The Philippine Statistician

JAN. - AUG., 1970 • VOL. XIX • 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 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. XIX, Nos. 1 and 2, 1970

1970 GRADUATION: U.P. Statistical Center

EXEQUIEL S. SEVILLA**

This morning, I will not talk to you about Statistics because I know you have heard enough on that subject. With your indulgence, I would like to speak about two types of LEADERSHIP. The first type is the one about which most of us are familiar with and may be called as the authoritative or conservative or paternalistic kind. The second type is less known and may be called the participative kind or leadership by participation.

It is generally admitted that people must be led. They also say that people perform best under leaders who are creative, imaginative, and aggressive—under leaders who lead. It is the responsibility of the leader to marshall the forces of the organization, to stimulate effort, to capture the imagination, to inspire people, to coordinate efforts, and to serve as a model of sustained effort.

The leader should keep an appropriate social distance, show no favorites, control his emotions, command respect and be objective and fair. He must know what he is doing and where he wants to go. He must set clear goals for himself and for the group or the corporation and then properly communicate these goals to all members of the organization. He must listen to advice and counsel before making decisions. However, it is *his responsibility* to make decisions and to set up mechanisms for seeing that the decisions are implemented. After weighing the facts and seeking expert counsel, he must make policy

* Speech delivered at the Commencement Exercises, Statistical Center and Population Institute, April 11, 1970.
** President, National Life Insurance Co.

and rules, set reasonable boundaries, and see that these are administered with justice and wisdom, even compassion.

The leader should reward good performance and learn effective ways of showing appreciation. He must be equally ready to criticize where warranted and to appraise performance frequently, fairly and without error. He must command strong discipline, not only because people respect a strong leader, but because strength and firmness indicate care and concern. Good leadership requires good followers. People tend to follow good leaders. Methods of election and selection are therefore very important. In many cases, finding the right chairman or president is the critical variable in the success of a program or a company. The quality of an organization is often judged by the visible quality of its leadership.

I HAVE JUST DESCRIBED TO YOU the first type of leadership: the AUTHORITATIVE kind which is familiar to most of us. I WILL NOW PROCEED TO DESCRIBE TO YOU the second type of leadership, the PARTICIPATIVE or GROUP APPROACH to leadership.

People grow, produce, and learn best when they set their own goals, choose activities that they see as related to these goals, and have a wide degree of freedom of choice in all parts of their lives. Under most conditions, persons are highly motivated, like to take responsibilities, can be trusted to put out a great deal of effort toward organizational goals, are creative and imaginative, and tend to want to cooperate with others.

The most effective leader is one who acts as a catalyst, a consultant, and a resource to the group. His job is to help the group to grow, to emerge and to become more free. He acts in such a way as to facilitate group strength, individual responsibility, diversity, non-conformity and aggressiveness. The good leader tends not to lead. He permits, feels, acts, relates, fights, talks — acts human as do other members of the group.

The participative leader who trust his colleagues and subordinates and has confidence in them tends to be open and frank, to be permissive in goal setting, and to be non-controlling in personal style and leadership policy. The self-adequate person tends to assume that others are also adequate and that, other things being equal, they will be responsible, loyal, appropriately work-oriented when work is to be done and adequate to carry out jobs that are commensurate with their levels of experience and growth. People naturally tend to share their feelings with those whom they trust and this is true at the simplest and most direct level of inter-personal relationships, as well as at more complex levels of organizational communication.

The participative leader is permissive in his relations with subordinates, for he assumes that people grow, learn to assess their own aptitudes, discover their deep-laying interests, and develop their basic potentials. Therefore, he gives his subordinates every opportunity to maximize self-determination and self-assessment, to verbalize their goals, to try new jobs or enlarge the scope of the work they are doing. He also trusts them to make mature judgments about job assignments. When he is dealing with a work-team or a group, he lets the group to make decisions about job allotments and work assignments. Basic to participative leadership is this process of allowing people to be responsible for their own destinies, for setting their own targets, assessing their own development needs, searching out resources to aid in job accomplishment and participating in setting up organizational objectives.

The participative administrator helps in creating a climate in which he has no need to impose controls. He knows that in a healthy group, controls emerge from group processes as the need is felt. Then controls are medicated by group or organizational objectives and by such relevant data as deadlines and target dates. People or groups who have set their own goals build internal tension-systems which maintain goal orientation and create appropriate boundaries.

Formal and written rules about such things as work space, canteen and washroom use, stockroom neatness are less and less necessary when people are engaged in a common task with others whose feelings and views they freely share. When there is trust and mutuality, people are inclined to respect the rights of fellow members. Moreover, the feeling that one is trusted encourages exploration, diversity and innovation, for the person spends little time trying to prove himself. His time and energy are freed to define and solve problems, accomplish work and create new dimensions for his job. On the other hand, a fearful person uses a great deal of energy in defending himself against actual or imagined threat or attack. Direct and open flow of information serves to create an atmosphere which encourages people to share information with those above as well with those below. In general, openness and information giving and expressing feelings enhances consensus and the more nearly a group can reach consensus on operational problems, the higher the quality of the decision and the greater the group's commitment to the program.

Which type of leadership is better: the AUTHORITATIVE or the PARTICIPATIVE?

The authoritative type is particularly appropriate to organizational form inherited from the medieval church and military; to a life of vertical hierarchy, prescribed role responsibilities and delegated authority; to a highly competitive economic and educational system; to the current dominant values of efficiency, excellence, productivity, task performance and perfectionism; to the impersonality, alienation, loneliness, and indifference in our people; to a world of automation, programming, data processing and engineering; to a persuasive, public relations and marketing mode of interpersonal commerce; to a world continually at war, threatened by war or preparing for war, i.e., to a *world of machines.*

The participative type of leadership is excellent for education for growth, intimacy, humanness and creativity; with the Christian ethics of love, honesty, faith, cheek-turning and brotherhood. It provides a good climate for research, inquiry, scholarship, contemplation and learning. It works well with cooperation, group planning, team building and various successful forms of group effort. It has been used with effectiveness in many new and exciting developments in education, architecture, the creative arts, economics, management and all phases of modern life, in brief, with the *world of human beings,* with people.

However, one conclusion reached after exhaustive studies by behavioral scientists at Harvard, M.I.T., Michigan, Chicago, U.C.L.A., Stanford and elsewhere is that the PARTICIPATIVE or GROUP approach to leadership does not seem to work *with all people and all situations.* Research has shown that satisfied, happy workers are sometimes more productive — and sometimes merely happy. Some managers and workers are able to take only limited responsibility, however much the company tries to give them. Some people recognize the need to delegate but simply "can't let go". It has also been discovered that in a profit squeeze the only way to get costs under control fast enough is by centralized, authoritative "get tough" management.

Nevertheless few, if any, behavioral scientists would advocate a general return to authoritative management. Instead, they are seeking a more thorough, systematic way to apply participative principles on a sustained schedule that will give the theory a better chance to work. Others believe that management must be tailor-made, suited to the work or the people, rather than packaged in a standard mixture. Some people are not and never will be suited for democracy on the job, according to other experts.

It, therefore, remains as a challenge to each one of you as well as to me to find the type of leadership best suited to a particular situation. I am confident that the training you received here at the U.P. Statistical Center will stand you in good stead in meeting this challenge and arriving at the proper solution.

The Philippine Statistician
Vol. XIX, Nos. 1 and 2, 1970

ON THE APPLICATION OF STATIONARY PROCESSES IN ECONOMICS

by

JOSE S. GUTIERREZ*

Models embodying two basic types of relationships are derived in economic theory. The first kind are purely definitional for example, savings plus investment equals income and expenditure on a commodity is its price times the quantity bought. The other type is of structural relationships, which describe the behavior of the individuals in the economy. Production functions, supply functions and demand functions are examples of the latter type. Structural parameters are involved in the latter type of relationships which have to be estimated by means of statistical methods. Elasticity of demand with respect to price, elasticity of demand with respect to income, marginal propensity to consume and marginal productivity, are examples of these structural parameters. The estimation of these parameters is difficult because the empirical data are actually the result of the interaction of many structural relationships. But, economic theory can only be useful when it has been connected with empirical facts which in general are of numerical kind[8]. However, even the use of appropriate descriptive statistical methods fails to present clear patterns which can be easily interpreted in terms of economic theory. Besides, there are usually other factors which are unmeasurable other than those considered influencing the behavior of such patterns thus there is no assurance that the behavior of such patterns are not due to random causes not considered in the theory. It is for this reason that the use of theoretical statis-

* Secretary and Associate Professor, U.P. Statistical Center.

tics is necessary in an attempt to sort out certain factors which may be regarded as random influence (5).

Limitations of some ordinary statistical models. Economic models based on theoretical statistics are set up according to some probability distribution with the assumption that such distribution fit the observation on hand best. Economic probability models may also throw some light on economic theory. Some dynamic models may be shown to be intrinsically unstable when small random influences are incorporated into the modal. A plausible theory explaining the quasi-cyclic behavior of some economic quantities which are determined by a set of equations is the system is disturbed by random shocks. Certain values of the constants although stable may show a tendency to overshot its equilibrium values and thus show a quasi-cyclic behavior (5).

Difficulties arises in estimating the parameters involved in the model. Consider a model involving the parameters $\theta_1, \ldots \theta_r$ and the observations are a sample $x_1, \ldots, x_n$ of a single variate which has a probability distribution $f(x/\theta_1, \ldots, \theta_r)$. The estimators of $\theta, \ldots, \theta$ are functions of the sample $t_1(x_1, \ldots, x_n)$, $\ldots, t_r\ (x_1 \ldots, x_n)$. The fundamental problem is whether the θ_i are in principle estimable, disregarding the desirable criteria of such estimators. For example, given a set of observations $\{x_i\}$ such that $x = u + v$, where u and v are independent normal variates with means m_1 and m_2 and standard deviations σ^2_1 and σ^2_2. It is easy to see that these four quantities are not estimable separately and are referred to as unidentifiable. However, $m_1 + m_2$ and $\sigma^2_1 + \sigma^2_2$ are estimable and are therefore identifiable.

Consider an empirical situation in which n pairs of values $(x_1\ y_1), \ldots, (x_n, y_n)$ have been measured and to make inferences from the data a probability model was set up. The relationship between x and y can be set up in at least four different models.

First, consider a simple linear regression model with the usual assumptions that the x_i are simply mathematical quantities or variable with no probability distribution attached to them. While y_i have a probability distribution whose mean is a linear function of x. Further that this distribution is normal and the random error is constant and independent of x. Thus the probability set to which our references are related, then consists of the joint distribution of all the y's for the fixed observed sets of the x's. Using this distribution, various tests of significance and estimation can be set up.

Suppose the following model is used,

$$y_i = \beta_o + \beta_1 x_i + \epsilon_i (i = 1, 2, \ldots n) \qquad (1)$$

where x_i is coded so that $\bar{x} = O$ and $\beta_o = \mu_y$, the x's are assumed equally spaced and ϵ_i is a random error with variance σ^2. If a simple linear equation of the type $y_1 = b_o + b_1{}^x$ is fitted the $b_o = \bar{y}$ and $b_1 = \frac{S_{xy}}{S_x{}^2}$ and the variance of the predicted value $\vec{y_i}$ for a given $x = x'$ is

$$V_1 = V(\hat{y}_1) = \sigma^2 \left[\frac{1}{n} + \frac{x'^2}{S_x{}^2} \right] \qquad (2)$$

Now, suppose, instead a quadratic function is fitted, of the form

$$\hat{y}_2 = b_o + b_1 x + b_2 z \qquad (3)$$

where $z = x^2 - \frac{S_x{}^2}{n}$, is used, in order to have an orthogonal model so that $b_o = \bar{y}$. Since the x's are assumed equally spaced with $x = O$, $\frac{S_x{}^2}{n} = \frac{n^2\text{-}1}{12}$ then

$$V_2 = V(\hat{y}_2) = \sigma^2 \left(\frac{n}{1} + \frac{x'^2}{S_x{}^2} + \frac{z'^2}{S_z{}^2} \right) \qquad (4)$$

which is easily seen to be larger than V_2.

The conclusion is clear that the use of a quadratic estimation equation based on the data only may be a very inefficient procedure for determining expected values.

Suppose a linear model is fitted when the true model is quadratic. For example, the true model is

$$y_i = \beta_o + \beta_1 x_i + \beta_2 z_i + \epsilon_i \tag{5}$$

and the estimating equation used was

$$\hat{y}'_1 = b_o + b_1 x' \tag{6}$$

In this case $E(y') = \beta_o + \beta_1 x'$, whereas the true expected value is $E(y') = \beta_o + \beta_1 x' + \beta_2 z'$, indicating a bias of $\beta_2 z'$, If the concept of the mean square (MSE) is used to combine the bias and variance components, then

$$\begin{aligned} MSE(y') &= E[\hat{y}_i - E(y)]^2 \\ &= E[\hat{y}_1 - E(\hat{y})_1) + E(\hat{y}_1) - E(y')]^2 \\ &= V_1 + (\beta_2 z')^2 \\ &= \sigma^2 \left[\frac{1}{n} + \frac{x'^2}{S_x^2} \right] + (\beta_2 z')^2 \end{aligned} \tag{7}$$

As compared to V_2, on this basis the simpler linear estimating equation would still have a lower average mean square mean error if $\beta_2^2 < \frac{\sigma^2}{S_z^2}$. In other words, if the error variance is so large that the standard error of b_2 exceeds β_2 one should. not attempt to use more than a linear estimating equation (1).

When both variable are assumed to be random variables that is the pair (x_i, y_i) are jointly distributed in a biviarate probability distribution a model of a completely different kind is obtained. The statistic to show x_i and y_i are not distributed independently is the correlation coefficient

$$r = \frac{\Sigma(y_i - y)(x_i - x)}{\sqrt{\{ \Sigma(y_i - y)^2 \ \ \Sigma(x_i - x)^2 \}}} \tag{8}$$

and the test statistic is

$$t = \frac{\sqrt{n-2}}{\sqrt{1-r^2}} \tag{9}$$

Nicholson (1948) studied the problem of the shrinkage of the correlation coefficient when applying a given estimating equation to a new set of data. Excellent goodness-of-fit to the same data used in deriving the equation may not hold up if the same equation is used for another set of data. It seems reasonable to assume that a model which is based on general considerations (such as theory or previous experience) and not on the data being fitted (7).

Ordinary statistical models as those described above are not very useful for serially dependent data, hence the need to construct a new theory in which successive observation are serially dependent on each other. This leads to the theory of stochastic processes (5).

Some simple stochastic models. Stochastic processes involve time in a continuous or discrete manner. Data encountered in economics are always given at discrete intervals. Although such data are often not the value of a variate at particular time but over an interval, the simplification resulting from considering discrete moments of time separated by intervals of constant length of time is great.

The time interval considered is the year. Shorter time interval such as months or weeks may lead to the trouble of eliminating seasonal effects. Also, if there is a lag or serial dependence between the two series which will be spread over many more intervals than before will result to heavier computations.

Consider the processes which are stationary. Where a process is said to be stationary, if the joint distribution of any set (x_t, x_u, . . ., x_v) depends only on the differences between t, u, . . . , v and not their absolute values. It is convenient to assume that the joint distribution of any set of the x's is a multivariate normal distribution with mean zero and variance σ^2, that is, $E(x_t) = O$ and $E(x_t^2) = O$.

Define the serial correlation coefficients by

$$\sigma^2\rho_s = E(x_t x_t + s) = \sigma^2\rho\text{-}s \qquad (10)$$

and a serial covariance generating function,

$$S_x(z) = \sigma^2 \sum_{s=\infty}^{\infty} \rho_s z^s \qquad (11)$$

where z is a complex variable (Quenouillie, 1947). The series converges in a ring $1 - \delta \leq | z | \leq 1 + \delta$ ($\delta > 0$). This simplifies the mathematics involved with the simple processes since if a new process $\{y_i\}$ is defined by an equation of the form

$$y_t = \sum_{l=0}^{\infty} a_l y_{l-t} \qquad (12)$$

where $\sum_{t}^{\infty} a_l$ is dominated by a convergent series, with the serial covariance generating function given by

$$S_y(z) = (\sum_{l=0}^{\infty} a_l^{-l}) (\sum_{l=0}^{\infty} a_j z^j) S_x z \qquad (13)$$

A simple model obtained by defining a process $\{\epsilon_t\}$ in terms of a completely random stationary process $\{\epsilon_t\}$ by the relationship

$$x = \rho x_{t-1} + \epsilon_t \qquad (14)$$

which is an example of a simple Markov process. For $\{x_t\}$ to be a stationary process $| \rho | < |$ and by multiplying by x_{t-s} ($s > 0$) and taking the expectations, $\rho_s = \rho\text{-}s = \rho^s$. This process has the Markovian character that if x_t is known, the conditional distribution of any set of x's with suffixes greater than t, given x_t, is independent.

The process (14) can be generalized by defining $\{x_t\}$ in terms of $\{\epsilon_t\}$ by the relationship

$$x_t + a_1 x_{t-1} + \dots + a_r x_{t-r} = \epsilon_t \qquad (15)$$

which is a stationary process under the condition that all the roots of the equation

$$z^r + a_1 z^{r-1} + \dots + a_r = 0 \qquad (16)$$

lie inside the unit circle $|z| = 1$ is imposed. The serial covariance generating function equal to unity and is given by

$$(1 + a_1 z + \dots + a_r z^r)^{-1} (1 + a_1 z^{-1} + \dots + a_r z^{-r})^{-1} \qquad (17)$$

thus $\quad S_r(z) = (1 + a_1 z + \dots + a_r z^r)^{-1}$

$$(1 + a_1 z^{-1} + \dots + a_r z^r)^{-1} \quad .$$

This is sometimes known as an autoregressive series and by choosing the a's so that the roots of (16) are complex, the series show oscillating behavior with a tendency to overswing the mean.

