.) to

TABLE 13
REVISED LABOUR FORCE PROJECTIONS FOR 1970

	Population Estimate by Inter-Agency Demography as of July 1 Committee on	Adjusted Labour participation rates Force	Labour Force
Males	(In thousands)	%	(In thousands)
10-24 years	6199	45.00	2790
25-44 years	4144	98.7	4090
45-64 years	1867	98.2	1833
65 years & over	457	55.0	251
TOTAL	12667	70.7	8964
Females			
10-24 years	5972	30.0	1792
25-44 years	4099	43.7	1791
45-64 years	1924	44.3	852
65 years & over	517	20.0	103
TOTAL	12512	36.3	4538
		TOTAL LABOUR FORCE	13502 or 13.5 million

suggest that their participation in economic activity will decline by 2 points—to 55% in case of males and 20.0% for females (as against 58.2% and 22.7% for males and females respectively projected by the NEC Technical Unit). The sulting projections of the labour force in 1970 may then be as summarized in Table 13.

31. As of July 1, 1970, the Philippine labour force is estimated to be 13.5 million, which may be corroborated by the results of the BCS Survey of Households taking place in the second half of 1970 in October. It may be recalled that the purpose of this Paper was not just the labour force projection itself for the year 1970 (although this exercise is relevant and important in itself), but even more than that the Paper aimed to bring out and analyse the Philippine non-demographic (socio-economic and cultural) factors' impact on the composition and structure of the country's labour force in the years to come. This analysis may be especially helpful in the determination of long-term labour force growth, as also informulating a suitable employment policy which takes int account the job needs and expectations of different segments of the labour force.

The Philippine Statistician
Vol. XVIII, Nos. 1 and 2, 1969

THE USE OF THE COEFFICIENT OF VARIATION IN SAMPLE SIZE DETERMINATION AND ALLOCATION[1]

By

CEFERINO S. SINSIOCO[2]

1. *Introduction.* Let Y_j, $j = 1, 2, \ldots, p$ be population characteristics which are under consideration. Suppose that $\pi_1, \pi_2, \ldots, \pi_p$ are subpopulations in a population π (Note here that π_i π_j may be empty). It is desired to make estimates of Y_j; $j = 1, 2, \ldots, p$ from a sample to be obtained from π. But no information is available on the population relevant to these characteristics under consideration. Hence, it is difficult to decide on a sample size that will insure sufficiently reliable estimates. So, ancillary variables, some information of which is available, are made use of.

Let $X_1, X_2, \ldots, X_r$ be r variables ($r \geq 1$) known to be related to Y_j, $j = 1, 2, \ldots, p$.

Let $Y_j = f(X_1, X_2, \ldots, X_r)$.

The problem is to devise a method by which a sample size n can be allocated to $\pi_1, \pi_2 \ldots, \pi_p$ to achieve sufficiently reliable estimates for all Y_j; $j = 1, 2, \ldots, p$, with $n = \sum_{j=1}^{p} n_j$. This method will be of use to planners of multi-purpose sample surveys in determining sample size and in determining how sample allocation can be made to different sections of the population under study to achieve sufficiently precise estimates of a number of characteristics under study.

[1] Excerpts from a Master's thesis.
[2] Instructor in Statistics, Statistical Center, University of the Philippines.

1. *The Use of the Coefficient of Variation As a Measure of Relative Variability.* Norris (1936) observed that the coefficient of variation ($\frac{Eu}{\sigma u}$) is an efficient statistic for distributions well graduated by the normal or the the so-called Pearson Type VII curve. But for distributions other than the normal, the maximum likelihood estimate of relative dispersion which he developed is more efficient than the coefficient of variation. He admitted, however, that there are differences of opinions as to whether or not the degree of added efficiency achieved by the use of these likelihood estimates will merit departing from the use of such a time-honored statistic as the coefficient of variation. One consideration is that the maximum likelihood estimates are known to be efficient only when applied to their appropriate distributions, i.e., the arithmetic-geometric ratio is usable only for series well graduated by the Type III function. This would only mean that a relatively large amount of pertinent information should be extracted from the parent population, which may not always be possible. Another consideration is the fact that in a sample taken from a skewed population, if the mean is overestimated or underestimated, the variance is also overestimated or underestimated, respectively, so that the ratio $\frac{s}{x}$ (estimate of C. V.) partially compensates for the poor estimators of the numerator and the denominator (Gutierrez, 1965). In this regard, the coefficient of variation will be considered as a reliable and practical measure of relative dispersion which can be used even in distributions other than the normal.

3. *The Use of the Coefficient of Variation in the Specification of Precision.* The coefficient of variation of the sample mean [CV $(\bar{x})$] can be used to specify the desired relative precision of the estimate. An increase in the sample size decreases the value of CV $(\bar{x})$ and therefore increases the precision of estimate $\bar{x}$.

Hansen et al (1953) stated that the estimate of a standard deviation or a mean is sufficiently reliable if it has a coefficient variation of no greater than 10 per cent. In some cases, the precision may be lowered to 15 per cent.

4. *Analysis of the Problem.* In some instances, it may be impossible to measure a certain population characteristic, say Y_j, because no direct and reliable method of measure of the characteristic is known. The use of ancillary variables to estimate the characteristic Y_j; $j = 1, 2, \ldots, p$, may then be resorted to. Suppose that p different population characteristics Y_j's; $j = 1, 2, \ldots, p$, are to be estimated where Y_j is to be estimated from the j^{th} population. Known to be associated with Y_j in the p populations are the ancillary variables $X_1, X_2, \ldots, X_r$.

Thus,

$$
\begin{aligned}
y_1 &= a_1 + b_{11} x_1 + \ldots + b_{1i} x_i + \ldots + b_{1r} x_r \\
&\vdots \\
y_j &= a_j + b_{j1} x_1 + \ldots + {}_{ji} x_i + \ldots + b_{jr} x_r \qquad (4.1) \\
&\vdots \\
y_p &= a_p + b_{p1} x_1 + \ldots + b_{pi} x_i + \ldots + b_{pr} x_r
\end{aligned}
$$

And the estimated sample means are

$$
\begin{aligned}
\bar{y}_1 &= a_1 + b_{11} \bar{x}_1 + \ldots + b_{1i} \bar{x}_i + \ldots + b_{1r} \bar{x}_r \\
&\vdots \\
\bar{y}_j &= a_j + b_{j1} \bar{x}_1 + \ldots + b_{ji} \bar{x}_i + \ldots + b_{jr} \bar{x}_r \qquad (4.2) \\
&\vdots \\
\bar{y}_p &= a_p + b_{p1} \bar{x}_1 + \ldots + b_{pi} \bar{x}_i + \ldots + b_{pr} x_r
\end{aligned}
$$

where a_j and b_{j1} are the regression coefficients.

Suppose y_j (for any $j = 1, 2, \ldots, p$) is made a function of just one of the ancillary variables x_i; $i = 1, 2, \ldots, r$.

That is,

$$Y_j = a_{ji} + b_{ji} x_i \qquad i = 1, 2, \ldots, r$$
$$j = 1, 2, \ldots, p$$

This would mean, therefore, that for each subpopulations, there are r such equations. This procedure of setting up r equations for each subpopulation paves the way for the use of the programming technique. This procedure also tackles the problem of the covariance between X_i's because this time, only the sampling error of one variable is considered. Using the equations 4.1, however, the sampling variance of y_j is

$$\text{Var}(y_j) = \sum_{i=1}^{r} b_{ji}^2 \text{ Var}(x_i) + 2 \sum_{i=k}^{r} b_{ji} b_{jk} \text{ Cov } x_i x_k$$

$$j = 1, 2, \ldots, p.$$

4.1. *Predicted Value of the Rel-Variance.* The precision or variability of the actual sample mean $\bar{y}_j$ for a given sample size n_j is not necessarily the same as the precision or variability of the predicted sample mean y_j for the sample size n_j. It can be shown that if y_j is measured with a certain degree of precision, then y is calculated with a precision higher than the former. In this section is shown the relationship between the estimated coefficient of variation of the actual mean $[\hat{CV}(\bar{y})]$ and the coefficient of variation of the predicted sample mean $(CV(\hat{y}_i))$.

The total sum of squares is

SST = SSE + SSR

The ratio of SSR to SST as

$$\frac{SSR}{SST} = \frac{\Sigma (\hat{y}-\bar{y})^2/n \, (y)^2}{\Sigma (y-\bar{y})^2/n \, (\bar{y})^2}$$

$$= \frac{s^2 (\hat{y}) \,/\, (\bar{y})^2}{s^2 (\bar{y}) \,/\, (\bar{y})^2}$$

$$= \frac{[CV_{\hat{y}}]^2}{[CV_{y}]^2} = K \qquad (4.3)$$

Let $Rv\ (\hat{y}) \equiv$ rel-variance of $\hat{y}$.

$Rv\ (\hat{y}) \equiv$ the predicted value of the rel-variance of y.

Define

$$Rv\ (\hat{y}) = \frac{Rv\ (\hat{y})}{K} \quad (4.4)$$

With simple random sampling, the predicted value of the rel-variance of $\bar{y}$ would be

$$\hat{Rv}\ (\bar{y}) = \frac{N\text{-}n}{(N\text{-}1)n} \quad \frac{Rv\ (\hat{y})}{K} \quad (4.5)$$

4.2. *The Allocation for a Minimum Sum of Rel-Variances.* Suppose a researcher has prescribed for himself a certain precision which he would like to attain, say about $Rv_o\ (\bar{y}_j)$ for each of the $\bar{y}$'s in the j^{th} population. Consider the rel-variances of the means $\bar{y}_1, \bar{y}_2, \ldots, \bar{y}_p$: $Rv\ (\bar{y}_j)$, $j = 1, 2, \ldots, p$, i.e.,

$$Rv\ (\bar{y}_j) = \frac{Rv\ (\hat{\bar{y}}_j)}{K_j}.$$

$$\begin{aligned}
\sum^{p} Rv_o\ (\bar{y}_j) &= Rv_1\ (\bar{y}_1) + \ldots + Rv_1\ (\bar{y}_j) + \ldots + Rv_1\ (\bar{y}_p) \\
&\vdots \\
\sum^{p} Rv_o\ (\bar{y}_j) &= Rv_i\ (\bar{y}_1) + \ldots + Rv_i\ (\bar{y}_j) + \ldots + Rv_i\ (\bar{y}_p) \quad (4.6) \\
&\vdots \\
\sum^{p} Rv_o\ (\bar{y}_j) &= Rv_r\ (\bar{y}_1) + \ldots + Rv_r\ (\bar{y}_j) + \ldots + Rv_r\ (\bar{y}_p)
\end{aligned}$$

If N_j is the total number of population units of the j^{th} population, using (4.5) and simplifying, the following equations are obtained:

$$\sum_{j=1}^{p} Rv_o\ (\bar{y}_j) = \sum_{j=1}^{p} \frac{Rv_1\ (\hat{y}_j)}{n_j K_j} - \sum_{j=1}^{p} \frac{Rv_1\ (\hat{y}_j)}{(N_j\text{-}1)K_j}$$

$$\vdots$$

$$\sum_{j=1}^{p} Rv_o(y_j) = \sum_{j=1}^{p} \frac{Rv_i(\hat{y}_j)}{n_j K_j} - \sum_{j=1}^{p} \frac{Rv_i(\hat{y}_j)}{(N_j-1)K_j}$$

$$\vdots$$

$$\sum_{j=1}^{p} Rv_o(y_j) = \sum_{j=1}^{p} \frac{Rv_r(\hat{y}_j)}{n_j K_j} - \sum_{j=1}^{p} \frac{Rv_r(\hat{y}_j)}{(N_j-1)K_j} \tag{4.7}$$

The problem is to determine the value of n_j that will yield a minimum sum of rel-variances. By the use of Lagrange multiplier λ, let

$$\phi = \sum_{j=1}^{p} \frac{Rv(\hat{y}_j)}{n_j K_j} - \sum_{j=1}^{p} \frac{Rv(\hat{y}_j)}{(N_j-1)K_j} + \lambda (n_1 + \ldots + n_j + \ldots + n_p - n)$$

Applying the same procedures used in Neyman's minimum variance allocation, it can be verified that

$$n_j = n \frac{\sqrt{Rv(\hat{y}_j)/K_j}}{\Sigma \sqrt{Rv(\hat{y}_j)/K_j}} \tag{4.8}$$

Using this relationship, the allocation corresponding to the i^{th} equation (see equations 4.6 and 4.7) can be obtained.