For a finite moving scheme,

$x_t = a_o + a_1 \epsilon_{t-1} + \dots + a_r \epsilon_{t-r}$,

where $a_o > a_1$, $a_r > 0$ and $\{\epsilon_t\}$ is a completely random series and

$$S_x(z) = (a_o + a_1 z + \dots + a_r z^r)\ (a_o + a_1 z^{-1} + \dots + a_r z^r)$$

and $\rho_s = 0$ for $|s| < r$ (17a)

Taking a moving average of a completely random series does not necessarily introduces oscillation if all the $a_i \geq 0$. Oscillatory tendency occur if some of the weights a are negative. The seemingly oscillatory behavior arises from the fact that moving average with positive weights will convert an irregular looking series into one which shows an appearance of wandering about the mean and in which neighboring values tend to be alike.

Moving averages, especially with equal weights, have sometimes been used to remove trend. For a sequence $\{x_t\}$ the trend value at t may be estimated by

$$x_t = (x_{t-r} + x_{t-r+1} + \dots x_t + \dots + x_{t-r+1} + x_{t-r}) \quad (2r + 1)^{-1} \qquad (18)$$

and regard $\{x_t - x_t\}$ as the stationary process.

A word of caution may be in order, consider $\{y_r\}$ which is a sequence of mutually independent random variables. Moving averages of order r are defined by

$$x^n = (y_n + y_{n+1} + \dots y_{n+r+1})/r \qquad (18a)$$

It is easily seen that the $\{x^n\}$ are not a Markov process. If only positions can be measured the process is a non-Markovian even though it is indirectly defined in terms of a Markov process (2).

Suppose x_t is the sum of a finite number of strictly periodic terms together with a superimposed error given by

$$x_t = \sum_{1}^{r} A_s \text{ Cos } (a_s t + b_s) + \epsilon_t \qquad (19)$$

where $\{\epsilon_t\}$ is a completely random series. This model has been been referred to as that of concealed periodicities.

The theory of stochastic processes is useful in economics in the construction of models which may be fitted to observed series, thus shedding some light on economic theory. How economic statics requires a dynamic theory for the discussion of stability problems has been shown by Samuelson in 1947. With parameters appropriately chosen, most linear dynamic models produce damped oscillation of bounded amplitude as required in the explanation of trade cycle theory which if assured to be non-linear has to be explained in some other way. Endogenous model can be set up from the linear models by introducing stochastic element into the model. For example given a simple linear model of the form

$$x_t + ax_{t-1} + bx_{t-2} = 0 \qquad (20)$$

the constants could be appropriately chosen to produce damped oscillations. The process can be kept in continuing oscillatory state by introducing a random element on the right hand side of the equation. Thus the assumption of non-linearity and the existence of an exogenous oscillatory factor to account for the trade cycle could be avoided.

The problem of statistical ineference from observed series. First consider the estimation of the mean given by

$$\bar{x} = n^{-1} \sum_{1=1}^{n} x_1 \qquad (21)$$

with a variance

$$\frac{\sigma^2}{n} \{1 + 2 \sum_{s=1}^{n-1} (1 - \tfrac{s}{n}) \rho \} = \sigma^2_n \text{ . say} \qquad (22)$$

or if n is large

$$\frac{\sigma^2}{n} \{ 1 + 2 \sum_{S=1}^{\infty} \rho_s \} \qquad (23)$$

where σ^2 is Var x_t .

If there is no serial correlation $\sigma_n^2 > n^{-1}\sigma^2$ and $\bar{x}$ though unbiased and asymptotically efficient is usually not the most efficient. A weighted sum of the x_i with weights depending on the serial crrelations is the most efficient estimator. Since these serial correlation have to be estimated, the sample mean $\bar{x}$ is usually used.

To test for serial dependence, the first order serial correlation coefficient r, is used. Define the serial coefficient r, is used. Define the serial coefficient of order s :

$$r_{.} = n(n\text{-}s)^{-1} \frac{\sum_{t=1}^{n-s} (x_t - \bar{x})(x_{t-r} - \bar{x})}{\sum_{t=1}^{n} (x_t - \bar{x})^2} \tag{24}$$

where $\bar{x} = n \sum_{i=1}^{n} x_t$.

By setting $x_{n+t} = x_t$ this can be written

$$r_s = \frac{\sum_{t=1} (x_t - \bar{x})(x_{t+s} - \bar{x})}{\sum_{t=1}^{n} (x_t - \bar{x})^2} \tag{25}$$

The existence of tables make the application of this test not difficult.

T. W. Anderson, 1948, gives one sided 5%, 1% and 0.1% levels of significance for n = 4, 5, . . . , 60 for r_1, where

$$r_1 = \frac{1/2(x_1 - \bar{x})^2 + 1/2(x_n - \bar{x})^2 + \sum_2 (x_i - \bar{x})(x_{i-1} - \bar{x})}{\sum_1^n (x_1 - \bar{x})^2} \tag{26}$$

while in terms of Von Neumann's mean square successive difference

$$\frac{\delta^2}{S^2} = \frac{n}{n-1} \frac{\sum_2^n (x_i - x_{i-1})^2}{\sum_1^n (x_i - x)^2} = \frac{2n}{n_t}(1 - r_1) \tag{27}$$

B. I. Hart tables for $\frac{\delta^2}{S^2}$ and R. L. Anderson has given tables for r_1 (26).

In what sense these tests are optimal have been considered by Anderson (1948). Take the alternative hypotheses that the x_t are generated by the simple Markov process $x_t = \rho x_{t-1} + \epsilon_t$ where $| \rho | <|$ the $\{ \epsilon_t \}$ are a completely random series, no uniformly most powerful test exists even for one sided alternative. The test based on Von Neumann's ratio or on the first order serial correlation may be regarded optimal since r as defined above is closed to the maximum likelihood estimate of ρ for a simple Markov process generated by

$$(x_t - m) = \rho(x_t - m) + \epsilon_t \tag{27a}$$

when m is unknown.

It should be mentioned, however, that r is quite a biased estimator of ρ for small samples.

Consider a simple Markov process generated by a relationship of the form

$$x_n = \rho x_{n-1} + \epsilon_n \tag{27b}$$

where $\rho_s = \rho^s$ and define the partial serial correlation by

$$\rho_{2 \cdot 1} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2} \tag{28}$$

which is estimated by

$$r_{2.1} = \frac{1 - r_1^2}{r_2 - r_1} \tag{29}$$

This can be used to test the hypotheses that the process is a simple Markov process against the alternative that further terms need to be included in the generating relation. This test approximate the likelihood ratio criterion.

Periodigram can be a useful method if the population correlogram defined by $\{\rho_s, s = 0, + 1 = \dots\}$ and the integral spectrum defined by $W(\theta)$ are in a kind of Fourier transformation as given the Wold-Khintchine theorem.

The necessary and sufficient condition that any arbitrary sequence of real constants $\{\rho_s\}k = O, \pm 1, \dots$ are the serial correlations of some stationary process is that there exists a non-decreasing function $W(\theta)$ such that

$$W(\theta) = O \qquad W(\pi) = \pi \tag{30}$$

and

$$\rho_r = \frac{1}{\mu} \int_O^{\pi} \cos r\,\theta\, d\, W(\theta)$$

$\sigma^2 W(\theta)$ is then known as the integrated power spectrum of the process and $\sigma^2 W(\theta)$ if it exists is known as the spectral density. If the $W(\theta)$ exists everywhere in the interval (O, π) and satisfies certain wide regularity conditions, it is given by

$$W^1(\theta) = 1 + 2 \sum_1^{\infty} \rho_r \text{ os } r\,\theta = S(e^{\,i\theta}) \tag{31}$$

For an observed series $\{x_t\}$ with say, a known zero mean, the periodigram ordinate I_p, for a given value p is defined by

$$I_p = A_p^2 + B_p^2$$

where

$$A_p = \sqrt{\frac{2}{n}} \sum_1^n x_t \cos\left(\frac{2\pi pt}{n}\right) \tag{32}$$

$$B = \sqrt{\frac{2}{n}} \sum_1^n x_t \sin\left(\frac{2\pi pt}{n}\right)$$

Setting the serial coveriance $c_s = \frac{1}{n\text{-}s} \sum_1^{n-s} x_t x_t + s$ then

$$I_p = 2 \sum_{s=n+1}^{n-1} \left(1 - \frac{|\delta|}{n}\right)(s \cos\left(\frac{2\pi ps}{n}\right) \tag{33}$$

and $E(I_p). = 2\sigma^2 \sum_{s=n+1}^{n-1} \left(1 - \frac{|s|}{n} ps \cos\left(\frac{2\pi ps}{n}\right)\right)$

The spectrum may be calculated directly and a goodness of fit of Kolmogoroff's type may be used.

It seems many of the methods used in the analysis of economic time series are not very robust with respect to serial dependence. This is more serious matter than non-robustness with respect to normality since if there is doubt about normality, transformation or parametric methods could be used.

Multivariate processes. Suppose all the variables are variates and if there are p of them and represent them by a prandom vector x_t with p components and a transpose

$$x_t' = (x_t^{(1)} \dots x_t^{(p)}) \tag{34}$$

The serial covarianσe matrix for any given S(= O, ± 1, . . .) is defined;

$$(C_s^{\ jr}) = E\{x_t x' \ \} = (C^{rj})' = \begin{Bmatrix} C_s^{11} & C_s^{12} & C_s^{13} & \dots & C_s^{1p} \\ C_s^{21} & C_s^{22} & C_s^{23} & \dots & C_s^{pr} \\ \vdots & & & & \\ C_s^{p1} & C_s^{p2} & C_s^{p3} & \dots & C_s^{pp} \end{Bmatrix}$$

where $C_p^{\ jr} = E\{x_t^{(j)} x^r_{\ t\text{-}r}\} = C_{-s}^{\ rj}$

The generalization of Wold-Khintσhine theorem given by Cramer: The nevessary and sufficient conditions that any given set of matrices $\{C_s^{\ jr}\}$ (s = O, ± 1 . . .) be the covariance matrices of a stationary p-variate process are that there exist p (possibly complex) functions of W(θ) that are defined and of bounded variation for $-\pi \leqq \theta \leqq \pi$ and such that

$$C_s^{\ jr} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \epsilon^{-is\theta} a W_{jr}(\theta) \tag{36}$$

For j = k $W_{jr}(\theta)$ is real and nondecreasing.

The matrix covariance generating function given by

$$S(z) = (\sum_{-\infty}^{\infty} C_s^{\ jr} z^s) \tag{37}$$

will be convergent in an annulus $1 - \delta \leqq | z | \leqq 1 + \delta$ $(\delta > O)$ $W'_{jr}(\theta)$ will exist and be given by

$$W'_{jr} (\theta) = \sum_{-\infty}^{\infty} C_s^{rj} \epsilon^{i s \theta} \quad (38)$$

Now define a vector process $\{y_t\}$ as a moving average matrix as follows:

$$y_t = A_o x_t + A_r x_{t-1} + \ldots \quad (38a)$$

where each A_i is a p x p matrix, such that each of the (i,j) elements of the A is an absolutely convergent series. Then if the matrix covariance generating function of the $\{y_i\}$ process is denoted by $S_y(z)$, it can be easily seen that

$$S_y z = \sum_{m=O}^{\infty} (A_m z^{-m}) \; S(z) (\sum_{r=O}^{\infty} A^m z^n) \quad (39)$$

The vector analogue of the stationary autoregressive process in one variate is given by the equation

$$X_t + A_1 X_t + \ldots + A_r X_t - r = \eta_t \quad (40)$$

where the A's are p x p amtrices and η is a column vector of variates whose variance-covariance matrix is $B = b_i$ and and such that the $\eta_{t,s}$ for different values of t are all independent. To insure that such an equation will generate a stationary process it is necessary that to impose some conditions on the A's and this is found to be that the roots of the equation

$$| 1 + \sum_{1=1}^{r} A_l z^{-1} | = 0 \quad (41)$$

must all lie outside the unit circle $|z| = 1$. When this is satisfied the matrix covariance generating function of the process is found to be

$$(1 + \sum_{1=1}^{r} A_l z^l)^{-1} \; B (1 + \sum_{1=1}^{r} A_l z^l)^{-1} \quad (42)$$

which exists inside some annulus $1 - \delta \leqq |z| \leqq 1 + \delta$ $(\delta > 0)$

For the estimation of the means the process (40) is put in the more general form

$$X_t + A_1X_{t-1} + \dots + A_rX_{t-r} + a = \eta_t \qquad (43)$$

where $E(x_t) \neq 0$. The η_t are assumed to be serially independent vectors. In the fitting a system of the type (40) to a set of observations the choice of r such that the resulting estimated η_t may be regarded as serially independent needs further study. Assuming that the elements of η_t have a diagonal variance-covariance matrix (σ_{ij}) and also that they are normally distributed, the maximum likelihood estimators will be consistent and asymptotically normally distributed the maximum likelihood estimators of the A's and are found by minimizing the sum of squares and, therefore, coincide with the latest square estimates. Even if the x's are not normally distributed these variances and covariances which can be estimated (Mann and Wald).

Now consider two stationary series $\{x_n\}$ $\{y_n\}$ with serial correlations ρ_s and ρ'_s. The variance of the product moment correlation coefficient between these two series based on n pairs of observation is given approximately

$$\frac{1}{n-1}\{1 + 2\sum_{s=1}^{n-1}(1 - \frac{s}{n})\ \rho_s\ \rho'_s \qquad (44)$$

where the second factor is usually greater than unity in many economic applications.

An approximation of the distribution of r is given by the following

$$1 + (n-1)\ [1 + 2\sum_{S=1}^{n-1}(1 - \frac{S}{n})]^{-1} \qquad (45)$$

The estimate r_s, r'_s cannot simply be substituted for ρ_s and ρ'_s since the standard error of sampling of the second is comparable to its mean. The best procedure is to fit an autoregressive series of low order and calculate the higher order serial correlation coefficients from the coefficients of this relation.

If the two rocesses are simple Markovian rocesses $\rho_s = \rho_1^s$ and $\rho'_s = \rho'_s)^s$ hence,

$$1 + 2 \sum_{s=1}^{n=1} (1 - \frac{s}{n}) \rho_s \rho'_s = n \frac{1 + \rho_1 \rho_1'}{1 - \rho_1 \rho_1'} -$$

$$2 (1 - \rho_1 \rho_1')^{-2} \{ (\rho_1 \rho_1')^n (1 - \rho \rho_1 \rho_1') - \qquad (46)$$

$$\rho_1 \rho_1' [1 - n(\rho_1 \rho_1')^{-n} + n-1(\rho_1 \rho_1')^n] \}$$

which is asymptotically equal to $n \frac{1 + \rho_1 \rho_1'}{1 - \rho_1 \rho_1'}$.

The estimates r and r^1 may now be substituted but these are biased in the small samples and may lead to an under-estimated variance of r .

Consider the case where regression is the appropriate model. Suppose $\{y_t\}$ is the series whose regression on $\{x_t\}$ is to be examined. Assume that $y = \alpha + \beta x_n + \epsilon_n$ where $\{\epsilon_n\}$ is another process. To test whether $\{\epsilon_r\}$ is serially correlated or not is calculate the serial correlation coefficient of the residuals and use it as a test criterion.

The best linear unbiased estimator for β could be obtained by minimizing the quadratic form

$$(y - x\beta)^1 \Gamma^{-1} (y - x\beta) \qquad (47)$$

where y′ is the row vector $(y_1, \ldots y_n)$, β is the column vector of the regression coefficients and x is the matrix of regression variables. Γ is the variance-covariance matrix of the residuals, usually this is unknown and assuming the serial correlations of the residuals are known.

This is just a partial review of some of the stationary processes which could be used in shedding some light on some economic theory. It might be worthwhile in this connection to quote Hicks remarks on Samuelson's econometric work; "... It may well be that for econometric work a theory of Professor Samuelson's type is all we need; it gives a superb model for statistical fitting. But for the understanding of the economic system we need something more, something which does refer back, in the last resort, to the behaviour of people and the motives of their conduct..."

LITERATURE CITED

1. Anderson, R. L. 1956. Some statistical problems in the analysis of fertilizer response data. Economic and Technical Analysis of Fertilizer Innovations and Resource Use ISC Press. Ames 187-206.

2. Feller 1950. An introduction to probability theory and its applications. J. Wiley. New York.

3. Hicks, J. R. 1961. Value and capital. An inquiry into some fundamental principles of economic theory. Oxford.

4. Mann, H. B. and Wold, A. 1943. On the statistical treatment of lenear stochastic difference equations. Econometrica 11: 173-220.

5. Moran, P. 1959. Random processes in economic theory and analysis. Sankhya: The Indian Journal of Statistics 21: 99-125.

6. ———— 1956. A test for the serial dependence of residuals. Bimetrika 37: 178-181.

7. Nicholson, G. 1948. The application of a regression equation to a new sample. (cited by R. L. Anderson, I).

7a. Quenouillie, M. 1947. Notes on the calculations of auto-correlations of linear autoregressive schems. Biometrika 34: 365-367.

8. Tintner, G. 1958. Economertics. J Wiley. New York.

The Philippine Statistician
Vol. XIX, Nos. 1 and 2, 1970

THE DISTRIBUTION OF MANILA ANNUAL RAINFALL DATA[1]

by

RUDY H. TAN[2]

1. Introduction

The distribution of Manila annual rainfall data is usually assumed to be normal especially in climatological and forecasting works. But the real distribution is slightly skewed to the right, and any estimates at the tail ends of the curve have very wide confidence limits. The estimates, however, can be further improved by choosing a non-symmetrical distribution.

In this paper, it is assumed that the probability distribution of the Manila annual rainfall data is a gamma distribution. The choice is not arbitrary and is the most reasonable after considering several other distributions. It is based on the fact that annual rainfall data are known to be positively skewed with "cut-off" observation at the origin and unlimited to the right.

2. Source of the Data,

The data for this study are from both published and unpublished reports of the Climatological Division, Philippine Weather Bureau. Observations are for Manila area only. The data began in 1965, the first year of weather observations in the Philippines.

[1] This paper is part of the thesis of the same title submitted by the author in partial fulfillment of the requirements for the degree of M. A. Statistics, U.P., 1969.

[2] Formerly meteorologist of the Weather Bureau. Now instructor at the Statistical Center, U.P.

3. Estimation of Missing Observations

Data for the war years from 1941 to 1945 are not available, and are considered missing observations. Although it would be reasonable to exclude from the analysis the missing observations, it is sometimes necessary to have objective estimates before certain analysis can be performed like testing for serial correlation. Also, any increase in the number of observations usually improve the fit of the theoretical distribution to the observed frequencies.

The scheme employed here to estimate the missing observations assumes that the data can be arranged in p rows and q columns. Denoting by $\bar{R}_i$ the mean of the ith row, $\bar{C}_j$ the mean of the jth column, and $\bar{G}$ the total mean, the estimate $\hat{z}_{ij}$ of the missing observation in the ith row and jth column is given by the following formula:

$$\hat{z}_{ij} = \bar{R}_i + \bar{C}_j - \bar{G} \qquad (1)$$

It can be shown that if pq = n, the above estimation equation reduces to the formula for estimating a single missing value in a p x q classification. Unlike the well-known method of estimating several missing observations by the iterative application of the formula for a single missing value, the present scheme eliminates iteration completely by simply formulating as many equations as there are missing observations and solving these equations simultaneously. Using this scheme, the following estimates of the missing observations are obtained: $X_{1941} = 80.42$, $X_{1942} = 76.07$, $X_{1943} = 73.41$, $X_{1944} = 75.74$, $X_{1945} = 69.09$.

4. The Gamma Distribution

A random variable X is said to be distributed as the gamma distribution if its probability density function is

$$f(x) = \frac{1}{\Gamma(\alpha)\,\beta^{\alpha}}\, x^{\alpha-1}\, e^{-x/\beta} \qquad 0 < x < \infty \qquad (2)$$

where $\Gamma(\alpha)$ is the gamma function of α.

The gamma distribution is completely determined by the parameters α and β. α is the shape parameter and β the scale parameter. Its lower limit is zero for positive values of x and is unlimited to the right. It is also positively skewed, the amount of skewness depending inversely on the shape parameter α. For this reason it has been found useful in fitting distributions of meteorological variates which are restricted to positive values and characterized by few extremely high values. An example is rainful distribution.

4.1. Maximum Likelihood and Moment Estimators of α and β.

The problem is to estimate the parameters α and β from the sample observations. There are several estimation methods that will produce estimators with different attributes of "goodness" but the easiest to apply and one which possesses many desirable attributes is the method of Maximum Likelihood devised by R.A. Fisher.

Let $x_1, x_2, \ldots, x_n$ be a random sample of size n from the gamma p.d.f. $f(x;\ \alpha,\ \beta)$, so that the likelihood function is

$$L(x_i;\alpha,\beta) = \prod_{i=1}^{n} \frac{1}{\Gamma(\alpha)\,\beta^{\alpha}}\ x_i^{\alpha-1}\ e^{-x/\beta} \qquad (3)$$

The logarithm of the likelihood function is

$$L^*(x_i;\alpha,\beta) = \ln L(x_i;\alpha,\beta)$$

$$= -n \ln \Gamma(\alpha) - n\,\alpha \ln \beta + (\alpha - 1)\,\Sigma \ln x_i - \frac{1}{\beta}\,\Sigma\, x_i \qquad (4)$$

Differentiating partially with respect to α and β gives

$$\frac{\partial L^*}{\partial \alpha} = -n\,\frac{\Gamma'(\alpha)}{\Gamma(\alpha)} - n \ln \beta + \Sigma \ln x_i \qquad (5)$$

$$\frac{\partial L^*}{\partial \beta} = -n\,\frac{\alpha}{\beta} + \frac{1}{\beta^2}\ \Sigma\, x_i \qquad (6)$$

where $\Gamma'(\alpha) = \dfrac{\partial\, \Gamma(\alpha)}{\partial \alpha}$, and ln is natural logarithm to the base e. The maximum likelihood equations involving the estimates of α and β are obtained on equating (5) and (6) to zero. After simplifying, these are:

$$\ln \hat{\beta} = \frac{\Sigma \ln x_i}{n} - \Psi(\hat{\alpha}) \quad (7)$$

and

$$\bar{x} = \hat{\alpha}\hat{\beta} \quad (8)$$

where the caret (^) is used over a parameter to indicate that it is a maximum likelihood estimator of the parameter. $\Psi(\hat{\alpha}) = \frac{\Gamma'(\hat{\alpha})}{\Gamma(\hat{\alpha})}$ is called the "digamma function" and its values are extensively tabulated in "Tables of the Digamma and Trigamma Functions" by E. Pairman.

The complex nature of the maximum likelihood equations (7) and (8) does not readily yield to a direct solution of $\hat{\alpha}$ and $\hat{\beta}$. However, the method of successive approximations can be employed to effect a solution by choosing $\hat{\alpha}$ first and then solving for $\hat{\beta}$ by (8). A first trial value can easily be obtained by the method of moments.

Consider a random sample $x_1, x_2, \ldots, x_n$ of size n from the gamma distribution. Then the sample moments are

$$m'_1 = \frac{1}{n} \Sigma x_i \quad (9)$$

and

$$m'_2 = \frac{1}{n} \Sigma x_i^2 \quad (10)$$

The population moments can be obtained easily by differentiating the moment-generating function of the gamma distribution.

These are

$$\mu'_1 = \alpha\ \beta \tag{11}$$

and

$$\mu'_2 = \alpha(\alpha + 1)\beta^2 \tag{12}$$

The moment estimators are obtained by equating the population and sample moments. After simplifying these are:

$$\tilde{\alpha} = \frac{\bar{x}}{\tilde{\beta}} \tag{13}$$

and

$$\tilde{\beta} = \frac{\frac{1}{n} \Sigma\, x_i^2 - \bar{x}^2}{\bar{x}} \tag{14}$$

where the tilde ($\approx$) is used over a parameter to indicate that it is a moment estimator of the parameter. Although the moment estimators are simple, they are known to be inefficient and good only for obtaining a first approximation to the maximum likelihood estimators.

4.2. Variances of $\hat{\alpha}$ and $\hat{\beta}$

The maximum likelihood estimators θ_1 and θ_2 for the parameters of a density function $f(x;\theta_1,\theta_2)$ from a sample of size n follow, for large samples, approximately a bivariate normal distribution with means θ_1 and θ_2 and with information matrix nR where the elements of R are given by

$$r_{ij} = -E\left\{ \frac{\partial^2}{\partial\theta_i\,\partial\theta_j} \ln f(x;\theta_1;\theta_2) \right\} \quad i = 1,2;\ j = 1,2 \tag{15}$$

The variance-covariance matrix of the estimators is $\frac{1}{n}V$ where $V = R^{-1}$.

For a gamma distribution with parameters α and β,

$$\ln f = -\ln \Gamma(a) - a \ln \beta + (a - 1) \ln x - x/\beta \quad (16)$$

Taking the first and second partial derivatives:

$$\frac{\partial \ln f}{\partial a} = - \Psi(a) - \ln \beta + \ln x \quad (17)$$

$$\frac{\partial^2 \ln f}{\partial a^2} = - \Psi'(a) \quad (18)$$

where $\Psi'(a) = \frac{\partial \Psi(a)}{\partial a}$ is called the "trigamma function" of a and its values are also extensively tabulated in Pairman's.

$$\frac{\partial \ln f}{\partial \beta} = - \frac{a}{\beta} + \frac{x}{\beta^2} \quad (19)$$

$$\frac{\partial \ln f}{\partial \beta^2} = \frac{a}{\beta^2} - 2 \frac{x}{\beta^3} \quad (20)$$

and $$\frac{\partial \ln f}{\partial a \, \partial \beta} = - \frac{1}{\beta} \quad (21)$$

Taking expected values of (18), (20), and (21), the following elements of the R matrix are obtained:

$$r_{11} = - E \left\{ \frac{\partial^2 \ln f}{\partial a^2} \right\} = \Psi'(a) \quad (22)$$

$$r_{22} = - E \left\{ \frac{\partial^2 \ln f}{\partial \beta^2} \right\} = - \frac{a}{\beta^2} + \frac{2}{\beta_3} E(x)$$

$$= \frac{a}{\beta^2} \text{ since } \mu = E(X) = a \beta \quad (23)$$

$$r_{11} = r_{21} = - E \left\{ \frac{\partial^2 \ln f}{\partial a \, \partial \beta} \right\} = \frac{1}{\beta} \quad (24$$

Now,

$$\frac{1}{n} V = \frac{1}{n} \begin{bmatrix} \Psi'(\alpha) & \frac{1}{\beta} \\ \frac{1}{\beta} & \frac{\alpha}{\beta^2} \end{bmatrix}^{-1}$$

$$= \frac{\beta^2}{n\,[\alpha\,\Psi'(\alpha) - 1]} \begin{bmatrix} \frac{\alpha}{\beta^2} & -\frac{1}{\beta} \\ -\frac{1}{\beta} & \Psi'(\alpha) \end{bmatrix}$$

Thus, the variances and covariances of $\hat{\alpha}$ and $\hat{\beta}$ are the following:

$$\text{Var}\ (\hat{\alpha}) = \frac{\hat{\alpha}}{k} \qquad (26)$$

$$\text{Var}\ (\hat{\beta}) = \frac{\hat{\beta}^2\,\Psi'(\hat{\alpha})}{k} \qquad (27)$$

$$\text{Cov}\ (\hat{\alpha}, \hat{\beta}) = -\frac{\hat{\beta}}{k} \qquad (28)$$

$$\text{where} \quad k = n\,[\hat{\alpha}\,\Psi'(\hat{\alpha}) - 1] \qquad (29)$$

4.3. Confidence Intervals for α and β

For a large sample distribution,

$$\hat{\alpha} \text{ is } N\,[\alpha,\ \text{var}\ (\hat{\alpha})] \qquad (30)$$

Hence,

$$z = \frac{(\hat{\alpha} - \alpha)\sqrt{k}}{\sqrt{\hat{\alpha}}} \text{ is } N(0,1) \qquad (31)$$

where var $(\hat{\alpha})$ and k are given in (26) and (29), respectively. To obtain a 95% confidence interval estimate for α, the probability that the standard normal variate z will fall between −1.96 and + 1.96 is computed, for example,

$$\Pr(-1.96 < z < +1.96) = 0.95 \qquad (32)$$

Substituting for z in (32) gives

$$\Pr\left(-1.96 < \frac{(\hat{\alpha} - \alpha)\sqrt{k}}{\sqrt{\hat{\alpha}}} < 1.96\right) = 0.95 \qquad (33)$$

or equivalently,

$$\Pr\left(\hat{\alpha} - 1.96\sqrt{\frac{\hat{\alpha}}{k}} < \alpha < \hat{\alpha} + 1.96\sqrt{\frac{\hat{\alpha}}{k}}\right) = 0.95 \quad (34)$$

so that the 95% confidence limits for α are

$$\hat{\alpha} \pm 1.96\sqrt{\frac{\hat{\alpha}}{k}} \qquad (35)$$

Similarly,

$$\hat{\beta} \text{ is N } [\beta, \text{var}(\hat{\beta})] \qquad (36)$$

Hence,

$$z = \frac{(\hat{\beta} - \beta)\sqrt{k}}{\beta\sqrt{\Psi'(\alpha)}} \text{ is N(0,1)} \qquad (37)$$

It follows that the 95% confidence limits for β are given by

$$\hat{\beta} \pm 1.96\,\beta\sqrt{\frac{\Psi'(\hat{\alpha})}{k}}$$

4.4. Confidence Interval for a Point Estimate X_o

Consider the problem of estimating the value of X, say X_o, corresponding to some probability value $Pr(X \geq X_o)$. Accordingly, X_o will be a function of α and β, the parameters of the gamma distribution. Mathematically,

$$X_o = f(\alpha, \beta) \qquad (39)$$

Defining ΔX_o, $\Delta \alpha$, $\Delta \beta$ as the variations from their true values, and if these variations are sufficiently small,

$$\Delta X_o = \frac{\partial X_o}{\partial \alpha} \Delta \alpha + \frac{\partial X_o}{\partial \beta} \Delta \beta \qquad (40)$$

Squaring and averaging both sides of the equation:

$$\frac{\Sigma (\Delta X_o)^2}{n} = \left(\frac{\partial X_o}{\partial \alpha}\right)^2 \frac{\Sigma(\Delta \alpha)^2}{n} + 2\left(\frac{\partial X_o}{\partial \alpha}\right)\left(\frac{\partial X_o}{\partial \beta}\right)\frac{\Sigma(\Delta\alpha)(\Delta\beta)}{n}$$

$$+ \left(\frac{\partial X_o}{\partial \beta}\right)^2 \frac{\Sigma(\Delta\beta)^2}{n} \qquad (41$$

or,

$$\sigma^2 x_o \doteq \left(\frac{\partial X_o}{\partial \alpha}\right)^2 \operatorname{var}(\hat{\alpha}) + 2\left(\frac{\partial X_o}{\partial \alpha}\right)\left(\frac{\partial X_o}{\partial \beta}\right) \operatorname{cov}(\hat{\alpha},\hat{\beta}) +$$

$$\left(\frac{\partial X_o}{\partial \beta}\right)^2 \operatorname{var}(\hat{\beta}) \qquad (42)$$

where the partial derivatives are computed at the points $\alpha = \hat{\alpha}$ and $\beta = \hat{\beta}$. Note that this is only an approximate variance.

Now

$$X_o = u_o \beta \sqrt{\alpha} \qquad (43)$$

where u_o is an argument in the Pearson's integral so that $I(u,p) = Pr(X \geq X_o)$.

Thus,

$$\frac{\partial X_o}{\partial \alpha} = \frac{u_o \hat{\beta}}{2\sqrt{\hat{\alpha}}} \tag{44}$$

and

$$\frac{\partial X_o}{\partial \beta} = u_o \sqrt{\hat{\alpha}} \tag{45}$$

give approximate values for the partial derivatives as unit change in X_o for unit change in the corresponding parameter.

For a 95% confidence interval for X_o, the following are the limits:

$$X_o \pm 1.96\ \sigma_{X_o} \tag{46}$$

where σ_{X_o} is the standard error of X_o obtained from (42).

5. Fitting Gamma And Normal Distributions To The Annual Rainfall Data

5.1. Testing the Data for Serial Correlation

The data are first tested for serial correlation. Although there is sufficient reason to believe that the successive observations are independent, there is no better way of showing this than by actual computation. For this purpose, the familiar Pearson's product-moment coefficient of correlation r is used with only a slight modification in form.

The serial correlation coefficient of the first order, or briefly the first serial correlation, is defined as

$$r_1 = \frac{(n-1)\sum_{i=1}^{n-1} x_i x_{i+1} - \left(\sum_{i=1}^{n-1} x_i\right)\left(\sum_{i=1}^{n-1} x_{i+1}\right)}{\sqrt{(n-1)\sum_{i=1}^{n-1} x_i^2 - \left(\sum_{i=1}^{n-1} x_i\right)^2}\sqrt{(n-1)\sum_{i=1}^{n-1} x_{i+1}^2 - \left(\sum_{i=1}^{n-1} x_{i+1}\right)^2} \tag{47}$$

The above formula is essentially the same as Pearson's product-moment coefficient of correlation r. In the same manner that r measures the extent of association between two variables X and Y, r_1 measures the degree of relationship between successive values within a series in which the correlated values are a constant time interval apart. The coefficient r_1 may have also a value between -1 and $+1$. A value of $\pm$ 1 indicates serial correlation. This means that there is a trend, linear or curvilinear. A zero value indicates that the sequence of observations arises from a purely random process.

For the data, the following quantities are computed:

$$n = 104 \quad \sum_{i=1}^{103} x_i = 8{,}384.58 \qquad \sum_{i=1}^{103} x_{i+1} = 8{,}370.60$$

$$\sum_{i=1}^{103} x_i^2 = 717{,}641.0350 \qquad \sum_{i=1}^{103} x^2_{i+1} = 715{,}705.5982$$

$$\sum_{i=1}^{103} x_i x_{i+1} = 678{,}993.0006$$

Substituting these values in (47) gives $r = -0.07$ which is negligible. This result is to be expected since rainfall data become serially correlated only if rains come from a single rain-producing mechanism. This is particularly true of daily and weelky rainfall data. Annual rainfall totals are made up of rainfalls caused by different weather disturbances such as tropical cyclones, easterly waves, cold fronts, the monsoons, the intertropical convergence zone, and local disturbances like thunderstorms and orographic lifting. Besides, there are months without any rainfall. Thus, the sequence of annual rainfall observations are serially independent, that is, the value of X_i is unaffected by any of the remaining value of X.

5.2. Fitting the Gamma Distribution to the Data

5.2.1. Computations of the Estimated Parameters.

The problem is to fit the gamma distribution to the original data on annual rainfall totals. For these data, the following are the required computations.

$n = 104$ $\qquad \Sigma \log x = 197.44176$

$\Sigma x = 8{,}447.14$ $\qquad \Sigma \ln x = 454.62645$

$\Sigma x^2 = 721{,}554.7886$ $\qquad \frac{\Sigma \ln x}{n} = 4.37141$

$\bar{x} = 81.22250$

$\bar{x}^2 = 6{,}597.29756$ $\qquad s = \sqrt{\frac{104(721{,}554.7886)-(8{,}447.14)^2}{(104)(103)}}$

$\frac{\Sigma x^2}{n} = 6{,}938.02681$ $\qquad = 18.55377$

The moment estimators are

$$\tilde{\beta} = \frac{\frac{\Sigma x^2}{n} - \bar{x}^2}{\bar{x}} = \frac{6{,}938.02681-6{,}597.29756}{81.22250} = 4.19501$$

and

$$\tilde{\alpha} - \frac{\bar{x}}{\tilde{\beta}} = \frac{81.22250}{4.19501} = 19.36$$

The maximum likelihood equations are

$$\ln \hat{\beta} = 4.37141 - \Psi(\hat{\alpha})$$

and $$\hat{\beta} = \frac{81.22250}{\hat{\alpha}}$$

Choosing successive values of $\hat{\alpha}$ close to the trial value $\hat{\alpha} =$ 19.36, say 19.00, 19.10, 19.30, 19.40, 19.50, 19.52, 19.54, and 19.56 the corresponding values of $\hat{\beta}$ are calculated independently from both maximum likelihood equations until the same value of $\hat{\beta}$ results. From Table 1, this value of $\hat{\beta}$ is 4.16 which corresponds to a value of $\hat{\alpha} = 19.54$. These are now the maximum likelihood estimators of α and β, respectively.