TABLE 1. ALLOCATION BY THE ANCILLARY VARIABLE X_i

Ancillary Variable (X_i)	n_1	...	n_j	...	n_p
X_1	$n_1^{(1)}$	...	$n_j^{(1)}$	...	$n_p^{(1)}$
.	.				.
.	.				.
X_i	$n_1^{(i)}$	...	$n_j^{(i)}$	.	$n_p^{(i)}$
.	.				.
.	.				.
X_r	(r)	...	(r)	.	(r)

By the method described above, there are as many different allocations as there are ancillary variables, X_i; $i = 1, 2, \ldots, r$. The problem is, which of these allocations should be used? Or possibly, what should be done to arrive at just one allocation? This problem is tackled in the next section.

4.3. *A Solution to the Problem*: *The Programming Technique.* If M_{o1} is written in place of

$$\sum_{j=1}^{p} Rv_o (\bar{y}_j)$$

to represent the specified sum of precision of y_j's, then,

$$\sum_{i=1}^{p} \frac{Rv_1 (\hat{y}_j)}{n_j K_j} - \sum_{j=1}^{p} \frac{Rv_1 (\hat{y}_j)}{(N_j-1) K_j} \leqq M_{o1} \tag{4.9}$$

$$\vdots$$

$$\sum_{i=1}^{p} \frac{Rv_i (\hat{y}_j)}{n_j K_j} = \sum_{j=1}^{p} \frac{Rv_i (\hat{y}_j)}{(N_j-1) K_j} \leqq M_{or}$$

$$\vdots$$

$$\sum_{i=1}^{p} \frac{Rv_r (\hat{y}_j)}{n_j K_j} - \sum_{j=1}^{p} \frac{Rv_r (\hat{y}_j)}{(N_j-1) K_j} \leqq M_{o1}$$

The inequality signs are used because the most economical allocation may supply sum of rel-variances smaller than the desired M_{o1}. It will be shown later on that minimizing $n = \Sigma n_j$ for a fixed sum of rel-variances (M_{o1}) will give an allocation that is exactly the same as the allocation obtained in section 4.2. And if n is minimized subject to the constraints in 4.9, a programming problem is now on hand.

4.3.1. *Allocation by the graphical method.* Dalenius (1957) developed a graphical method of allocation which is very useful in multiparametric stratified sampling. This method, a programming one, is usable only when there are two strata.

If there are more than two strata, the author suggested the use of the techniques devised for linear programming.

The same technique used in Dalenius' graphical method of allocation will be applied to the present problem. The method is therefore usable only when $p = 2$, i.e., there are only two populations. With $p = 2$, the following equations will be obtained:

$$\frac{Rv_1(\hat{y}_1)}{n_1K_1} + \frac{Rv_1(\hat{y}_2)}{n_2K_2} \leq M_{o1} + \sum_{i=1}^{2} \frac{Rv_1(\hat{y}_j)}{(N_j-1)K_j} \qquad (4.10)$$

$$\vdots$$

$$\frac{Rv_1(\hat{y}_1)}{n_1K_1} + \frac{Rv_r(\hat{y}_2)}{n_2K_2} \leq M_{or} + \sum_{i=1}^{2} \frac{Rv_1(\hat{y}_j)}{(N_j-1)K_j}$$

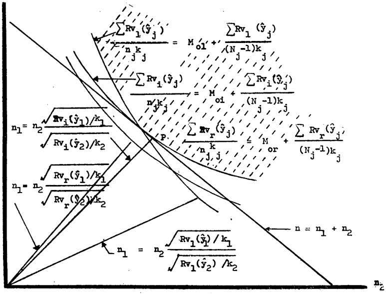
$$\vdots$$

$$\frac{Rv_1(\hat{y}_2)}{n_1K_1} + \frac{Rv_r(\hat{y}_1)}{n_2K_2} \leq M_{o1} + \sum_{i=1}^{2} \frac{Rv_r(\hat{y}_j)}{(N_j-1)K_j}$$

In figure 1, any point that lies within the shaded area satisfies the inequalities in (4.10). The point that will give a *minimum sum of rel-variances allocation,* however, can be located by the intersection of the line

$$n_1 = n_2 \frac{\sqrt{Rv(\hat{y}_1)/K_1}}{\sqrt{Rv(\hat{y}_2)/K_2}}$$

and the edge of the shaded area. It can be shown that this point of intersection is also the same point of tangency of the line $n = n_1 + n_2$ and the curve in question. The proof for this follows:



n_1
n_2
$\frac{\sum Rv_1(\hat{y}_j)}{n_j k_j} = M_{o1} + \frac{\sum Rv_1(\hat{y}_j)}{(N_j - 1)k_j}$
$\frac{\sum Rv_i(\hat{y}_j)}{n_j k_j} = M_{oi} + \frac{\sum Rv_i(\hat{y}_j)}{(N_j - 1)k_j}$
$\frac{\sum Rv_r(\hat{y}_j)}{n_j k_j} \leq M_{or} + \frac{\sum Rv_r(\hat{y}_j)}{(N_j - 1)k_j}$
P
$n_1 = n_2 \frac{\sqrt{Rv_i(\hat{y}_1)/k_1}}{\sqrt{Rv_i(\hat{y}_2)/k_2}}$
$n_1 = n_2 \frac{\sqrt{Rv_r(\hat{y}_1)/k_1}}{\sqrt{Rv_r(\hat{y}_2)/k_2}}$
$n_1 = n_2 \frac{\sqrt{Rv_1(\hat{y}_1)/k_1}}{\sqrt{Rv_1(\hat{y}_2)/k_2}}$
$n = n_1 + n_2$

FIGURE 1

Using any inequality in (4.10),

$$-\frac{Rv\ (\hat{y}_1)}{n^2{}_1 K_1}\ dn_1 \quad -\frac{Rv\ (\hat{y}_2)}{n^2{}_2 K_2}\ dn_2 = 0$$

or,

$$\frac{dn_1}{dn_2} = -\ \frac{Kn_1\ n^2{}_1\ Rv\ (\hat{y}_2)}{Kn_2\ n^2{}_2\ Rv\ (\hat{y}_1)} \qquad (4.11)$$

and from the equation $n = n_1 + n_2$

$$dn_1 + dn_2 = 0$$

or,

$$\frac{dn_1}{dn_2} = -1 \qquad (4.12)$$

Equating (4.1) and (4.2),

$$n_1 = n_2 \frac{\sqrt{Rv\ (\hat{y}_1)/K_1}}{\sqrt{Rv\ (\hat{y}_2)/K_2}} \qquad (4.13)$$

The problem is therefore simplified by just finding the point of tangency of the line $n = n_1 + n_2$ and the edge of the shaded area.