Referring to Table 1, column (1) contains successive values of $\hat{\alpha}$ with their corresponding values of $\Psi(\hat{\alpha})$ tabulated in column (2). Column (3) contains the logarithms of $\hat{\beta}$ ob-

TABLE 1

COMPUTATIONS FOR $\hat{\alpha}$ AND $\hat{\beta}$

$\hat{\alpha}$	$\Psi(\hat{\alpha})$	$x = \ln \hat{\beta}_1$	$\hat{\beta}_1 = e^x$	$\hat{\beta}_2 = \bar{x}/\hat{\alpha}$	$\hat{\beta}_1 - \hat{\beta}_2$
19.00	2.91789	1.45352	4.27806	4 27487	+0.00319
19.10	2.92328	1.44813	4.25502	4.25259	+0.00253
19.30	2.93397	1.43744	4.20973	4.20842	+0.00131
19.40	2 93928	1.43213	4.18748	4.18673	+0.00065
19.50	2.94455	1.42686	4.16576	4.16526	+0.00050
19.52	2.94561	1.43580	4.16118	4.16099	+0.00019
19.54	2.94666	1.42475	4.15673	4.15673	+0.00000
19.56	2.94771	1.42370	4.15245	4.15248	—0.00003

tained by subtracting each entry in (2) from 4.37141. Taking the antilogarithms of entries in (3), the values of $\hat{\beta}$ (indicated by $\hat{\beta}_1$) shown in column (4) are obtained. Column (5) contains successive values of $\hat{\beta}$ (indicated by $\hat{\beta}_2$) obtained by dividing $\bar{x} = 81.22250$ by the corresponding successive values of $\hat{\beta}$. The small differences between $\hat{\beta}_1$ and $\hat{\beta}_2$ are shown in column (6). Note that $\hat{\beta} = 19.54$ yields the difference of 0.00000 between $\hat{\beta}_1$ and $\hat{\beta}_2$.

5.2.2. Grouping the Data into a Frequency Distribution and Computations of Expected Frequencies

Before testing whether or not the observed distribution fits the pattern of the gamma distribution described by the estimated parameters $\hat{\alpha} = 19.54$ and $\hat{\beta} = 4.16$, the first step is to group the data into a frequency distribution. The range of the observations is 154.35 - 35.69 = 115.66 so that grouping the data into intervals of size 10 would give 12 classes. Since the observations form a continuous series and also to facilitate definition to which class an observation belongs, the classes are constructed from 35-45, 45-55, 55-65, . . . , 145-155. The next step is to calculate the value of

$$u = \frac{x}{\hat{\beta}\sqrt{\hat{\alpha}}}$$

for each class limit or class boundary. To illustrate, consider the class interval 75-85:

$$u_{75} = \frac{75}{4.16\sqrt{19.54}} = 4.08$$

and $$u_{85} = \frac{85}{4.16\sqrt{19.54}} = 4.62$$

Using Pearson's Table I and performing bi-variate interpolation, the following values for I(u,p) are obtained:

$$I(4.08,\ 18.54) = 0.39171$$
$$I(4.62,\ 18.54) = 0.60829$$

where $p = 19.54 - 1 = 18.54$.

The area under the curve between 75 and 85 is

$$0.60829 - 0.39171 = 0.21658$$

This implies that 21.66% of the annual rainfall totals could be expected to have a range of 75 to 85 if the distribution were really close to the gamma distribution having the same parameters as the actual data. The expected frequency is 21.66% of 104, the total frequency, or 22.5, which is close enough to the actual frequency of 22.

For the entire distribution, it is convenient to organize the work as shown in Table 2.

TABLE 2

COMPUTATION FOR THE GAMMA DISTRIBUTION

Class Limits	x	u	I(u,18.54)	Probabilities	Expected Frequencies	Observed Frequencies
35 - 45	35	1.90	.00069	.00855	0.9	2
45 - 55	45	2.45	.00924	.05311	5.5	4
55 - 65	55	2.99	.06235	.12407	12.9	10
65 - 75	65	3.53	.18642	.20529	21.4	25
75 - 85	75	4.08	.39171	.21658	22.5	22
85 - 95	85	4.62	.60829	.17782	18.5	20
95 - 105	95	5.17	.78611	.11008	11.5	12
105 - 115	105	5.71	.89619	.06032	6.3	4
115 - 125	115	6.25	.95651	0.2665	2.8	3
125 - 135	125	6.80	.98316	.01123	1.2	1
135 - 145	135	7.34	.99439	.00382	0.4	0
145 - 155	145	7.89	.99821	.00127	0.1	1
	155	8.43	.99948			

Referring to Table 2, column (1) contains the class limits, column (2) contains the class boundaries, and column (3) contains the corresponding u's. Column (4) contains the interpolated probabilities from Pearson's Table I corresponding to the u's in column (3), and column (5) contains the probabilities within the classes obtained by taking the differences of successive entries in column (4). Column (6) contains the probabilities in column (5) by 104, the total frequency. Column (7) contains the observed frequencies for comparison with the expected frequencies.

5.2.3. Testing for Goodness of Fit

To compare on precise basis the observed and expected frequencies, the chi-square criterion is calculated with the following formula:

$$\chi^2 = \Sigma \frac{(o_i - e_i)^2}{e_i}$$

where o_i is the observed frequency and e_i is the expected frequency. If the computed chi-square with k-3 degrees of freedom is very small compared with the tabulated chi-square the fit is good; if it is large the fit is bad. The number of degrees of freedom equals the number of pairs of frequencies k which are to be compared, minus the number of quantities used to calculate the expected frequencies; the total frequency n and the two parameters α and β.

Thus, the computed chi-square is equal to

$$\chi^2 = \frac{(6-6.4)^2}{6.4} + \frac{(10-12.9)^2}{12.9} + \frac{(25-21.4)^2}{21.4} + \frac{(22-22.5)^2}{22.5}$$

$$+ \frac{(20-18.5)^2}{18.5} + \frac{(12-11.4)^2}{11.4} + \frac{(9-10.8)^2}{10.8}$$

$$= 1.747$$

Note that it is necessary to combine the first two classes and the last five classes in following the rule that the expected frequencies must be at least 5. For d.f. = 7 — 3 = 4, $\chi^2_{0.95}$ = 9.488. Then since 1.747 < 9.488, it is concluded that the observed distribution does not differ very significantly from the gamma distribution; hence, the fit is very good at 95% level of confidence.

5.2.4. Confidence Intervals for the Parameters

The variances and covariance of $\hat{\alpha}$ and $\hat{\beta}$ are computed below:

$$k = n[\hat{\alpha}\Psi'(\hat{\alpha}) - 1]$$

$$= 104\,[(19.54)(0.05250895) - 1] = 2.70659$$

$$\text{Var}(\hat{\alpha}) = \frac{\hat{\alpha}}{k} = \frac{19.54}{2.70659} = 7.22$$

$$\text{Var}(\hat{\beta}) = \frac{\hat{\beta}^2\Psi'(\hat{\alpha})}{k} = \frac{(4.16)^2(.05250895)}{2.70659} = 0.34$$

$$\text{Cov}(\hat{\alpha}, \hat{\beta}) = -\frac{\hat{\beta}}{k} = -\frac{4.16}{2.70659} = -1.54$$

A 95% confidence interval for α is given by

$$19.54 \pm 1.96\sqrt{7.22}$$

that is, 19.54 ± 5.27, or 14.27 and 24.81. Similarly, a 95% confidence interval for β is

$$4.16 \pm 1.96 \sqrt{0.34}$$

or 3.62 and 5.30.

5.3. Fitting the Normal Distribution to the Data

It would also be interesting to fit the normal distribution to the same set of data and compare the resulting frequencies with that of the gamma distribution. The fitting entails no difficulty and details are shown in Table 3. In calculating z for the class limits, the following formula is used:

$$Z = \frac{x - \bar{x}}{s}$$

where $\bar{x} = 81.22250$ and $s = 18.55377$.

Referring to Table 3, column (1) contains the class limits, column (2) contains the class boundaries, and column (3) contains the corresponding z's. Column (4) contains the normal curve areas corresponding to the z's of column (3), and column (5) contains the differences between the successive entries of column (4). The last column (6) contains the expected frequencies obtained by multiplying each area of column (5) by 104, the total frequency.

For the fitted normal curve, the computed chi-square is

$$\chi^2 = \frac{(6-7.6)^2}{7.6} + \frac{(10-11.7)^2}{11.7} + \frac{(25-18.2)^2}{18.2} + \frac{(22-22.1)^2}{22.1} + \frac{(20-19.9)^2}{19.9} + \frac{(12-13.4)^2}{13.4} + \frac{(9-10.53)^2}{10.53}$$

$$= 3.821$$

$\chi^2_{.95} = 9.488$ for d.f. = 7–3 = 4 . Since 3.821 < 9.488, the fit

TABLE 3
COMPUTATION FOR THE NORMAL DISTRIBUTION

Class Limits (1)	x (2)	z (3)	Area under Normal Curve from - ∝ to z (4)	Area for Each Class (5)	Expected Frequencies (6)	Observed Frequencies (7)
	35.0	—2.49	.0064			
35 - 45				.0192	2.0	2
	45.0	—1.95	.0256			
45 - 55				.0537	5.6	4
	55.0	—1.41	.0793			
55 - 65				.1129	11.7	10
	65.0	—0.87	.1922			
65 - 75				.1747	18.2	25
	75.0	—0.34	.3669			
75 - 85				.2124	22.1	22
	85.0	+0.20	.5793			
85 - 95				.1911	19.9	20
	95.0	+0.74	.7704			
95 - 105				.1293	13.4	12
	105.0	+1.28	.8997			
105 - 115				.0659	6.9	4
	115.0	+1.82	.9656			
115 - 125				.0253	2.6	3
	125.0	+2.36	.9909			
125 - 135				.0072	0.8	1
	135.0	+2.89	.9981			
135 - 145				.0016	0.2	0
	145.0	+3.44	.9997			
145 - 155				.0003	0.03	1
	155.0	+3.97	1.0000			

is good, but not as good as the gamma distribution. Figure 1 compares the "goodness" of fit of the two distributions.

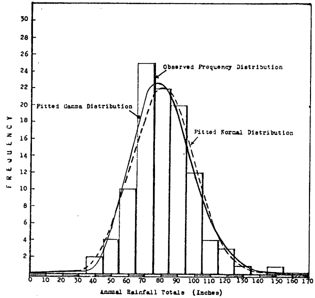
6. Estimating Probable Extreme Values

6.1. Probable 1000-Year Maximum Rainfall Total

Consider the problem of estimating the probable value of an exceedingly high annual rainfall total which will not be exceeded more than once in 1000 years, that is,

$$\Pr(X \geqq X_o) = \frac{1}{1000}$$

FIGURE 1


Observed Frequency Distribution
Fitted Gamma Distribution
Fitted Normal Distribution
FREQUENCY
30
28
26
24
22
20
18
16
14
12
10
8
6
4
2
0 10 20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170
Annual Rainfall Totals (Inches)

This probability corresponds to an area under the left-tail end of the fitted gamma curve. From Pearson's Table $u_o = 8.16$ to make $I(u_o, 18.54) = 1 - 0.001 = 0.999$, and therefore,

$$X_o = u_o \hat{\beta} \sqrt{\hat{\alpha}}$$
$$= 8.16\ (4.16) \sqrt{19.54}$$
$$= 150.04 \text{ inches}$$

Thus $X_o = 150.04$ inches or more is the estimated value of annual rainfall total exepcted to occur only once in 1000 years. The 1919 total is 154.35 inches and this observation has never been exceeded in the course of 100 years of weather observations in Manila. It is significant to note that this observation is singularly phenomenal considering that the next highest recorded observation is only 134.96 inches in 1923. Will the 1919 total stands out to be the highest in the next 900 years? Only time can tell. But it should be emphasized that in a real situation one would not normally make a very simple statistical prediction especially about the weather without considering possible changes in various climatological factors within the next 900 years.

Using the fitted normal distribution, $z_o = 3.09$ such that

$$\Pr\ (z > 3.09) = 0.999$$

Thus,

$$X_o = \bar{x} + z_o s$$
$$= 81.22 + 3.09(18.55)$$
$$= 138.54 \text{ inches}$$

which is quite a low estimate. To show further that the fitted gamma distribution gives a better estimate at the tail end of the curve, the confidence limits for each estimate are computed.

From (44), (45), and (42):

$$\frac{\partial X_o}{\partial \alpha} = \frac{(8.16)(4.16)}{2\sqrt{19.54}} = 3.84$$

$$\frac{\partial X_o}{\partial \beta} = (8.16)\sqrt{19.54} = 36.07$$

$$\sigma^2 x_o = (7.22)(3.84)^2 + 2(-1.54)(3.84)(36.07) + (0.34)(36.07)^2 = 122.21$$

Thus, the approximate standard error of the estimate from the fitted gamma distribution is ± 11.05 inches and a 95% confidence limits for X_o are 138.54 ± 1.96 x 11.05 = 128 and 172 inches. The confidence limits for X_o from the fitted normal distribution are 138.54 ± 1.96 x 18.55 = 102 and 175 inches. It should be noted that in both confidence limits only the upper limits are meaningful since the lower limits are very much lower than the observed 1919 total. Though the upper limits are close to each other from both distributions, the confidence limits from the normal distribution are considerably wider than that obtained from the gamma distribution. This result is to be expected since the gamma distribution asymptotically approaches the ordinate more rapidly than the normal distribution.

6..2. Probable 1000-Year Minimum Rainfall Total

For the problem of estimating the lowest amount of annual rainfall total to occur once in 1000 years, consider the area under the right tail-end of the fitted gamma curve from O to u_o such that $I(u_o, 18.54) = 0.001$. From Pearson's Table, $u_o = 1.95$ and therefore

$$X_o = 1.95(4.16)\sqrt{19.54}$$

$$= 35.86 \text{ inches or less}$$

For this estimate

$$\frac{\partial x_o}{\partial \alpha} = \frac{(1.95)(4.16)}{2\sqrt{19.54}} = 0.92$$

$$\frac{\partial x_o}{\partial \beta} = 1.95\sqrt{19.54} = 8.62$$

$$\sigma^2 x_o = (7.22)(0.92)^2 + 2(-1.54)(0.92)(8.62) + (0.34)(8.62)^2$$

$$= 6.94$$

Thus, a 95% confidence interval for X_o is $35.86 \pm 1.96\sqrt{6.94}$ or 31 to 41 inches. The upper limit is meaningless since the lowest recorded total is 35.69 in 1885.

Using the fitted normal distribution, $z_o = -3.09$ such that

$$Pr(z<-3.09) = 0.001$$

Thus $\chi_o = 81.22 - 3.09(18.55) = 23.90$ inches with confidence limits $23.90 \pm 1.96 \times 18.55 = -12.46$ and 60.26 inches. Here, the lower limit is absurb whereas the upper limit is meaningless. Also, the confidence limits for this estimate are very much wider thanthat obtained from the fitted gamma distribution. It should be noted also here that although the upper limits of the maximum probable estimates from both curves tend to converge to a certain finite value, the lower limits of the minimum probable estimates do not show any such tendency. This is due to the fact that whereas the gamma curve terminates at the zero point of the ordinate, the normal curve extends further to -oo.

7. Conclusion

Even for long period records, say 100 years, the skewness of the rainfall frequency distribution persists. Positive skewness is a characteristic of rainfall data, and this is borne out statistically by comparing the fitted symmetrical normal distribution with the positively skewed gamma distribution. Though both curves fit the data well, the gamma distribution appears to be somewhat a better fit than the normal distribution. It should be noted that the Manila annual rainfall data are only slightly skewed.

It is seen that the gamma distribution gives better estimates of extreme probable values than the normal distribution. The confidence limits of the estimates from the gamma distribution are very much narrower than that obtained from the normal distribution. In other words, though both distributions are unbounded to the left, the gamma distribution is more efficient asymptotically than the normal distribution.

APPENDIX A

MANILA ANNUAL RAINFALL TOTALS[1] 1865 - 1968
(In Inches)

	0	1	2	3	4	5	6	7	8	9	TOTALS
1860						76.48	80.99	117.28	76.53	100.35	451.63
1870	78.85	63.94	77.86	67.47	47.52	66.42	78.36	99.41	58.25	67.99	706.07
1880	94.75	83.56	90.00	88.50	74.87	35.69	63.06	89.18	71.15	68.21	758.97
1890	82.48	101.72	50.50	56.75	66.08	82.08	70.29	51.32	87.67	110.01	758.90
1900	83.69	72.90	67.59	40.57	84.13	71.85	89.39	72.47	97.68	72.31	752.58
1910	75.89	67.62	76.15	75.64	95.87	75.39	64.66	91.71	85.01	154.35	862.29
1920	86.65	103.19	73.95	134.96	100.01	99.55	73.85	88.83	67.56	91.34	919.89
1930	72.79	116.19	95.41	75.68	106.01	112.67	90.84	119.17	75.78	86.31	950.85
1940	88.76						75.15	93.23	74.80	57.32	389.26
1950	67.02	72.97	111.28	95.12	71.64	50.62	86.97	61.56	95.41	62.70	775.29
1960	100.33	89.77	89.92	74.06	85.33	73.43	92.95	78.33	62.56		746.68
TOTALS	831.21	771.86	732.66	708.75	731.46	744.18	866.51	962.49	852.40	870.89	8072.41

Source: Climatological Division, Philippine Weather Bureau.
[1] Original data with missing observations.