4.3. *Allocation by the linear programming approach.* The method previously discussed is usable when $p = 2$. When $p \geq 2$, linear programming can be used. This method is concerned with maximizing or minimizing a linear expression called the "objective function" subject to linear constraints. The latter are either inequalities or equations. In the present case, the objective function is $U = \Sigma\ u_j$ where $u_j = \frac{1}{n_j}$. The linear programming problem is then that of maximizing $U = \Sigma u_j$ subject to the following constraints:

$$\sum_{j=1}^{p} \frac{Rv_i\ (\hat{y}_j)}{K_j}\ u_j \leq M_{oi} + \Sigma\ \frac{Rv_i\ (\hat{y}_j)}{(N_j\text{-}1)\ K_j}$$

$$\sum_{j=1}^{p} \frac{Rv_1\ (\hat{y}_j)}{K_j}\ u_j \leqq M_{o1} + \Sigma\ \frac{Rv_1\ (\hat{y}_j)}{(N_j\text{-}1)\ K_j}$$

$$\vdots$$

$$\sum_{j=1}^{p} \frac{Rv_r\ (\hat{y}_j)}{K_j}\ u_j \leqq M_{or} + \Sigma\ \frac{Rv_r\ (\hat{y}_j)}{(N_j\text{-}1)\ K_j}$$

To this set of constraints are added the following:

$$u_1 \geqq \frac{1}{n_{o1}}$$

$$\vdots$$

$$u_j \geqq \frac{1}{n_{oj}}$$

$$\vdots$$

$$p_p \geqq \frac{1}{n_{op}}$$

Where n_{oj} is the largest value of n_j among all $n_{j's}$ in the j^{th} population (argest $n_j^{(i)}$) and is obtained by the *minimum sum of rel-variance* allocation (see Table 1). This problem is illustrated in Figure 2. For simplicity, consider only $p = 2$ and $r = 2$.

In Figure 2, the line $U = u_1 + u_2$ is moved as far as possible from the origin until it has exactly one point in common with the feasible region (shaded area). With linear programming, this point is always a corner point. It must be pointed out that this method is usable only when $p = 2$. For $p > 2$ various techniques devised for linear programming should prove helpful. One such technique, the Simplex Method,[1] with some modifications, converts the constraints (inequalities) into equations. The problem is now written as: Maximize $U = \Sigma\ u_j$ subject to the following constraints:

[1] For the purpose of this study the Simplex Method in the form of extended tableaux and modified into the m-method, will be used.

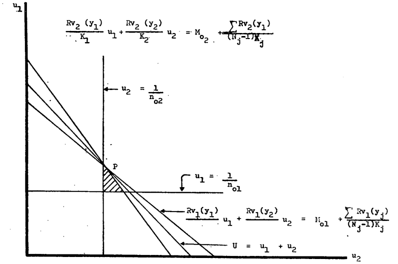

u1
Rv2 (y1)/K1 u1 + Rv2 (y2)/K2 u2 = Mo2 + ∑Rv2(y1)/(Nj-1)Kj
u2 = 1/no2
P
u1 = 1/no1
Rv1(y1)/K1 u1 + Rv1(y2)/K2 u2 = Mo1 + ∑ Rv1(yj)/(Nj-1)Kj
U = u1 + u2
u2

FIGURE 2

$$\frac{Rv_1\ (\hat{y}_1)}{K_1} u_1 + \ldots + \frac{Rv_1\ (\hat{y}_j)}{K_j} u_j + \ldots + \frac{Rv_1\ (\hat{y}_p)}{K_p} u_p + u_{p+1}$$

$$= M_{1o} + \frac{\Sigma\ Rv_1\ (\hat{y}_j)}{(N_{1-j})\ K_j}$$

$$\vdots$$

$$\frac{Rv_i\ (y_1)}{K_1} u_1 + \ldots + \frac{Rv_i\ (y_j)}{K_j} u_j + \ldots + \frac{Rv_i\ (y_j)\ u_p}{K_p} + u_{p+i}$$

$$= M_{oi} + \frac{\Sigma\ Rv_i\ (\hat{y}_j)}{(N_{j-1})\ K_j}$$

$$\vdots$$

$$\frac{Rv_r\ (\hat{y}_1)}{K_1} u_1 + \ldots + \frac{Rv_r\ (\hat{y}_j)}{K_j} u_j + \ldots + \frac{Rv_r\ (\hat{y}_p)}{K_p} u_p + u_{p+r}$$

$$= M_{or} + \frac{\Sigma\ Rv_r\ (\hat{y}_j)}{(N_j\text{-}1)\ K_j}$$

$$u_1 - u_{p+r+1} = \frac{1}{n_{o1}}$$

$$\vdots$$

$$u_j - u_{p+r+j} = \frac{1}{n_{oj}}$$

$$\vdots$$

$$u_p - u_{2p+r} = \frac{1}{n_{op}}$$

$u_1, u_2, \ldots, u_{p+1}, \ldots, u_{p+r}, u_{p+r+1}, \ldots, u_{2p+r} \geqq 0$ where $u_{p+1}, \ldots, u_{2p+r}$ are the artificial variables that cause the inequalities to be converted into equalities. After the optimum

u_j is known, the value of n_j is obtained by the relationship $n_j = \frac{1}{u_j}$.

5. *An Illustrative Example.* The crop year period for palay is from July 1 to June 30 of the next year. To be able to estimate mean palay production per farming household, a sample can be taken during or after the harvest season late in the reference period. A forecast, however, can be made on palay production (Y_j) using the population characteristics: area planted with palay (X_2), and total farm area (X_1). The data on area planted with palay and the total farm area can actually be collected at the beginning of the crop year period.

In this example, the object of study is the measurement of palay production in irrigated areas (Y_1) and palay production in non-irrigated areas (Y_2) for the crop year 1967-1968. Since Y_1 and Y_2 can only be predicted, sampling must therefore be made from the population units of X_1 and X_2. The two populations in this study are: π_1: the farming households working in irrigated farms, and π_2: the farming households working in non-irrigated farms.

The problem, therefore, is to determine the sample sizes n_1 and n_2 to be allocated to π_1 and π_2.

5.1. *The data.* The data used in this example were taken from the results of the 1966-1967 listing operation of the Crop and Livestock Survey conducted by the Bureau of Agricultural Economics. In that operation, all households (farming and non-farming) were listed in each sample barrio. Other information such as the total farm area, area planted with palay, production of palay for this year's crop and last year's crop, quantity of fertilizer used, etc., were also listed.

The populations π_1 and π_2 are confined only to four sample barrios in Nueva Ecija. The sampling unit is the *farming household* which is defined as a household which derives income from operating a *farm.* A *farm* includes all parcels, lots, or pieces of land operated by one person, whether by his own labor or with the assistance of the members of his household or hired employees. The total number of farming households working in irrigated farms (π_1) is 545 and the total number of farming households working in non-irrigated farms (π_2) is 484.

5.2. *Estimates of the Parameters.* Using data obtained from the 1966-1967 listing operation and the model $Y = \alpha + \beta_1 x_1 + \beta_2 x_2 + \epsilon$, estimates of Y_1 and Y_2 are obtained as:

$$\hat{y}_1 = 33.422 - 1.49 x_1 + 30.19 x_2 \tag{5.1}$$

$$\hat{y}_2 = 37.798 - 11.07 x_1 + 42.44 x_2$$

where $\hat{y}_1$ is the predicted palay production (in cavans) per farming household working in irrigated farms and $\hat{y}_2$ is the predicted palay production per farming household working in irrigated farms and y_2 is the predicted palay production per farming household working in non-irrigated farms. The unit of measurement of X_1 and X_2 is the hectare. The predicted mean production per farming household can now be written as:

$$\hat{y}_1 = 33.422 - 1.49 x_1 + 30.19 \hat{x}_2 \tag{5.2}$$

$$\hat{y}_2 = 37.798 - 11.07 \hat{x}_1 + 42.44 \hat{x}_2$$

From the same data, the coefficients of variation of the predicted palay production per farming household are computed. The computed values are shown in Table 2. Also listed in this table are the regression equations for y_j's and their corresponding K-values ($K = \frac{SSR}{SST}$). It will be noted that $\hat{y}_j$ is now expressed as a function of only one ancillary variable. The last column of the table gives the predicted values of the rel-variances.

TABLE 2. PREDICTED VALUES OF THE REL-VARIANCES

Equation Form	$K = \frac{SSR}{SST}$	$[\hat{CVy}\%]^2$	Predicted $[CVy\%]^2$
$\hat{y}_1 = 71.80 + 16.259x_1$	0.24596	$(21.21)^2$	1828.9
$\hat{y}_1 = 32.07 + 28.90\ x_2$	0.52993	$(36.59)^2$	2526,3
$\hat{y}_2 = 64.098 + 20.06\ x_1$	0.15600	$(24.73)^2$	3920.3
$\hat{y}_2 = 29.827 + 33.16\ x_2$	0.37320	$(35.32)^2$	3342.7

5.3. *Solution to the Problem.* Suppose it is desired to measure the sample mean with a precision of about 10% [(CVy%) = 100%]. The sum of precision of the sample means can therefore be specified as 200%. The problem then can be written as: Minimize $n = n_1 + n_2$ subject to the following constraints:

$$\frac{545 - n_1}{544\ n_1} 1828.9 + \frac{484 - n_2}{483\ n_2} 39920.3 \leq 200$$

$$\frac{545 - n_1}{544\ n_1} 2526.3 + \frac{484 - n_2}{483\ n_2} 3342.7 \leq 200$$

or simply,

$$\frac{1828.9}{n_1} + \frac{3920.3}{n_2} \leq 211.42$$

$$\frac{n_1}{2526.3} + \frac{3342.7}{n_2} \leq 211.55$$

The graphical solution to this problem is shown in Figure 3.

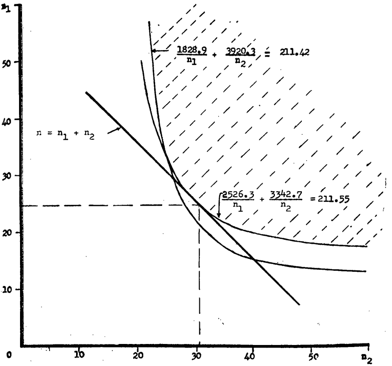


FIGURE 3

$n_1 = 25$ n_1 = No. of farming households working in irrigated areas.

$n_2 = 31$ n_2 = No. of farming households working in non-irrigated areas.

When the near programming approach is used, the *minimum sum of rel-variances* allocation is first resorted to, to determine the restrictions on the sample size n_j. The two sets of allocation corresponding to ancillary variables X_1 and X_2 are tabulated in Table 3.

TABLE 3. ALLOCATIONS FOR MINIMUM SUM OF REL-VARIANCE

Ancillary Variable (X_i)	$n = \dfrac{[\Sigma(CVy_j)]^2}{\Sigma Rv_o(y_j) + \dfrac{\Sigma Rv(y_j)}{N_j - 1}}$	n_1	n_2
Total Farm Area (X_1)	52.4324	21.3748	31.2685
Area Planted with Palay (X_2)	55.2355	25.6771	29.5481

From these tabulated values, the following restrictions are obtained:

$$n_1 \leq 25.6771$$

$$n_2 \leq 31.2685$$

Let $u_1 = \dfrac{1}{n_1}$ and $u_2 = \dfrac{1}{n_2}$. If the specified sum of precision is 200%, the problem can be written as: Maximize $U = u_1 + u_2$ subject to the following constraints:

$$1828.90\ u_1 + 3920.30\ u_2 \leq 211.42$$

$$2526.30\ u_1 + 3342.70\ u_2 \leq 211.55$$

$$u_1 \geq \frac{1}{25.6771}$$

$$u_2 \geq \frac{1}{31.2685}$$

Using the Simplex Method of linear programming, the constraints can be rewritten as:

$$1828.90\ u_1 + 3920.30\ u_2 + u_3 = 211.42$$

$$2526.30\ u_1 + 3342.70\ u_2 + u_4 = 211.55$$

$$u_1 - u_5 = 0.038945$$

$$u_2 - u_6 = 0.031981$$

$$u_1,\ u_2,\ u_3,\ u_4,\ u_5,\ u_6 \geq 0$$

For this particular type of problem, the m-method of the Simplex Method[1] is used. The successive tableau of the method yield the following solution:

$$u_1 = 0.041427$$

$$u_2 = 0.031981$$

To get $n_1 = n_2$,

$$n_1 = \frac{1}{u_1} = 24.97 \rightarrow 25$$

$$n_2 = \frac{1}{u_2} = 31.27 \rightarrow 32$$

6. *Conclusion.* This study attempts to find a method of sample size allocation among several subpopulations that is useful when the characteristic under consideration can not be directly measured but can only be predicted or estimated by ancillary variables. Some advantages in using this method are the following:

(1) There is no need to compute for the covariance between ancillary variables when estimating the coefficients of variation

[1] Reference, Heady and Candler (1958), "Method to Use in Meeting Minimum Restriction", pp. 202-204.