BIBLIOGRAPHICAL REFERENCES

1. Brooks, C. and Carruthers, N., *Handbook of Statistical Methods in Meteorology,* Her Majesty's Stationery Office, London, 1953.
2. Chapman, D., *Estimating the Parameters of a Truncated Gamma Distribution,* Annals of Mathematical Statistics, Vol. 27, 1956, pp. 498-506.
3. Deutsch, R., *Estimation Theory,* Prentice-Hall Inc., Englewood Cliffs, N.J., 1965.
4. Hogg, R. and Craig, A. *Introduction to Mathematical Statistics,* The Macmillan Company, New York, 1959.
5. Gupta, S. and Groll, P.,*Gamma Distribution in Acceptance Sampling Based on Life Tests,* Journal of the American Statistical Association, Vol. 56, No. 296, 1961, pp. 942-970.
6. Kendall, M., *The Advanced Theory of Statistics,* Vols. I, II, Charles Griffin & Co., Ltd., London, 1946.
7. Mood, A. and Graybill, F., *Introduction to the Theory of Statistics,* McGraw-Hill Book Co., Inc., New York, 1963.
8. Moran, A., *The Statistical Treatment of Flood Flows,* Transactions, American Geophysical Union, Vol. 38, No. 4, 1957, pp. 519-523.
9. Panofsky, H. and Brier, G., *Some Applications of Statistics to Meteorology,* The Pennsylvania State University, 1968.
10. Wilkinson, G., *Estimation of Missing Values for the Analysis of Incomplete Data,* Biometrics, Vol. 14, No. 2, 1958, pp. 257-286.

LIST OF STATISTICAL TABLES

1. Pairman, E., *Tables of the Digamma and Trigamma Functions,* Tracks for Computers No. 1.
2. Pearson, K.,*Tables of the Incomplete Gamma Function,* Cambridge University Press, 1951.
3. Thompson, A., *Tables of the Coefficients of Everett's Central Difference Interpolation Formula,* Tracks for Computers No. 5.
4. *Tables of the Exponential Function* e^x, U.S. Department of Commerce, National Bureau of Standards Applied Mathematics Series No. 14.

The Philippine Statistician
Vol. XIX, Nos. 1 and 2. 1970

THE USE OF ALPHABETICALLY LISTED FRAME IN STRATIFIED SAMPLING[1]

by

Generoso G. de Guzman[2]

1. Introduction

A major limitation in the use of probability sampling design is the absence of a suitable frame. Quite often alphabetically listed names of individuals, places, towns and barrios, commercial establishments, members in various professional associations and honor societies are readily available. However on account of the non-random arrangement of the units in these alphabetically enumerated lists, they cannot just be used with a high degree of objectivity in probability sampling design.

It is a common observation in this country that in certain regions (or any suitable administrative or political subdivision) there is a dominance of family names beginning with only a few of the letters of the English or Pilipino alphabets.

This dominance of a few family names is still a common phenomenon in rural areas on account of the tendency of families of the same clan to reside close to each other.

Another characteristic of these alphabetically listed names is that the names may be classified or grouped according to some criteria. For example in the case of membership in professional associations and societies, the members, etc., and in

[1] This paper is part of the thesis of the same title submitted by the author in partial fulfillment of the requirements for the degree of M.A. (Statistics), U.P., 1970.

[2] Instructor, Statistical Center, University of the Philippines.

the case of telephone directory, the commercial establishments may be classified according to products produced or services rendered.

These lists are feasible frames for a two-way (or multiple) classipication sampling scheme, hence this study.

2. Characteristics of the 36-letter alphabet

Previous studies have shown that some letters of the alphabet appear more often than others, as the first letter of surnames of people in an alphabetical list; that is, the letters are distributed in a non-random fashion.

Thus, if the relative frequency distribution of the 26-letters in the English alphabet can be constructed, this distribution may be used in determining the probability of selecting a particular letter.

The relative frequency distribution of these 26-letters, therefore, may serve as a frame in as much as it possesses properties of a good frame.

3. Material used in constructing the frame

The material used in constructing the frame is the 1968 Philippine Long Distance Telephone Company Directory. Schools, business establishments, and the like are excluded.

Duplication in the list is considered as existing if two or more similarly named individuals are listed as residing in the same residence. Otherwise, they are considered different persons if they do not have the same address in the list.

A total of 38,815 names have been counted and the distribution of the first letter of the surnames of these people is prepared as shown in Table 1.

Table 1. RELATIVE FREQUENCY DISTRIBUTION OF THE FIRST LETTERS OF SURNAMES (PLDT Directory)

Letters	Relative Frequency	Letters	Relative Frequency
A	.073566	N	.019397
B	.053962	O	.022931
C	.116436	P	.052002
D	.036436	Q	.006087
E	.024015	R	.058373
F	.030257	T	.059559
G	.066730	S	.090977
H	.018727	U	.008873
I	.009673	V	.037944
J	.020223	W	.004617
K	.012768	X	.000026
L	.075010	Y	.016766
M	.077951	Z	.006655

4. The construction of the frame

After computing the relative frequency distribution of the 26 letters of the English alphabet among initial letters of surnames in the PLDT Directory. This is then arranged into groups of letters representing the first letters of the surnames of individuals in such a way that each group of letters includes about the same proportion of all names of the large population. That is, for the i^{th} letter in alphabetical array.

$$\Sigma p_{i1} = \Sigma p_{i2} = \ldots = \Sigma p_{ij} = \ldots = \Sigma p_{in}$$

and

$$\Sigma p_{i1} + \Sigma p_{i2} + \ldots + \Sigma p_{ij} + \ldots + \Sigma p_{in} = 1.00$$

where $j = 1, 2, \ldots, n$;

p_{ij} is the probability of the occurence of the i^{th} letter which is grouped in the j^{th} row;

Σp_{ij} is the probability of the occurence of all the letters which are grouped in the j^{th} row;

and n is the number of rows in the array.

It can be seen from Table 1 that for every 100 individuals in an alphabetical list, about 11 or 12 have their surnames beginning with a C. It is clear, therefore, that if one does not split this proportion of letter C, he can group the 26 letters into at most a 9-row alphabetic array. In this study however, the array is extended up to a 12-row alphabetic array by splitting C into two groups for a 10-row array and C and S into two groups each for a 12-row array. The splitting of these two letters are made possible by considering the distribution of the first two-letters of the surnames in the list.

Table 2 shows the 4-to 12-row alphabetic array prepared from the relative frequency distribution of the English alphabet among surnames in the PLDTCo Directory.

Table 2. THE 4- TO 12-ROW ALPHABETIC ARRAY

4-Row Alphabetic Array	ΣPij
C D O R X Y	.251005
E J K S T V W	.250103
H G I M N P Q	.250567
A B P L Z U	.248323
6-Row Alphabetic Array	
C J O Z	.166245
K L N T	.166734
D M P X	.166452
E F S W Y	.166632
B I Q R V	.166039
A G H U	.167896
8-Row Alphabetic Array	
A F J	.124046
B N P	.125361
C U X	.125335
D M Q W	.125128
E I S	.124665
G H O Y	.125154
K L V	.125722
R T Z	.124597
9-Row Alphabetic Array	
A F Z A	.110504
B D J	.110 58
C	.116436
E M U	.110839
G Q V	.110761
H I T W Y	.109342
K L O	.110709
N S	.110374
P R	.110375

10-Row Alphabetic Array

A N Z	.099618
B Cj-Cz	.100856
Ca-Ci F	.099799
D Q R	.100933
E T Y	.100340
G J K	.099721
L H W	.098354
M O	.100882
P V I	.099619
S U X	.099876

12-Row Alphabetic Array

A I	.083239
B O Z	.083548
Ca-Ci K X	.082336
Cj-Cz D	.083367
E T	.083574
F P	.082259
G Y	.083496
H Q R	.083755
J N Se-Sz	.084038
L U	.083883
M W	.082568
Sa-Sd V	.084503

After getting these t-row alphabetic arrays, each array is assigned to a particular two-way sampling design by considering two cases.

4.1 Case 1. The (r x c) cells = t-row alphabetic array

For a 2 x 2 up to a 3 x 4 or 6 x 2 tables, the following is assumed:

$$r \times c = t,$$

that is, a particular alphabetic array is assigned to a r x c table such that the number of groups of letters in the said array is equal to the number of cells of the table. Thus, a 4-row alphabetic array is used as the frame for a 2 x 2 table, a 3 x 2 table makes use of a 6-row array, etc.

Under this method, each group of letters in a given t-row alphabetic array is assigned a number. Random numbers from 1 to t are drawn without replacement and the groups of letters corresponding to the random numbers drawn are matched with the cell of the r x c table in a serpentine way starting from cell c_{11}. The serpentine way of allocation is illustrated in section 5.2.

4.2 Case 2. r x c > 12

Because of the limitations inherent in the relative frequency distribution of the 26-letter alphabet, only up to a 12-row alphabetic array has been prepared. However this does not invalidate the study in cases where the total number of cells exceed 12, that is, rc > 12.

A frame can still be prepared using the same t-row alphabetic array established here.

A method for the construction of a frame is proposed in this section.

The choice of a particular t-row array that could be used for a frame is in general influenced by the overall sampling fraction to be used in the survey. The working equation is

$$1) \qquad f = \frac{n}{N} \leq p_{ij}$$

that is, the overall fraction, f, should always be less than or equal to the probability of occurence of all the letters which are grouped inthe j^{th} cell. This is so, because Σp_{ij} is actually the proportion of the population in the ij^{th} stratum that is included in the frame. As the sample is drawn from the available names in the frame it is obvious that the sampling fraction should not exceed this proportion Σp_{ij}.

After eliminating the k-row arrays that do not satisfy Eq. 1, a particular array is selected from the remaining (t-k)-row arrays with eual probability. The groups of letters in the chosen array are then numbered from 1 to s where s is the number of groups of the letters of the particular array so chosen. Random numbers from 1 to s are drawn with replacement and the groups of letters corresponding to the numbers drawn are matched with the rc cells in a serpentine way similar to case 1.

4.3. Revision of the frame when $f > \Sigma p_{ij}$

The frame constructed in this study has the peculiar characteristic that samples are drawn not from the whole population but rather from a portion only of the population. Because of this limitation, the sampling fraction, f, is obviously restricted by the proportion of the population included in the frame. If one will use a sampling fraction larger than the proportion Σp_{ij}, he has to revise the frame to suit his need.

Provided $f \leq 2\ \Sigma p_{ij}$, the fram may be revised by randomizing first the rows and then the columns of the original frame and then combining cell by cell the randomized two-way frame with the original one. This method will double, more or less, the original Σp_{ij}, thus increasing the proportion of the population covered by the frame.

5. Application

The alphabetically listed frame is applied to a study conducted to determine the average annual income of the faculty members teaching collegiate course in the educational institutions in the Greater Manila Area. Greater Manila area includes places such as Manila, Quezon City, Pasay City, Caloocan City, Malabon, Navotas, Makati, Mandaluyong, Marikina, San Juan, Parañaque and Las Piñas.

A two-stage, two-way stratified sampling design is used. The educational institutions being considered as the primary sampling units and the faculty members as the secondary sampling units. In this study a complete list of educational institutions is available but the complete list of all faculty members teaching in these institutions is assumed not readily available, hence the necessity of the alphabetically listed frame is imperative to get samples from these secondary sampling units.

5.1. Stratification of the population

To carry out the study, the population of educational institutions is stratified according to two criteria — according to size of enrolment and according to location. The size of enrolment is divided to four strata namely:

A. less than 1000
B. 1000 to 4999
C. 5000 to 10000
D. over 10000

and the location to three strata:

I. Suburbs
II. Manila, First and Second Districts
III. Manila, Third and Fourth Districts

Table 3. DISTRIBUTION OF EDUCATIONAL INSTITUTIONS IN THE GREATER MANILA AREA, STRATIFIED BY LOCATION AND BY SIZE OF ENROLMENT.

		LOCATION			
		Suburbs	1st and 2nd Districts	3rd and 4th Districts	Total
ENROLMENT	Less than 1000	23	9	13	45
	1000-4999	2	4	6	12
	5000-10000	1	4	1	10
	over 10000	1	4	1	6
	Total	27	21	25	73

5.2. The 4 x 3 frame used in this study

As shown in Table 3, the population of educational institutions has been stratified into 12 strata. This calls for the use of a 4 x 3 frame. To construct this frame the 12-row alphabetic array is used in as much as there are 12 groups of letters in this array and there are 12 cells to be filled up. Each group of letters is numbered from 1 to 12. Random numbers from 01 to 12 are drawn without replacement and the group of letters corresponding to the first random number drawn is assigned to the cell of the first row and first colum (c_{11}). The remaining groups of letters are matched with the rest of the cells in a serpentine way in the order of the random numbers as shown in Figure 1.

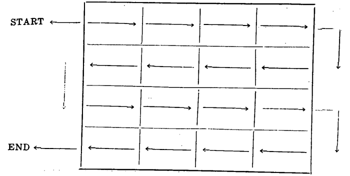


Figure 1. THE SERPENTINE ..AY OF ALLOCATION OF THE 12-ROW ALPHABETIC ARRAY TO THE 4 x 3 TABLE.

M W .082568	H Q R .083755	J N Se-Sz .084038
Sa-Sd V .084503	A I .083239	Cj-Cz D .083496
E T .083574	F P .082259	G Y .083496
Ca-Ci K X .082336	L U .083883	B O Z .083548

Figure 2. THE FRAME FOR A 4 x 4 STRATA[1]

Note that on the average, this 4 x 3 frame (Fig. 2) gives p_{ij}= .0833 or that only 8.33% of the population in the ij[th] stratum has been included in the frame. This means that the frame may be used only for a maximum 8.33% sampling fraction per stratum.

To obtain samples for a 10% sampling fraction, the frame is revised following the procedure in section 4.3. The resulting revised frame is shown in Figure 3.

M Cj-Cz W D .165935	H Q R Sa-Sd V .168258	J N Se-Sz A I .167277
Sa-Sd V B O Z .168051	A I Ca-Ci K X .165575	Cj-Cz D L U .167250
E T J N Se-Sz .167612	F P M W .164827	G Y H Q R .163030
Ga-Gi K X G Y .165832	L U E T .167457	B O Z F P .165807

Figure 3. REVISED FRAME FOR A 4 x 3 STRATA

[1] The figure below each cell is equal to pij which represents the proportion of the population in the ijth stratum that is included in the frame.

3.3. *Allocation of the sample psu and ssu*

Using the alphabetically listed frame developed here, the number of sample psu is determined by the sampling fraction and the proportion of the stratum population included in the frame.

To illustrate this, using Figure 2 and a sampling fraction of 1%, it is desired to estimate the number of institutions to be included as sample for the suburbs stratum where the enrolment is less than 1000 students.

For proportional allocation,

$$f = f_{ij}$$

but $$f_{ij} = f_{1ij} f_{2ij}$$

here $f_{ij} = .01$ and $f_{2ij} = p_{11} = .082568$

then $f_{1ij} = f_{ij}/f_{2ij}$

$= .121112$

Now from Table 3 there are 23 schools under this stratum, therefore the total number of institutions to be included as sample psu is equal to (.121113)(23) = 2.786, and rounding to the next highest integer we have 3 schools as sample psu. Correcting the ssu sampling fraction,

$f_{2ij} = f_{ij}/m_{ij}/M_{ij})$

$= (.01)/(3/23)$

$= .76667$

This means that for every 100 faculty members in this stratum whose surnames begin with letters M and W, about 92 or 93 will be included as samples.

Table 4 shows how the psu and ssu are to be alloted for 1, 5, and 10% sampling fractions.

Table 4. DISTRIBUTION OF SIMPLE EDUCATIONAL INSTITUTIONS (PSU) AND FACULTY MEMBERS (SSU)[2]

I. 1% Sampling Fraction

		LOCATION		
		I	II	III
ENROLMENT	A	3 92.85%	2 53.73%	2 77.35%
	B	1 23.67%	1 48.05%	1 71.97%
	C	1 11.96%	1 48.63%	1 59.88%
	D	1 12.14%	1 47.68%	1 11.97%

[2] The upper entries are the number of educational institutions (psu) while the lower entries are the percentage of faculty members (ssu) in the frame used as samples.

II. 5% Sampling Fraction

		LOCATION		
		I	II	III
ENROLLMENT	A	14 99.48%	6 89.55%	8 96.68%
	B	2 59.17%	3 80.09%	4 89.96%
	C	1 59.83%	3 81.04%	3 99.80%
	D	1 60.73%	3 79.48%	1 59.85%

III. 10% Sampling Fraction

		LOCATION		
		I	II	III
ENROLLMENT	A	14 99.01%	6 89.15%	8 97.14%
	B	2 59.51%	3 80.53%	4 89.69%
	C	1 59.66%	3 80.89%	3 99.77%
	D	1 60.30%	3 79.62%	1 60.31%

6. The Analysis

After all sample psu and ssu have been determined, the average annual income of the faculty members teaching in educational institutions in the Greater Manila Area are estimated.

The coefficient of variation is computed to compare the variability among the different sampling fractions and also to determine the relative efficiency of one type of frame with the other.

To establish the behavior of these estimates — the mean variance, and coefficient of variation — their sampling distributions are prepared by replicating the sampling scheme 20 times for each of the three sampling fractions.

It has been expected that the coefficient of variation will decrease with increasing sampling fraction. However, it will be noticed in Table 5 that contrary to what is expected, the 1% sampling fraction gave the smallest coefficient of variation. A closer look at Table 4 will reveal that except for 3 strata only one sample psu per stratum was drawn for the 1% sampling fraction while more than one psu were drawn per stratum for the 5 and 10% sampling fractions. Hansen, Hurwitz, and Madow (7) pointed out that the variance will be smaller if only one sample psu stratum is drawn as compared to a design where there are two or more sample psu's per stratum.

Nevertheless, comparing the 5 and 10% sampling fractions, it will be observed that there is a considerable reduction in the coefficient of variation.

Table 5. THE MEANS OF THE SAMPLING DISTRIBUTIONS OF THE MEAN, STANDARD DEVIATION, AND COEFFICIENT OF VARIATION OF THE ANNUAL INCOME.

Sampling Fraction (%)	1	5	10
Average Annual Income (₱)	7125	7292	7327
Standard Deviation	274.02	337.48	326.37
Coefficient of Variation (%)	3.85	4.64	4.45
Variance of the C.V.	.008933	009362	003192

The efficiency of the two-way stratification is now compared with the one-way stratification. As shown in Table 6, it is evident that the two way stratified sample design gives a considerable reduction in the variability of the estimate most especially when the location is used as the criterion for stratification.