(2) The method allows partial relaxation of the precision of some estimates if the variability of these estimates are very high.

(3) The method gives a compromise allocation where no obvious compromise is possible.

BIBLIOGRAPHY

Anderson, R. L. and Bancroft, T. A. "Statistical Theory in Research." McGraw-Hill Book Company, New York, Toronto and London. 1952

Candler, W. and Heady, E. O. "Linear Programming Methods." The Iowa State College Press, Ames, Iowa, U.S.A. 1958

Castillo J. V. and Damasco, M.D. "The Present Situation and Outlook of the Estimation of Rice Supply of the Country" *Papers and Reviews on Rice and Related Statistics.* 1965

Cochran, W. G. "Sampling Techniques." John Wiley and Sons Inc., New York and London, 1953 Dalemus, T. "Sampling in Sweden: Contribution to the Methods and Theories of Sample Practice" Almqvist and Wiksell, Stockholm, Sweden, 1957.

Gutierrez, J. S. "Some Theoretical Considerations on the Uses of Coefficients of Variation in Sample Survey." *The Philippine Statistician,* Volume XIV November 3. 1965

———. "On the Comparison of the Ratio of Squared Coefficients of Variation and Ration of Variances as Indicators of Relative Efficiency." Master's Thesis, Department of Experimental Statistics, North Carolina State College, Raleigh. 1958

Hansen, M. H., Hurwitz, W. N., Madow, W. G. "Sample Survey Methods and Theory." 2 Vols. John Wiley and Sons, Inc., New York. Chapman and Hall, Ltd., London. 1953

Keeping, E.S. and Kenney, J. F. "Mathematics of Statistics." D. Van Nostrand Company, Inc., Princeton, New Jersey, Toronto, London, London, and New York. 1953.

Norris, N. "Some Efficient Measures of Relative Dispersion." Annals of Mathematical Statistics. Volume9, p. 214. 1938

Vajda, S. "Mathematical Programming." Addison-Wesley Publishing Co., London, England. 1761

The Philippine Statistician
Vol. XVIII, Nos. 1 and 2, 1969

MODELS FOR ESTIMATING INTERNAL MIGRATION

ELSA L. GUTIERREZ* AND JOSE S. GUTIERREZ**

The demographic variables, fertility and mortality have much more developed methodology than has migration. Collection and analysis of data on migration are relatively not well developed. In particular, methods of analysis of migration data are among the weakest of demographic tools, hence the relative difficulty in analysis.

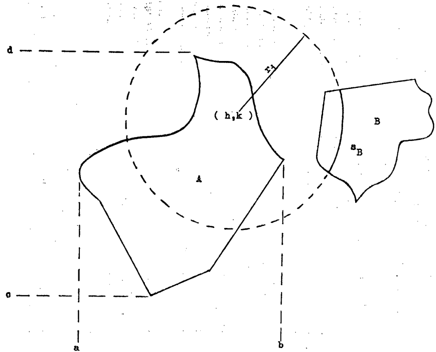
Different mathematical models of migration have been proposed by Ravenstein, Young, Stuffer, Stewart, Zipf, Dodd, Kulldorf, Bogue and many other social scientists. Recently, Thomlison (1961) presented a mathematical model which will enable demographers to compare migration rates by controlling spatial variables such as size of area of origin, shape of area of destination, destribution of population within area of origin, distribution of population within area of destination and distance moved.

The Thomlinson model can be stated as follows: Given a circle of radius r_i with its center in area A, if part of the perimeter of the circle falls in a second area B, then the probability that a person starting at the center of the circle and moving a distance r will be an interarea migrant is equal to

$$\frac{\text{Length of arc in area B}}{\text{Circumference of circle}} \qquad (1)$$

For example, consider a source area A of irregular shape and density D, a terminal area B of irregular shape, a distance of migration r_i, a point of origin (h, k) and an arc A_B of circle intersecting area B as shown in Fig. 1.

* Project Leader, Socio-Economic Survey of Quezon City, Office on Social Justice, Office of the Mayor, Quezon City.
** Assoc. Professor, Statistical Center, University of the Philippines.


d
(h,k)
B
A
c
a
b

A conditional probability statement can be stated as follows:

Pr. (migrant settles in B / move of r_i miles from point (h, k) =

$$\frac{A_B D_{h,k}}{2\Pi r} \quad (2)$$

The probability that any migrant traveling r_i miles from a point in A settles in B is given by the formula

$$\Pr(A,B;r_i) = \frac{\int_c^d \int_a^b A_{B,xy} \, D_{xy} \, dxdy}{2\Pi r \int_c^d \int_a^b D_{xy} \, dxdy} \quad (3)$$

The total number of migrants $M(A,B;r_i)$ is estimated as follows:

$$M(A,B;r_i) = [\Pr(A,B;r_i)] \, [M(A;r_i)] \quad (4)$$

where $M(A;r_i)$ is the number of migrants going r_i miles from A.

Models similar to that of Thomlinson's approach are formulated in this paper. Four models designed to compute the number of migrants attributable to spatial variables were used. The constructed models depend on two factors; namely, a probability factor and a computed potential number of migrants moving within a radius, r_i from a point of origin.

Estimation of proportions of migrants by barrio.

The probability factor in the total migrants is illustrated as follows:

$$P_T(a,A;r) = \frac{\text{Population of barrio, a}}{\text{Sum of the population of all barrio (A) within a radius (r) including that of a}} \quad (5)$$

$$P_T(a,A;r) = \frac{Pa}{\Sigma P\acute{a}}$$

It should be noted that the probability factor $P_T(a,A;r)$ will also yield the proportion of migration of both men and women coming from barrio (a) and residing in other barrios in A within radius, r.

The proportion $P_T(a,A;r)$ is the expected of total migrants coming from barrio, a. Comparison between the expected and observed proportions of migrants coming from different barrios of Bulacan, Bulacan is given Table 1. This square test yield a close agreement between the expected and observed proportions of out-migrants (χ^2 computed = 0.7978, χ^2_{12}, 05 = 21.026).

TABLE 1. POPULATION 1960, NUMBER OF OUT MIGRANTS, AND EXPECTED AND OBSERVED PROPORTION OF OUT-MIGRANTS BY BARRIO IN BULACAN, BULACAN, 1967

Barrio	Total Pop'n 1960	No. of Out Migrants	Proportion Expected	Out-Migrants Observed
Bagumbayan	807	6	4.39	3.10
Balubad	660	6	3.59	3.10
Bambang	3,989	34	21.68	31.00
Matungao	2,054	19	11.17	11.73
Maysantol	687	5	3.73	3.09
Pitpitan	702	7	3.82	4.32
Poblacion	904	8	4.91	4.94
Tabang	676	8	3.67	4.94
San Nicolas	1,126	10	6.12	6.17
Sta. Ana	2,567	22	13.96	13.58
Sta. Ines	1,508	13	8.20	8.02
Taliptip	2,395	21	13.02	12.96
Tibig	320	3	1.74	1.85
Total	18,395	162	100.00	100.00

$\chi^2 = 0.7978$

Estimation of proportions of male and female out-migrants by barrio.

The probability factor $P_T(a,A;r)$ which gives the total proportion of out-migrants coming from barrio (a) may be split into two probability factors $P_M(a,A;r)$ or the proportion of male out-migrants from barrio (a) and $P_F(a,A;r)$ of the proportions of female out-migrants from barrio (a) and residing in barrios A within a radius, r.

Sympoblically then

$$P_T(a,A;r) = P_M(aA;r) + (P_F(a,A;r)$$

where

$$P_M(a,A;r) = \frac{\text{Male population from barrio (a)}}{\Sigma\ P\acute{a}}$$

and

$$P_F(a,A;r) = \frac{\text{Female population form barrio a}}{\Sigma\ P\acute{a}}$$

Non-significant results were obtained when the χ^2 test was applied to the expected and observed proportions of male and female out-migrants (χ^2 values obtained; male: 1.2612, female; 3.3785).

TABLE 2. MALE POPULATION 1960, NUMBER OF MALE OUT-MIGRANTS AND EXPECTED AND OBSERVED PROPORTION OF MALE OUT-MIGRANTS BY BARRIO IN BULACAN, BULACAN, 1967

Barrio	Male Pop'n 1960	No. of Out-migrants	Proportion of Out-migrants Expected	Observed
Bagumbayan	419	4	2.28	2.47
Balubad	338	3	1.84	1.85
Bambang	1.929	20	10.49	12.34
Matungao	1,032	10	5.61	6.17
Maysantol	352	3	1.91	1.85
Pitpitan	348	3	1.89	1.85
Poblacion	452	5	2.46	3.09
Tabang	342	3	1.86	1.85
San Nicolas	541	5	2.94	3.09
Sta. Ana	1,281	13	6.96	8.02
Sta. Ines	755	8	4.10	4.94
Taliptip	1,208	10	6.57	6.17
Tibig	189	1	1.03	.62
Total	9,186	88	49.94	54.31

TABLE 3. FEMALE POPULATION 1960, NUMBER OF FEMALE OUT-MIGRANTS AND EXPECTED AND OBSERVED PROPORTION OF FEMALE OUT-MIGRANTS BY BARRIO IN BULACAN, BULACAN, 1967

	Female Pop'n 1960	No. of out-migrants	Proportion of Out-migrants Expected	Observed
Bagumbayan	388	2	2.11	1.23
Balubad	322	3	1.75	1.85
Bambang	1,060	14	11.19	8.64
Matungao	1,022	9	5.56	5.55
Maysantol	335	2	1.82	1,23
Pitpitan	354	4	1.93	2.47
Poblacion	452	3	2.45	1.85

	Female Pop'n 1960	migrants No. of out-	Expected Proportion of	Observed Out-migrants
Tabang	334	5	1.81	3.09
San Nicolas	585	5	3.18	3.09
Sta. Ana	1,286	9	6.99	5.55
Sta. Ines	753	5	4.10	3.09
Taliptip	1,187	11	6.45	6.79
Tibig	131	2	0.71	1.23
Total	9,208	74	50.06	

Estimation of expected number of migrants

The four estimators considered in estimating the theoretical (expected) number of migrants within a radius, r, from a point or area, a, are given below. The estimators under consideration in this study consist essentially of two factors, a probability factor and potential number of migrants. The theoretical number of migrants M(a,A;r) is estimated as follows:

$$M(a,A;r) = P(a,A;r)\ M(a,r)$$

where P(a,A;r) is the probability factor and M(a,r) is the potential number of migrants within a radius, r, from a point (area) of origin, a.

The first estimator $M_1(a,A;r)$ is computed as follows:

$$M_1(a,A;r) = P(a,A;r)\ [100^2\ P(a,A;r) - 100]$$

where M(a,r) is estimated by $[100^2\ P(a,A;r) - 100]$
The potential number of migrants is a certain proportion of a thousand population of which 100 may stay or will not migrate completely.

The second estimator $M_2(a,A;r)$ is given as follows:

$$M_2(a,A;r) = P(a,A;r)\ [\frac{100^2}{\Pi r} - P(a,A;r)]$$

$$= P(a,A;r)\ [\frac{100^2}{\Pi} - P(a,A;r)]$$

where M(a,r) is given by $[\frac{100^2}{\Pi}\ P(a, A;r)]$ and where r is taken as unity. The assumption of a unit radius seems reasonable for estimating internal migration, since the area in this country practically contigous and the means of transportation used is practically the same throughout the place.

The third estimator is a simple one and is computed as follows:

$$M_3(a, A;r) \; [1] \frac{Pa}{100}$$

where the probability factor P(a,A;r) is unity and the potential migrant is estimated to be 1/100*th* of the total population of the barrio. A proportional factor other than 1/100 can be used and can be established on the basis of actual observations.

The fourth estimator, $M_4(a,A;r)$, utilizes a different probability factor, that is, the probability that the potential migrant will move is proportional to the total population of the town minus the population of the barrio. The fourth estimator is computed as follows:

$$M_4(a,A;r) = [1\text{-}P(a,A;r)] \frac{Pa}{100}$$

where M(a,r) is the same as in third estimtor.