Table 6. THE RELATIVE EFFICIENCY OF THE T..O-..AY STRATIFICATION TO THE ONE-..AY STRATIFICATION.

Sampling Fraction (%)	1	5	10
CV (Enrolment) / CV (Two-Way)	3.6052	1.6379	1.6652
CV (Location) / CV (Two-Way)	7.3980	3.3362	3.4562

7. Conclusion

The frame prepared in this study utilized the distribution of the first letter of the surnames from an alphabetically list — the PLDT Directory. One special feature of this frame is that it does not depend on the identity of the sampling unit; it only utilized the fact that the distribution of the first letter of the surnames in an alphabetically enumerated list is stable for large population. Because of this, the frame may find a universal application in surveys when ever the sampling units are individuals.

Although the frame was prepared for a two-way stratified design, it may also be used for a one-way stratified design by considering only the row or the column, as the case may be.

It was also shown in this study that two-way stratification causes a high reduction in the variability of the estimate compared with a one-day stratification.

It is hoped that the results of this study will help people engaged in survey work who are always confronted with frame construction problems. However, in as much as the study was concentrated on using the frame for a two-way stratified sampling design using proportional allocation of the samples, perhaps other researchers may be interested on studying the feasibility of the frame for other sampling designs, for example, systematic sampling design.

BIBLIOGRAPHY

1. Bryant, E.C., Hartley, H.O., and Jensen, R.L. "Design and Estimation in Two-Way Stratification," *Journal of the American Statistical Association*, 55 (1960), pp. 105-124.
2. Buechley, R.W., "A Reproducible Method of Counting Persons of Spanish Surnames," *Journal of the Americal Statistical Asseciation*, 56 (1961), pp. 88-97.
3. Cochran, W.G. *Sampling Techniques*. John ileyWiley & Sons, Inc. New York. 1963.
4. Frankel, A. and Wright P. "The Alphabetical Array Sampling Technique," *American Statistical Association Proceeding of the Social Section* (1964), pp. 20-24.
5. Frankel L. and Stock, J.S. "On the Sample Survey of Unemployment," *Journal of the American Statistical Association*, 45 (1950), pp. 350-372.
6. Gutierrez, J.S. "Some Theoretical Considerations on the Uses of Coefficients of Variation in Sample Surveys" *The Philippine Statistician*. xiv (1965) pp. 222-243.
7. Henson, M.H., Hurwitz, ...N., and Madow, W.G., *Sample Survey Methods and Theory*. John Wiley & Sons, New York.
8. Kish, L. *Survey Sampling*. John Wiley & Sons, Inc. 1965.
9. Yamane, T. *Elementary Sampling Theory*. Prentice-Hall, Inc., New Jersey. 1967.
10. Yates F. *Sampling Methods for Census Surveys* (3rd ed.) Hafner Publishing Co., Inc. New York. 1960.

The Philippine Statistician
Vol. XIX, Nos. 1 and 2, 1970

1970 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 Elecctric Company

Menzi and Company

National Life Insurance Company of the Philippines

National Investment and Development Corportaion

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 member:

ANDRES ROXAS

Republic of the Philippines
Department of Public Works and Communication
BUREAU OF POSTS
Manila

SWORN STATEMENT
(Required by Act 2580)

The undersigned, CRISTINA P. PAREL, editor of THE PHILIPPINE STATISTICIAN, published in English at The Statistical Center, 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: Cristina P. Parel	Rizal Hall, Padre Faura, Mla.
Owner: The Philippine Statistical Assn.	Rizal Hall, Padre Faura, Mla.
Publisher: The Philippine Statistical Assn.	Rizal Hall, Padre Faura, Mla.
Printer: U.S.T. Press	España, Manila
Office of Publication: The Statistical Center	Rizal Hall, Padre Faura, Mla.

In case of publication other than daily, total number of copies printed and circulated of the last issue dated Sept.-Dec. 1969.

1. Sent to paid subscribers	475
2. Sent to others than paid subscribers	50
Total	525

CRISTINA P. PAREL
Editor

SUBSCRIBED AND SWORN to before me this 27th day of May, 1970 at Makati, Rizal, the affiant exhibiting her Residence Certificate No. A-4293792 issued at Quezon City on January 13, 1970.

Doc. No. 205; Page No. 42;
Book No. II; Series of 1970

Notary Public

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

ACKNOWLEDGMENT

The officers of the Philippine Statistical Association, Inc. and the Editorial Board of *The Philippine Statistician* (Vol. XIX) acknowledge the financial grant etended by the National Science Development Board for the publication of this issue.

The Philippine Statistician

SEPT. - DEC., 1970 • VOL. XIX • NOS. 3-4

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 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. XIX, Nos. 3 and 4, 1970

"THE USE OF STATISTICS FOR SETTING INDUSTRIAL PRIORITIES" *

By

RODOLFO VILLARICA**

The paper I am honored to deliver today is on "The use of Statistics for Setting Industrial Priorities" in the Board of Investments. Without questioning the merits of the topic, since it takes for granted two givens, namely, that the Board of Investments (BOI) sets priorities and that the BOI uses statistical methods of analysis, both functions of which the BOI performs, let me describe how the BOI fulfills its mission of implementing the law, R.A. 5186 by reviewing its significance and investment policies for a moment.

R.A. 5186, otherwise known as the Investment Incentives Act, is a law that prescribes the incentives and guarantees being offered by the Philippine Government to both domestic and foreign enterprises in preferred areas of Investment in the Philippines. The Act also grants incentives to investors in enterprises registered with the BOI. Additional inducements are provided for export-oriented enterprises although these would be transferred to the Export Incentives Bill, approved by Congress in their second special session and now awaiting the signature of the President for final enactment.

R.A. 5186 is a milestone in economic legislation because it is the first to spell out a comprehensive guideline for channelling both Filipino and foreign resources into preferred investment areas. It also specifies the various kinds of incentives to be granted to investments and investors. Although

* Read at the Annual Conference of the Philippine Statistical Association, September 18, 1970.

** Director, Chemical Department, Board of Investments.

it seeks to insure the control of the country's patrimony by Filipinos, as is the practice in most developing countries, it also provides such adequate incentives and safeguards as are essential for profitable foreign investment in our country.

Interested investors and enterpreneurs frequently ask us what our investment policy is. In general, the twin policies are: (1) to encourage Filipino and foreign investments in projects to develop agriculture, mining and manufacturing industries which increase the national income with the least cost, increase exports, provide more opportunities for employment, raise the living standards of the masses and provide for an equitable distribution of wealth and (2) to welcome and encourage foreign capital to establish pioneer enterprises that are capital intensive and would utilize substantial amounts of local raw materials, in joint venture with Filipino capital, whenever possible. From these twin policies emanate the criteria that determine the priority or non-priority of industrial projects. The Act requires the Board of Investments to set forth the different activities in the agricultural, mining and manufacturing sectors to be accorded preferred pioneer or preferred non-pioneer status as fields of investments for purposes of enjoying the incentives and guarantees offered under the Acts. This annual plan, called the Investment Priorities Plan, is the result of the analysis, synthesis and projection of data collected by the BOI from private and public sources.

Our government, as a matter of policy, adheres to a relatively free enterprise economy. The private businessmen are actually the person who make decisions on whether to invest or not in a particular enterprise. Two emergent factors, therefore, determine the planning process at BOI: first, the institutional constraints within which it operates and second, its ability to extend its planning process beyond the preparation of the investment priorities plan. The first factor forces the BOI to act and react within the framework of the real world where bureaucracies, obsolete laws, outdated regulations, irrelevant institutions cannot be disregarded nonchalantly and where incomplete figures and uncertain information more often than not limit choices to solutions that are not optimum. The second factor endows more meaning and responsibility on the BOI to plan not only the priority areas but also the implementation of the program through assistance in all facets of realization, availment of incentives, and supervision over the registered enterprises throughout the different phases of the life of their projects.

Having thus briefly described the objectives and policies of the BOI governing the planning responsibilities in the context of reality and implementation, we are now ready to describe the planning process itself, the preparation of the Investment Priorities Plan.

In the preparation of the first Investment Priorities Plan, a formula was evolved which took into account the following factors: (1) Social Rate of Return; (2) Foreign Exhange earnings and/or savings; (3) Labor Intensiveness; (4) Potential Forward and Backward Linkages; and (5) Insufficiency Gap.

The *Social Rate of Return* measured the net or real gain to the economy on a given investment. This method started with the estimation of the commercial (or private) profitability of a project, after which adjustments were made to reflect the real profit to the economy by netting out the taxes and duties for revenues and costs and by adjusting for opportunity costs of materials and manpower.

The *Foreign Exchange Earnings and/or Savings* was a factor which gauged the effect of a project on the balance of payment by computing the project's revenue, priced at CIF values if the product was an import substitution or FOB value if export-oriented, and then deducting the cost of imported raw materials, and amortization on imported equipment. As an index, the amount of foreign exchange earned or saved per peso invested was used.

Labor intensiveness measured the employment generative potential of projects. The index used was the number of jobs generated per peso of investment.

Potential Forward and Backward Linkage measured the necessity of the project viewed from the need to integrate manufacturing and thereby to create self-sufficiency.

The Insufficiency Gap represented the volume of demand which cannot be met by existing production facilities.

The five criteria, assessed individually for each project, were combined by weighted proportions into a formula for ranking projects, called the Index of Productive Output.

The weight assigned to each parameter was: 30 points for the Social Rate of Return, 30 points for Foreign Exhange Earnings and/or Savings, 15 points for Labor Intensiveness, 15 points for Linkage Effects and 10 points for the Insufficiency Gap. The Projects were each normalized individually within a given criterion and the normalized numbers were plugged into the Index of Productive Output formula. Thus a ranking was achieved among the sixty (60) odd projects recommended for priority implementation in the first plan.

Soon after the first Investment Priorities Plan was approved by the President, copies of the Plan were distributed among economists for their comments regarding the application of the formula which they had in principle previously approved. Constructive criticism, verbal and written led the Board to think twice about the formula it had adopted. It was pointed out, for instance, that implicit in such formula is a set of marginal rates of substitution between criteria of choice, which, if followed, purports to allocate investment funds optimally. But since the weights of the criteria were somewhat arbitrarily made and assigned, and the interrelationships between the criteria were not fully known, it was possible that the marginal sets of substitution between such criteria were not being optimized. In other words, it would be extremely difficult to determine what would be the trade-off between criteria if the interrelationship among them were not fully known. For instance, what would happen to our objective of pursuing a very favorable effect on the balance of payments through a project's foreign exchange savings/or earnings effect if were to overemphasize labor intensiveness to an extreme degree?

In the preparation of the second Investment Priorities Plan, therefore, the data on the five criteria were longer plugged into the Index of Productive Output formula. Emphasis was centered on the social rate of return criterion which was modified and refined to suit the needs of the times. It was also felt that the foreign exchange implications of a project were to be deemed important since indications prognosticated the weakening position of the peso vis-a-vis the dollar. Towards the end of the preparations of the second IPP, there came a suggestion from two quarters, NEC and PES, to utilize the input/output analysis as a criterion for the ranking of projects. This would involve a determination of the economy's target performances and a definition of those economic activities which had the greatest purpose of accelerated economic development consonant with nationalism.

This approach would have entailed two parts: first a projection of all types of final demand in appropriate detail; and second, the input-output analysis.

Th first part would necessitate analysis and processing of the project data with regard to the target performance of the GNP and of various sectors and subsectors. The GNP figures from the NEC would be the starting point and the aggregate data broken down into sectors. The final demands of each of these sectors would be projected on the accelerated basis. Then by means of correlation, the degree of association of the performance of the economy, with sectors classified according to the PSIC could be estimated, and from regression analysis, the values of the expected performance of the various sectors would have to be provided. The various projected demands would be checked against minimum economic size plants to find out whether and at what time during the projection period there would be a probability for additional expansion and/or new plant capacities to be installed.

With the completion of first part, the Input-Output analysis would then be the basis or criterion for the inclusion or noninclusion of a project in the IPP. The matrix of input coefficients and other parameters of input-output statistics would be taken from the NEC input-output table unless knowledge of change in techniques of production would justify a change in input coefficients. Using the NEC "inverse matrix" and the projections of final demand, one could then work back to the impact of these projected final demand on the various industries: the ultimate impact of demand for specific commodity on all industries is provided in the appropriate column of the matrix; similarly, the effects of final demand upon sales of a specific industry are derived by reading across the appropriate row. A projection of expansion of demand on individual industries and an industry-by-industry discussion of the main consideration affecting development could be made and used as the basis for selecting the preferred areas to be included in the IPP.

Mr. Ramon Katigbak, Jr. then detailed from PES to the BCI, worked on this input-output approach. He soon realized that this involved an enormous amount of work and that within the time constraint of IPP preparation, such a job could not be completed. But there were other more real problems in that this analysis which involved optimization within a given set of projects being realized and implementeed within specific and stated times could not work within the context of BOI planning since there is no assurance that specific projects will be operational by such and such a year nor even, that any given project would at all have taken.

Perhaps, this paper should have been entitled, *In Search of a Statistical Model* for *Industrial Planning*. Indeed the very dynamism of the process of economic development seems to be beyond the static precision of an Input-Output, or a social marginal productivity formula to measure. And yet we would like to believe that the elasticities of statistical models are positive enough to be able to create a still more precise measurement of priority regarding industrial investments that would more clearly define truly impact areas where socio-economic benefits could more speedily flow to a wider segment of our society:.The challenge stands.

The Philippine Statistician
Vol. XIX, Nos. 3 and 4, 1970

THE USE OF STATISTICS FOR ECONOMIC DEVELOPMENT PLANNING*

By
ARTURO Q. TRINIDAD **

A. *Introduction*

This paper is being presented to discuss another paper entitled "Planning and Statistics" *** written by Mr. Bernardino G. Bantegui, for the SEMINAR ON STATISTICS AND PLANNING held in Moscow, USRR on September 22 — to October 12, 1969.

It was the feeling of Mr. Bantegui that it is "old hat" to present a paper on this subject anew. However, we believe that the work of Mr. Bantegui is worthwhile discussing since it reported lengthily on the procedures adopted by the National Economic Council in assessing the feasibility of the Board of Investment's Second Investment Priorities Plan (IPP) but at the same time, it fails to report on the results of their evaluation.

One could rationalize on the probable reason for this since the paper was presented before a foreign audience. It would not be good to criticize one's own government agency if the findings had been unfavorable. There remains however, an impression of omission which we hope can correct in our discussion this morning. The results, of course, are of particular interest to us in the financing field who should go by the priorities set forth in the Investment Priorities Plan.

* Read at the Annual Conference of the Philippine Statistical Association, September 18, 1970.
** Actg. Vice President, National Investment Development Corporation.
*** Published in the Statistical Reporter Vol XIII (3), 1969, OSCAS, National Economic Council.

The underlying theme that runs throughout the paper of Mr. Bantegui is that statistics are indispensable to planning for economic and social development. That statistics are necessary for planning is hard to deny. This, however, should not obscure the fact that what matters in development planning is not mainly statistical but mainly policy. It is still possible to have a good Development Plan without using much statistic by concenrating on policies which will stimulate on upward movement of the economy. The Singapore experience is an example. It is also possible to write a Development Plan which is mathematically consistent, but which will nevertheless achieve nothing, because policies are inconsistent.

One of the fatal factors that undermines economic development planning is the promulgation of contradictory and/or incompatible policy actions. For example, a minimum wage policy is adopted which negates the desired effects of a price policy. A tariff policy is enunciated to cripple developing industries and an export tax policy to neutralize export incentives. These clashes of policies create repercussions that cannot be measured by special purpose models which through their inherent limitations neglect indirect side effects. It does not mean though that the results of these contradictions cannot be measured without the use of complicated scientific aids.

The paper of Mr. Bantegui mentioned in passing linear programming among the approaches to development planning. Given several targets of output, import and export for each industry which are all consistent and apparently reasonable, the mathematical programming approach aims at optimum allocation of funds in the light of return expectations and scarce financial resources. The determination of the social rate of return is a preliminary step in adopting this procedure which the paper of Mr. Bantegui ably presented. At present using industry statistics which are available from public and private agencies, this approach may be the simplest and the most practical. The results, however, needs tempering consistent with avowed national policies. An example toucning on a controversial sector — the Philippine sugar industry — will elaborate this point.

Regarding the use of the input-output table to evaluate the BOI's Priorities Plan, a more lengthy comment may be given.

B. *Planning and Prediction*

There is a distinction between the input-output table, regarded as a statistical description of certain features of the economy, and input-output analysis, regarded as a means of predicting the consequences of changes in underlying factors. The table is clearly an extremely useful and ingenious construction that can be regarded as a natural extension of national income accounts. As such, its value and usefulness are largely independent of the success of input-output analysis as a predictive technique. Any dispute about the table is only whether it is worth its cost, not whether it is worth having at all.

Viewed as a predictive device, as may be gleaned from Mr. Bantegui's paper, input-output analysis specifies a method of predicting the total output of a series of industries from the so-called "final-demand schedule", or "bill of goods." Its actual use to forecast total output for a future year involves, first, forecasting final demand, then predicting total output from this final demand; any error in the forecast can, in principle, be separated into (1) the error in the forecast of final demand; and (2) the error in the conditional prediction of total output, given final demand. In judging the analytical validity of input-output analysis, only the accuracy of the conditional prediction should be taken into account, for errors in forecasting final demand cannot be attributed to defects of input-output analysis.

The central feature of input-output analysis as a predictive device is that it proceeds to make predictions as if all coefficients of production were fixed, as if, in each defined industry, the amount of each input per unit of output were rigorously fixed, regardless of relative prices, levels of output, and so on. It is obvious that coefficients of production are not rigorously fixed, that all sorts of variations are possible and do occur. The lack of descriptive realism of fixed coefficients of production is not, however, the main issue against input-output analysis.