The logical explanation of the use of the probability factor [1-P(a,A;r)] in the fourth estimator is that population pressure exerts a strong influence on migration. If the population pressure on area a is stronger than outside a, then the tendency for people to move out of a is expected to be stronger

The models for estimating migration formulated in this study did not, however, take into consideration socio-economic factors, educational factors and other demographic factors.

The theoretical estimates using different estimators and the actual number of out-migrants are given in table 4. The last two estimators yielded non-significant differences when compared with the actual number of out-migrants. The first two, however, yielded significantly different estimates of out-migrant than what was actually observed.

Applications of estimators M_3 and M_4 on the estimation of male and female migrants are given in table 5. Non-significant results were obtained in both sexes and using both M_3 and M_4.

SUMMARY AND CONCLUSIONS

An attempt was made to formulate four mathematical models for estimating the number of migrants. The models studied consisted of a probability factor and an estimated optential number of migrants. Goodness of fit test between the estimated total number and observed total number of migrants yielded two significant results and two non-significant results.

The two formulas which yielded non-significant results in the estimation of total out-migrants were applied in the male and female out-migrants. Non-significant differences were obtained using the chi-square goodness of fit.

TABLE 4. TOTAL NUMBER OF OUT-MIGRANTS OBSERVED AND ESTIMATED USING VARIOUS ESTIMATORS.

Barrio	No. of Out-Migrants Observed	M_1	M_2	M_3	M_4
Bagumbayan	6	14.88	6.14	8.07	7.71
Balubad	6	9.29	4.10	6.60	6.36
Bambang	34	448.34	149.93	39.89	31.24
Matungao	19	113.59	39.79	20.54	18.24
Maysantol	5	10.18	4.43	6.87	6.61
Pitpitan	7	10.77	4.65	7.02	6.75
Poblacion	8	19.19	7.68	9.04	8.59
Tabang	8	9.79	4.29	6.76	6.51
San Nicolas	10	31.33	11.94	11.26	10.57
Sta. Ana	22	180.92	62.16	25.67	22.08
Sta. Ines	13	59.04	21.44	15.08	13.84
Taliptip	21	156.55	54.07	23.95	20.83
Tibig	3	1.28	0.96	3.20	3.14
TOTAL	162	1065.15	371.59	183.35	162.47
X^2 computed (X^2_{12} 21.02)		717.55	162.60	3.69	1.48

TABLE 5. NUMBER OF MALE AND FEMALE OUT-MIGRANTS, OBSERVED AND ESTIMATED USING TWO ESTIMATORS

Barrio	No. of Male Out Migrants			No. of Female out-migrants		
	Observed	M_3	M_4	Observed	M_3	M_4
Bagumbayan	4	4.19	4.09	2	3.88	3.79
Balubad	3	3.38	3.30	3	3.22	3.16
Bambang	20	19.29	17.26	14	3.88	3.79
Matungao	10	10.32	9.74	9	20.22	9.65
Maysantol	3	3.52	3.45	2	3.35	3.28
Piitpitan	3	3.48	3.41	4	3.54	3.47
Poblacion	5	4.52	4.40	3	4.52	4.40
Tabang	3	3.42	3.35	5	3,34	3.27
San Nicolas	5	5.41	5.25	5	5.85	5.66
Sta. Ana	13	12.81	11.91	9	12.86	11.96
Sta. Ines	8	7.55	7.24	5	7.53	7.22
Taliptip	10	12.08	11.28	11	11.87	11.10
Tibig	1	1.89	1.87	2	1.31	1.30
TOTAL	88	91.86	86.56	74	92.09	86.55
X^2 computed		1.19	2.23		6.78	5.69

LITERATURE CITED

Thomlinson, Ralph. A model for Migration Analysis. *Journal of the American Statistical Association* Volume 56 November 295 (1961)

1969 PSA INSTITUTIONAL MEMBERS

Central Bank of the Philippines

China Banking Corporation

Development Bank of the Philippines

Filipinas Cement Corporation

Insular Life Assurance Co,. Ltd.

Government Service Insurance System

Manila Electric Company

Menzi and Company

National Life Insurance Company of the Philippines

National Investment and Development Corporation

Philippine American Life Insurance Co.

Philippine National Bank

Philippine Refining Company

Procter and Gamble PMC

Provident Insurance Company of the Philippines

Social Security System

Sycip, Gorres, Velayo and Co.

Victorias Milling Company

OBITUARY

The Philippine Statistical Association regrets to announce the untimely demise of the following members:

MR. HONESTO R. BRINGAS

DR. FEDERICO M. SIOSON

Republic of the Philippine
Department of Public Works and Communcations

BUREAU OF POSTS
Manila

SWORN STATEMENT
(Required by Act 2580)

The undersigned, CESAR M. LORENZO, editor of THE PHILIPPINE STATISTICIAN published quarterly in English at Rizal Hall, Padre Faura, Manila, after having been duly sworn in accordance with law, hereby submits the following statement of ownership, management, circulation, etc, which is required by Act 2580, as amended by Commonwealth Act No. 201.

	Name	*Address*
Editor	CESAR M. LORENZO	Liberty Mines Incorporated
Business Manager	JOSE S. GUTIERREZ ..	U.P. Statistical Center
Owner	Philippine Statistical Association, Inc.	P. O. Box 3223, Manila
Publisher	— do —	— do —
Printer	U.S.T. Press	España, Manila
Office of Publication	— Rizal Hall, Padre Faura St. Manila	

If publication is owned by a corporation, stockholders owning one per cent or more of the total amount of stocks: — None —

Bondholders, mortgagees, or other security holders owning one per cent or more of total amount of security: — None —

In case of daily publication, average number of copies printed and circulated of each issue during the preceding month of ...., 19....

In case of publication other than daily, total number of copies printed and circulated of the last issue dated June 1968

1. Sent to paid subscribers 477
2. Sent to others than paid subscribers 48

TOTAL 525

CESAR M. LORENZO
Editor

SUBSCRIBED AND SWORN to before me this 25th day of April 1969, at Manila, the affiant exhibiting his Residence Certificate No. A-309128 issued at Manila on January 13, 1970.

Doc. No. 824
Page No. 66
Book No. VI
Series of 1969

(Officer Administering Oath)

NOTE: This form is exempt from the payment of documentary stamp tax.