The crucial question is not whether coefficients of production are, in a descriptive sense, rigorously fixed but whether treating them as if they were constant involves neglecting factors that are considered "minor" disturbances but which are actually the major determining influence. In many instances there are real shifts from one method of production to another. These changes usually involve substitution of productive stocks and raw materials which cannot be corrected by the application of technical coefficients or indices. This is particularly true in the chemical industry where process changes will distort any estimate of input and output based on an old set of stock-flow coefficients. A more effective technique must obviously determine to what extent the variation is the output of a particular industry will reflect the changes in its internal structural characteristics or vice versa. Whenever these internal disturbances are known to be present but they are inaccessible to quantification then it might be better to make use of a simpler technique that will simply describe their effects through some kind of time trend or cyclical behavior. In such cases the use of the much understood methods of economic time series analysis may be more practical than the sophisticated input-output tables derived from insufficient information.

There is also the question of numerical accuracy of the figures presented in input-output table. We can also surmised that none of the figures used were obtained through actual or direct observation. Most likely, the data used were derived from the usual indirect estimating procdures applied on standard sets of primary data whenever they are available. Actually, there is nothing wrong with such indirect procedures provided they are blended with actual experience. It is doubtful whether the estimates can be corrected by interviewing businessmen or executives even if they are directly involved in the different industries concerned although it may be helpful. But there are factual information deeply burried in private business records which are not also immediately available to the executive operating at a higher level or the researcher conducting a study over a short period of time covering several industries. Whether we like it or not there has not been much, if not constant, communicattion between the practitioners immersed in actual business operations and the economists or statisticians interested in furhering sophisticated techniques in their own fields.

It seems we can justify the use of the input-output tables by differentiating between planning and prediction. In planning, we are interested in what producers must do to maintain a given level of future output while in prediction we are concerned with what producers will chooose to do in response to recent levels of demand. We may then proceed to believe that what producers must do is not closely related at all to what they will do. If this is the case then we have reasons to doubt the practical value of the input-output technique.

For planning, we also have to know the relation between final demand and output in each industry. Assuming the relationship is known there is still the problem of forecasting final demand. Hence, like most techniques, the input-output approach is excellent if final demand can be dictated rather than forecasted.

C. *Cost and Capital-Output Ratios*

The input-output tables are supposed to yield cost ratios indicating pesos of given inputs per peso of given output. These ratios are much more than a historical description of transactions since they are supposed to apply not only to the past period covered by the tables but also to future periods, possibly more than a decade, after appropriate adjustments. In general, this supposition can only be true if the cost ratios change slowly or if the changes can be detected and incorporated in the basic tables. As pointed out earlier, these conditions may be difficult to meet in practice.

A good economic plan must present a realistic estimate of investment expenditure especially in a capital-starved country like the Philippines. Investment expenditure is the product of capital-output ratio and required capacity. In the absence of statistics on a sectoral level capital-output ratio is estimated by using the stock concept which is the cost of all existing plant and machineries divided by the output for that year, all at constant prices. While this choice of an alternative (over the incremental concept which requires sectoral level information) appears insignificant, it results in grossly misleading estimates. Most of the industries which went through substantial increase in productive stocks during the control era preceding 1962 and the period of tight credit after the decontrol in that year have inflated assets due to several reasons not to mention the dubious ones, among which were:

1) The need for high dollar allocations which can be sold or used to generate precious excess dollars;

2) The need to inflate owner's equity in private firms to qualify for higher borrowing capacity;

3) The need for over-valuation to cover financing of local costs by foreign financiers.

This observation certainly does not lessen the efficacy of the input-output technique but rather does it prove its worth.

D. Ranking of Sectors

For industrial planning where the sources of financing are mostly government institutions, there is a need for priority ranking of sectors. The criterion mentioned by Mr. Bantegui in his paper, which is the Social Rate of Return (SRR) is widely accepted in theory and in practice in allocating scarce funds to the different sectors of the economy. Attempts have been made in the past and up to the present to apply the SRR to justify the financial support given to projects in different industries. Unfortunately, it has been observed that the SRR alone will not assure a decision contributing to a balance growth of the economy. It is also controversial when applied to certain industry sectors.

The formula presented by Mr. Bantegui for determining the social profitability of capital is simple enough yet comprehensive. It is in fact an improvement over the so-called national economic profitability criterion, also called the Social Rate of Return by Murray D. Bryce[1]. However, Bryce's criterion brings together in one measure all the major measurable economic costs and benefits of a project to a developing country, which includes labor intensity, a factor which is considered separately by Mr. Bantegui in his formula for determining social profitability of capital. It is not known how Mr. Bantegui would comparatively measure labor intensity and demand gap for the purpose of assigning points to various sectors of the economy.

Actually, the problem is not the derivation of a formula which is subject to different presentation. The question is in the results. It is easy to show, for example, that applying the social rate of return criterion will justify a continuing and increasing support of the sugar industry, yet, this may run counter to the thinking of independent policy-makers at present and even that of the International Monetary Fund. Using a different approach but following the principle behind the social rate of return criterion mentioned by Mr. Bantegui, earlier studies were made as to whether allocating funds for the sugar industry can be justified considering:

1. The national income effect;
2. The balance of payment effect;
3. The national economic profitability criterion; and
4. Backward and forward linkages.

The National Income Effect

The contribution to the national income by the sugar industry can be measured and compared with those of other in-

[1] Bryce, Murray D. "Industrial Development", New York, McGrawhil Book Co., 1960, pp 141-146

dustries by getting the percentage ratio of the net domestic value added by the industry to the total value of scarce resources utilized in production.[2] The expression can be written:

$$N_e = \frac{D_v}{R_s} \times 100$$

where: D_v, the net domestic value added by the sugar industry can be determined by summing up the peso values of the factors of production used, namely:

a) Land — Rent of land directly used for actual agricultural and milling activities;

b) Labor — Salaries, wages bonuses and fringe benefits received by employees and laborers employed;

c) Capital — Depreciation of productive assets used i.e. machineries, buildings, equipment and facilities and interest paid on funds utilized in actual operation; and

d) Entrepreneurship — Cost of profit and direct taxes paid by the sugar planter and milling company.

Alternatively, D_v can be calculated by deducting from net sales the following forward and backward linkage values:

a) Materials and Supplies — Cost of materials, supplies and other goods purchased from other producers and used in agricultural production, milling, administration and sales;

b) External Services — Cost of all services rendered to the sugar planter and milling company by other parties i.e., professional services, advertising, communications, transportation;

c) Indirect Taxes — Cost of all indirect taxes paid by the sugar planter and milling company.

R_s, the scarce resources used in sugar production include the following items:

a) Depreciation of fixed and replaceable assets — Cost

[2] Ruggles and Ruggles, "National Income Accounts and Income Analysis", McGraw-Hill Book Co., Inc., New York, 1956, pp. 54-59.

of depreciation and deplation of fixed local and imported assets;

b) Maintenance of fixed assets — Cost of labor, spare parts and supplies incurred in maintaining all fixed local and imported assets; and

c) Use of foreign exchange — Cost of foreign exchange used on materials and supplies directly imported, salaries paid to foreign consultants and expenses on travels abroad.

The above criterion was applied to the sugar projects using prefloating figures and considering a typical operating year. Assumptions made are given in Appendix I which also shows the type of statistics that are necessary for studies of this nature. A summary is given below:

1) Agricultural

a) Size of Plantation 2,000 hectares

b) Total Investment ₱8,480,000.00

c) Assumed Net Selling Price of Raw Sugar (Ex-warehouse) — ₱30.00/picul

d) Sugar Field — 92.32[a] piculs/ha. x 2000 has. = 184,640 piculs

e) Net Sales at ₱30/picul = ₱5,539,200.00

f) Expenses Deductible from Net Sales

i) Materials and Supplies	=	₱ 931,960.00
ii) External Services	=	80,100.00
iii) Indirect Taxes	=	48,180.00
Total		₱1,060,240.00

g) Net Domestic Value Added. D_v = ₱ 4,478.960

h) Scarce Resources Used, R_s

[a] Average share of planter in Negros Occidental.

i) Depreciation of Fixed and Replaceable Assets = ₱ 234,560.00

ii) Maintenance of Fixed and Replaceable Assets = 96,220.00

iii) Foreign Exchange Used = 596,340.00

Total ₱ 927,120.00

i) National Income Effect, $N_e = \frac{D_v}{R_s} \times 100 = 483\%$

2. *Milling*

a) Capacity of Mill 3000 tons cane/day

b) Total Investment ₱50,500,000.00

c) Sugar Yield = 1.655 piculs/ton cane x 3000 tons x 126 days = 625,590 piculs

d) Net Sales @ 33% Mill Share = 208,530 piculs x ₱30 = ₱6,444,900.00

e) Expenses Deductible from Net Sales

i) Materials and Supplies	= ₱ 961,000.00
ii) External Services	= ₱ 189,000.00
iii) Indirect Taxes	= ₱ 300,000.00
Total	₱1,450,000.00

f) Net Domestic Value Added, D_v = ₱4,994,900.00

g) Scarce Resources Used, R_s

i) Depreciation of Fixed and Replaceable Assets = ₱2,374,000.00

ii) Maintenance of Fixed and Replaceable Assets = 580,000.00

iii) Foreign Exchange Used 672,000.00

Total ₱3,626,000.00

h) National Income Effect, $N_e = \frac{D_v}{R_s} \times 100 = 138\%$

The national income effect as measured above by N_e, was used to determine the priority to be given to the sugar sector. A value of 100% or a one to one ratio between net domestic value added and scarce rescources used or a value less than 100% would mean minimal or negative effect on the national income. Projects could be rated according to the magnitude of N_e. A big value of N_e would mean a high and favorable effect on the national income and the higher would be its priority for investment purposes.

Note for example that in the above analyses the national income effect of the agricultural phase of sugar production gave an estimated value of 483% and milling an estimated value of 138% which are high enough to justify further support of the sugar industry. These values also support the observation that in some cases much more can be added to the national income by capital invested in agriculture instead of a manufacturing enterprise[3]. In setting policiès, therefore in the sugar industry first priority should be given to investments in agricultural equipment or improved agricultural methods.

Corrolarily, however, increased agricultural production would require additional sugar milling facilities.

The Balance of Payment Effect

Another important criterion for investment planning mentioned by Mr. Bantegui is the expected contribution of a sector to the foreign exchange reserve position of the country. This is in view of the critical foreign exchange position of underdeveloped countries like the Philippines. Projects are considered of top priority if they are expected to become substantial earners of scarce foreign exchange. Assuming we will have a ready outlet in the U.S. market, a sugar project can be classified as a potential foreign exchange earner. The problem is in measuring relative potentiality which we shall call as the balance of payment effect.

[3] Eugene Staley, "The Future Underdeveloped Countries", Harper and Brothers, N.Y. 1954, pp. 304.

It is not enough to measure the foreign exchange effect in terms of absolute amount of say, U.S. dollars earned. An equally important value is the amount of scarce resources used in the process of generating the much needed foreign exchange. Correlating the two, a measure that can be used is: $B_e = \frac{F_n}{R_s} \times 100$

where: F_n is the net amount of foreign exchange earned expressed in pesos. It is determined by subtracting from the foreign sales proceeds the amount of foreign exchange used in production.

While R_s, the value of scarce resources used, also expressed in pesos is determined by summing up depreciation and depletion cost, maintenance cost and foreign exchange cost as shown earlier.

For the sugar industry, assuming only two-thirds of production is exported, the balance of payment effects using pre-floating rate figures are measured below:

1) Agricultural — 2000 — Hectare Sugar Farm

a) Gros Foreign Exchange Earned = $\frac{2}{3}$ x 184,640 piculs x 3.90 = ₱3,840,512.00

b) Foreign Exchange Used = ₱ 596,340.00

c) Net Foreign Exchange Earned, F_n = ₱3,244,172.00

d) Scarce Resources Used, R_s = ₱ 927,120.00

e) Balance of Payment Effect, B_e = 350%

2) Milling — 3000 Tons Cane per day Mill

a) Gross Foreign Exchange Earned[a] = $\frac{2}{3}$ (208,530 piculs x \$8 + 3780 x \$ 13) 3.90 = ₱4,465,188.00

b) Foreign Exchange Used ₱ 672,000.00

c) Net Foreign Exchange Earned, F_n = ₱3,799,188.00

d) Scarce Resources Used, R_s = ₱3,626,000.00

e) Balance of Payment Effect, B_e = 105%

[a] Including sales proceeds from molasses.

Note again that the agricultural phase of sugar production has a more strengthening effect on the balance of payment position of the country compared to the milling phase, but the two operations have to be considered together. Jointly, they contribute highly favorable effects to the country's balance of payment position. Under this criterion and using the data available on hand the priority that has been given to the industry can be justified.

The National Economic Profitability Criterion

National economic profitability is defined as the total net measurable rate of return to the economy of an investment. It is also called the "social" rate of return in contra-distinction with the "private" or commercial rate of return. As a criterion it brings together in one measure all the major measurable economic costs and benefits of an investment to a developing country. For example, the commercial and national economic profitability estimates of an investment in a 3000-ton per day mill are shown below:

	Commercial Profitability	National Economic Profitability
Sales Revenues — ₱30/picul	₱6,444,900.00	₱7,304,220.00[a]
Operating Costs		
Labor, Salaries and Wages	881,000.00	528,600.00[b]
Materials and Supplies	381,000.00	434,340.00[c]
Repair and Maintenance	580,000.00	661,200.00[c]
Depreciation	2,374,000.00	2,374,000.00
Interest	1,251,000.00	1,251,000.00
Insurance and Indirect Taxes	300,000.00	100,000.00[d]
Miscellaneous	189,000.00	189,000.00
Total Operating Costs	₱5,956,000.00	₱5,538,140.00

[a]. . . . 20% added to 2/3 of revenue as premium for foreign exchange receipts

[b]. . . . 40% deducted from labor costs to reflect estimated alternative-use value of labor in agriculture

[c]. . . . 20% added to 70% of item to reflect imported cost and value of foreign exchange

[d]. . . . Taxes removed as they are not a cost to the economy

Net Profit, Before Tax	488,900.00	1,766,080.00
Income Tax	138,670.00	— d
Net Profit	₱3,350,230.00	₱1,766,080.00

a) Adjustment of revenue — Income from milling is derived from sales of U.S. and domestic quota sugars. Aproximately two-thirds output is sold to the American market at prices higher than the domestic rates. To both the individual producer and the economy, sales to the U.S. market mean more income in terms of actual peso revenue and premium on the foreign exchange earned. An upward adjustment of 20% on the revenue generated by export sales would approximate the additional benefits to the national economy.

b) Adjustment of Operating Costs — Positive and negative cost adjustments are necessary to reflect real costs to the economy. For a country like the Philippines where there is underemployment, the real cost of labor depends on the alternative use to which it be put to use. If utilized in agriculture its cost would be much less, hence the downward adjustment to the cost equivalent to the percentage ratio of minimum agricultural wages to industrial wages or approximately 60%. Imported materials for operation and maintenance are adjusted upwards by 20% to reflect their real costs due to the drawdown on the country's foreign exchange reserves.

c) Adjustment of Income — Finally positive adjustments in net profits are made by removing both direct and indirect taxes which are not true costs to the national economy.

The above comparison shows how an investment can yield a social rate of return three times as much the private rate of return. For planning the use of scarce funds a government institution can rank projects in the order of their national economic profitability.

The national economic profitability criterion is not without any limitations. First, like any other method it is also dependent on available statistics and it is therefore subject to a certain margin of error. However, the data requirements are not as rigorous compared with other sophisticated technique. Second, it does not consider "structural" disequilibrium or the unbalanced growth of sectors in the economy and hence it cannot be considered a complete measure for setting a national investment plan. As long as it is bear in mind that the national economic profitability criterion does not include factors which are not directly measurable and it does not discount altogether the use of value judgement then its utility as an analytical tool is not at all diminished.

Project-Oriented Planning

Although considerable progress has been achieved in recent years in the development of statistics in Asian countries their use for purposes of planning is of comparatively recent origin. Most statistics are derived largely as by products of administration. The biggest bottleneck in most countries is funds. We do not want to be misunderstood in claiming that comprehensive development plans can still be proper planning is enhanced by the availability of adequate data but when available statistical series are still fragmentary from the standpoint of planning for economic development a more practice alternative should be adopted. As Mr. Bantegui stated in the case of developing economies.

> "Even in the case of developing economies where comprehensive planning may not be immediately practicable due to shortage of basic data and planning experience, substantial benefits would still be derived from the deliberation of individual projects and development policies within the general scheme of an overall plan of development."

Deficiency in statistical data has been an important reason why many of the less developed countries do not produce comprehensive development plans. The example given in this paper regarding the sugar industry would serve to emphasize the project-oriented approach to planning. Economic development cannot be easily achieved by drawing up grandiose socio-economic development plans. In most instances as proven by experience, the most effective approach is through actual project development in individual industry sectors of high priority, still, there is a need of establishing the so-called priority areas and in so doing, statistics are still necessary.

The type of data presented in Appendix I of this paper are usually available in private organizations like the Philippine Sugar Association, accounting offices like Miller and Cruz and financing institutions like the Philippine National Bank. Unfortunately, these data do not usually find their way to government statistical offices. If industry statistics are collected in the same way that the sugar industry, through the efforts of the people behind it collect numerical information on their operations and these in turn collated by a government office, then it is possible to derive an input-output table, which can stand scrutiny. In the meantime, as long as there is doubt behind the accuracy of the figures appearing in the inter-industry-matrix, decision-makers will continue entertaining doubts on its applicability. It may very well serve the Bureau of Census and Statistics to enlist the statistical collecting arms of offices and institutions like the Board of Investments, the Board of Industries, the Development Bank of the Philippines, the Philippine National Bank, and the similar organizations who are able to collect factual information in exchange for tax and financial benefits which they can grant local business firms.

APPENDIX I

CALCULATING NATIONAL INCOME AND BALANCE OF PAYMENT EFFECTS

I. *Basis for agricultural costs computation per hectare*

A. Animal Plowing — 6 man days at ₱5.50/day

B. Clearing — 5 man days at ₱5.50/day

C. Tractor Plowing

1) Tractor driver at ₱6.00/day
2) Accomplishment or 2 hectares/day

D. Cane Points

1) ₱3.50 per lacsa at 7 lacsa/hectare
2) Hauling, ₱2.00/lacsa

E. Planting — 12 man days at ₱5.00 each
F. Tractor Furrowing
   1) Tractor driver at ₱6.00/day
   2) Accomplishment of 5 hectares/day
G. Animal Furrowing — 2 man days at ₱3.50/day
H. Hand Cultivation — 23 man days at ₱5.50/day
I. Plow Cultivation — 4.5 man days at ₱5.50/day
J. Tractor Cultivation
   1) Tractor driver at ₱6.00/day
   2) Accomplishment — 3 hectares/day
K. Replanting
   1) ₱3.50/lacsa, 2 lacsa/hectare
   2) Hauling, ₱2.00/lacsa
L. Applying Fertilizer — 3 man days at ₱5.00/day
M. Drainage — 2 man days at ₱5.50/day
N. Irrigation — 2 man days at ₱5.50/day
O. Pest Control — 2 man days at ₱5.50/day
P. Harvesting — 30 man days at ₱5.50/day

II. *Agricultural costs calculations*

A. Cost of Plowing per Hectare

1. Cost of Tractor ₱22,000.00

a. Depreciation	$= \dfrac{C - S}{\text{Hectare} \times \text{Years} \times \text{Days}} =$	
	$\dfrac{22{,}000 - 2{,}200}{2 \times 8 \times 70}$	= ₱ 13.76
b. Interest	$= \dfrac{22{,}000 - 2{,}200)\ 0.12}{2 \times 2 \times 90}$	= 6.60
c. H. I. T.	$= \dfrac{22{,}000 \times 2}{2 \times 90 \times 100}$	= 2.46
d. Daily Service and Repairs	= ₱5.13/2	= 2.57
e. Tire Changes	$= \dfrac{1.72}{2}$	= 0.86
f. Fuel and Oil	$= \dfrac{11.39}{2}$	= 5.69
g. Labor (Tractor Driver)	$= \dfrac{6.00}{2}$	= 3.00
	Cost Per Plowing per Hectare	= ₱ 34.94

2. Cost of Plow — ₱4527.00 (4 Disc Plow Massey Ferguson)

a. Depreciation	$= \frac{4627 - 231}{2 \times 10 \times 90}$	= ₱	2.44
b. Interest	$= \frac{(4627 - 231)\ 0.12}{2 \times 2 \times 60}$	=	2.20
c. H.I.T.	$= \frac{4627 \times 1}{2 \times 60 \times 100}$	=	0.39
d. Repairs	$= \frac{4627 \times 4}{2 \times 60 \times 100}$	=	1.54
	Plow Cost Per Hectare	₱	6.27
	Total Cost of Plowing Per Hectare	₱	41.21

Accomplishments — 2 Hectares/day

B. Cost of Harrowing Per Hectare

1. Cost of Disc Harow ₱5950.00 (18 Disc Harrow Mossy Ferguson)

a. Depreciation	$= \frac{5750 - 287}{3 \times 10 \times 60}$	= ₱	3.03
b. Interest	$= \frac{(5750 - 287)\ 0.12}{3 \times 2 \times 60}$	=	1.84
c. H.I.T.	$= \frac{5750 \times 2}{3 \times 2 \times 60}$	=	0.64
d. Repairs	$= \frac{5750 \times 4}{3 \times 60 \times 100}$	=	1.27

2. Tractor Cost, Per Hectare = 2/3 × (34.94) = 23.30

Total Cost of Harrowing Per Hectare = ₱ 30.08

Accomplishment — 3 Hectares per day

C. Cost of Furrowing Per Hectare

1. Cost of Furrower — ₱5627.00

a. Depreciation	$= \frac{4627.00 - 231.00}{6 \times 10 \times 90}$	= ₱	0.81
b. Interest	$= \frac{(4627 - 231)\ .06}{6 \times 2 \times 60}$	=	0.44
c. H.I.T.	$= \frac{4627 \times 2}{6 \times 60 \times 100}$	=	0.26
d. Repairs	$= \frac{4627 \times 4}{6 \times 60 \times 100}$	=	0.51

Cost of Furrower Per Hectare = ₱ 2.02

2. Total Cost of Tractor Per Hectare (Plowing Tractor)
   1/3 of 34.94 = 11.64
   Total Cost of Furrowing Per Hectare ₱ 13.66
   Accomplishment — 6 Hectares/day

D. Cost of Subsoiling Per Hectare (Optional)

1. Cost of Tractor = ₱78,700 (Crawler Type TD — 9) IH Price

a. Depreciation	$= \frac{78,700 - 2361}{3 \times 10 \times 90}$	= ₱ 28.28
b. Interest	$= \frac{(78,700 - 2361)\ 0.12}{3 \times 2 \times 90}$	= 16.97
c. H.I.T.	$= \frac{78,700 \times 2}{3 \times 90 \times 100}$	= 5.82
d. Repairs	$= \frac{78,700 \times 4}{3 \times 90 \times 100}$	= 12.00
e. Daily Service	$= \frac{78,700 \times 1}{3 \times 90 \times 100}$	= 2.91
f. Fuel and Oil	$= \frac{₱\ 7.00}{3}$	= 2.33
g. Labor (tractor driver)	$\frac{₱8.00}{3}$	= 2.66
Tractor Cost Per Hectare		₱ 70.90

2. Cost of Subsoiler ₱548.00

a. Depreciation	$= \frac{548 - 27}{3 \times 10 \times 60}$	= ₱ 0.29
b. Interest	$= \frac{(548 - 27)\ 0.12}{3 \times 2 \times 60}$	= ₱ 0.17
c. H.I.T.	$= \frac{548 \times .02}{3 \times 60}$	= 0.03
d. Repairs	$= \frac{548 \times 4}{3 \times 60 \times 100}$	= 0.12
Cost of Sub-soiler Per Hectare		— ₱ 0.61
Cost of Sub-Soiling Per Hectare		= 71.51

Accomplishment — 3 Hectares/day

E. Cost of Planting Per Hectare

1. Cost of Cane Points

p 80 per 1,000 (includes cost of gathering, cutting and preparation)

70,000 cane points at 50 per cent utilization (Rate of planting per hectare — 35,000 cane points per hectare	= ₱ 56.00

2. Hauling with the use of Tractor and Trailer (4 hectares/day)

a. Tractor cost per hectare $\frac{69.88}{4}$	= 17.47
b. 2 trailers with daily total cost = $\frac{28.64}{4}$	= 7.16
c. Labor — 2 Laborers for loading and distributing cane points at various destinations at ₱5.50 = $\frac{11.0}{4}$	= 2.75
3. Seed Treatment (lime magnesium sulphate)	= ₱ 3.00
4. Labor for Planting 12 man days to finish one hectare at ₱5.50	= 66.00
5. Replanting and Laborers 20 percent to replanted at ₱5.50	= 22.00
Cost of Planting per Hectare	₱174.38

F. Cultivation

Cost of Hilling up Per Hectare

1. Cost of Tractor/hectare	$= \frac{\text{(Tractor Plowing} \times 2/3)}{\text{per Hectare}}$	
	$= 2/3 \times 34.94$	= ₱ 23.30
2. Cost of Cultivation	$= \frac{\text{Plow Cost / day}}{3}$	
	$= \frac{6.27}{3}$	= 2.09
Cost per hectare		₱ 25.39

Accomplishment — 3 hectares/day

G. Off Barring

1. Use of Tractor and Cultivation

a. Cost of Tractor (per hectare) = 2/3 × 34.94	=	₱ 23.30
b. Cost of Cultivation per hectare = $\frac{6.27}{3}$	=	2.20
Off Barring Cost Per Hectare		₱ 25.50

Accomplishment — 3 Hectares/day

H. Weeding

1. Cost of Sprayer — Knapsack 5 gallons		=	₱150.00
a. Depreciation	$= \frac{150 - 7.50}{5 \times 60}$	=	.47
b. Interest	= (150.00 − 7.50) 0.12	=	.29
c. H.I.T.	$= \frac{150 \times 2}{600}$	=	.05
d. Repair	$= \frac{150 \times 4}{60 \times 100}$	=	.10
Spraying Cost / Day / Hectare		=	.91
Labor Cost at ₱5.50		=	5.50
Cost of Spraying			
Accomplishment 1 Hectare/day		=	₱ 6.41

2. Use of Chemicals

a. Cost of Spraying	=	₱ 6.41
b. Cost of Chemicals		
Rate Per Hectare		
a) 3 kilograms at 120 gallons of water		
b) Cost	=	₱ 30.00
Cost of weeding by use of chemicals	=	₱ 36.41
3 times / season, 1 hectare / day		₱109.23

APPENDIX I

III. *Agricultural Costs Breakdown per Hectare*

ITEM	Salaries & Wages	Materials	Maintenance Labor	Maintenance Materials	Depreciation	Interest	Insurance & Indirect Taxes	Foreign Exchange	External Service & Fringe Benefits
Plowing	₱ 3.00	₱ 6.55	₱ 1.64	₱ 2.47	₱ 16.20	₱ 8.80	₱ 2.85	₱ 6.30	₱ 0.25
Harrowing	2.00	4.37	1.19	1.79	12.20	6.24	2.28	4.30	0.20
Furrowing	1.00	2.18	0.55	0.82	5.40	2.64	1.08	2.10	0.10
Subsoiling	2.00	2.33	6.01	9.02	28.57	17.14	5.85	7.95	0.20
Planting	132.25	21.95	0.77	1.15	10.32	9.60	1.84	16.17	10.60
OFF-Barring	2.00	4.37	0.88	1.33	10.27	2.70	1.74	3.99	0.20
Weeding	16.50	90.00	0.30	—	1.41	0.87	0.15	63.00	1.30
Fertilization	5.50	224.80	—	—	—	26.98	—	157.36	0.40
Pest Control	5.50	34.24	0.10	—	0.47	0.29	0.05	23.97	0.40
Hilling — Up	2.00	4.37	0.88	1.33	10.20	2.65			
Irrigation & Drainage	22.00	7.58	3.55	5.33	17.24	10.34	3.55	9.04	1.80
Harvesting	165.00	25.00	—	—	—	19.80	—	—	13.20
Supervision	90.00	—	—	—	—	10.80	—	—	7.20
Overhead	50.00	10.00	4.00	5.00	5.00	6.00	3.00	—	4.00
T O T A L	₱498.75	₱437.74	₱19.87	₱28.24	₱117.28	₱124.85	₱ 24.09	₱298.17	₱40.05

IV. *Basis for Milling Costs Calculations*

A. Investment on 3000 Tons Cane Per Per Day Mill	=	₱ 50,500,000,
B. Total Canes Crushed = 126 days	=	378,000 tons
C. Sugar Produced = 1.655 piculs x 378,000 tons	=	625,590 piculs
D. Sugar Mill Share = 33 1/3 per cent x 625,590 piculs	=	208,530 piculs
E. Molasses Mill Share = 33 1/3 per cent x 378,000 tons	=	3,780 tons
F. Direct Mill Costs[a]		
1. Salaries and Wages, Milling	=	₱ 340,000.00
2. Salaries and Wages, General Shops	=	74,000.00
3. Salaries and Wages, Fabrication	=	190,000.00
4. Lime - 0.2 per cent (378.000) (40.00)	=	30,240.00
5. Fuel and Oil - ₱300/day x 140 days	=	28,000.00
6. Repair and Maintenance, 1 per cent of Investment	=	505,000.00
7. Materials and Supplies ₱250/day x 250 days	=	62,500.00
8. SSS, WCC and Fringe Benefits at 10 percent of Salaries and Wages		60,400.00
9. Miscellaneous and Additions to Assets	=	100,000.00

G. Indirect Mill Costs

1. Administration	₱ 600,000.00	600,000.00
a. Salaries and Wages	₱ 200,000.00	
b. SSS, WCC and Medical	20,000.00	
c. Bonuses and Allowances	17,000.00	
d. Office supplies, Cables materials	10,000.00	
e. Rentals	7,000.00	
f. Travel and Transportation	12,000.00	
g. Entertainment and Representations	10,000.00	
h. General Maintenance	5,000.00	
i. Taxes and Licenses	120,000.00	
j. Insurance	160,000.00	
k. Miscellaneous	39,000.00	

[a] All Materials.

H. Transportation

1. Total Investment = ₱ 3,300,000.00
2. Total Cost —378,000 tons x 10 kms x 232 = 876,960.00

The mill requires 125 tons of cane per hour. This would mean using 25 trailers with 5 — ton loads delivered hourly. This is a minimum requirement and to allow for any discrepancy, 30 trailer should be delivered every hour.

Loading Time. One of the factors that affects hourly deliveries to the mill is the time the farmers take in loading their cane. Only the day light hours are used for loading cane or 12 hours loading time. Since the mill needs 30 trailers per hour over a twenty-four period, the farmers have to load 60 trailers over a twenty-four period, or a total of 720 trailer loads for the twelve hour period.

Tractor to Trailer Ration. After loading the trailers are picked up from the different farms and delivered either to the mill or to the transfer stations. The time it takes to deliver a set of trailers and return empty for another load is known as the travel time. It is dependent on the distance of a farm from the mill and on the speed with which the trailer will be delivered. It has been estimated that farms located 10 to 15 kilometers from the mill will have travel times varying from 1 to 3 hours and those closer will have less than 1 hour. Based on actual experience in Silay (AIDSISA) in Negros Occidental the workable tractor to trailer ratio is 1:12 or 60 tractors for 720 trailers.

I. Cost of Irrigation / Hectares

1. Deep well Pumps 150 H.P. discharging 1,400 gpm at depth of 750 ft. Cost of Unit ₱40,000 including installation

a. Depreciation $= \frac{48,000 - 1440}{3 \times 10 \times 90}$ = ₱ 17.24

b. Interest $= \frac{(48,000 - 1440)\ 0.12}{3 \times 2 \times 90}$ = 10.34

c. H.I.T. $= \frac{48,000 \times 12}{3 \times 90 \times 100}$ = 3.55

d. Repairs $= \frac{48,000 \times 4}{3 \times 90 \times 100}$ = 7.11

e. Daily Service $= \frac{48,000 \times 1}{3 \times 90 \times 100}$ = 1.77

f. Fuel and Oil $= \frac{22.75}{3}$ = 7.32

g. Laborers, 4 laborers at ₱5.50

$$= \frac{5.50 \times 4}{3} \qquad = \quad 7.32$$

Irrigation Cost / Hectare = ₱ 54.91

Accomplishment — 3 Hectares / day

J. Pest Control

1. Cost of Chemicals

Diazimon (Insecticides)

a. Rate Per Hectare, 3240 mt. at 120 gallon/hectare

b. Cost = 34.24

Total Cost / Hectare ₱ 39.43

V. *Milling Costs Breakdown* (000) *Per Year — 3,000 Tons Cane Per Day*

ITEM	Salaries & Wages	Fuel & Materials	Maintenance & Repair (Parts etc)	Depreciation	Interest	Insurance & Indirect Taxes	Foreign Exchange	External Services & Fringe Benefits
General Plant	—	₱ 191.	₱ 505.	₱2,000.	₱1,050.	₱ 250	₱ 487.	₱ 30.
Milling	₱ 340	—	—	—	17	—	—	—
Fabrication	74	—	—	—	4	—	—	8.
General Shops	190	—	—	—	10	—	—	19.
Transportation	60	180	70	372	165	20	175	10.
Administration	217	10	5	2	36	30	10	88.
T O T A L	₱ 881	₱ 381	₱ 580	₱2,374	₱1,251	₱ 300	₱ 672	₱189

The Philippine Statistician
Vol. XIX, Nos. 3 and 4, 1970

DISCUSSION OF DR. ARTURO A. TRINIDAD'S PAPER*

By

B. G. BANTEGUI**

Dr. Arturo A. Trinidad introduced his paper entitled "Statistics in Economic Development Planning" with his statement: "This paper is being presented to discuss another paper entitled "Planning and Statistics" written by Mr. Bernardino G. Bantegui for the Seminar on Statistics and Planning held in Moscow, USSR on 22 September — 12 October 1969.

At the outset, it might be stated that I tried to accomplish the following in the aforementioned paper:

Firstly, I attempted to present the nature and various approaches to development planning. Among the various approaches cited were:

(a) The aggregative type which applies to the entire economy and deals with production, consumption, investment and the like as a single aggregate;

(b) The sectoral type which applies to individual sectors; and

(c) The interindustry type which is concerned with the relationship of the productive sectors of the economy with one another and each of these sectors with the rest of the economy.

Secondly, I indicated the data requirements for each of aforementioned three approaches in development planning and cited some important considerations in statistical operations in countries at various stages of economic development.

*Read at the Annual Conference of the Philippine Statistical Association, September 18, 1970.

**Vice Chairman, Congressional Economic Planning Office.

Thirdly, I presented a practical definition of industrial planning as well as its basic concepts and presented the experience in the Philippines regarding the preparation of the annual investment priorities plan by the Board of Investments including the evaluation procedures employed by the National Economic Council in assessing such yearly annual invesment priorities plan as required by law.

While my paper attempted to present the entire range of activities, problems and issues in development planning in relation to statistics development planning, Dr. Trinidad in his paper concentrated mainly on a section of the paper i.e. the portion relating the use of input-output data and techniques in evaluating the annual investment priorities plan, although the avowed purpose of the paper was to discuss the whole paper.

Before anything else I would like to point out some inaccuracies in the paper of Dr. Trinidad.

1. *Inaccurate abstracting*

While the orginal text of Dr. Trinidad's paper made no comment whatsoever on whether the results of the NEC were favorable or not, the abstract stated rather strongly that "the paper of Mr. Bantegui failed to report the results of the NEC study although local sources reported that their findings were unfavorable."

Allow me to read the pertinent portions of the abstract and that of the original text;

On the one hand the abstract read as follows —

> "It was noticed that the paper of Mr. Bantegui failed to report on the results of the NEC study although local sources reported that their findings were unfavorable."

The full text of the paper, on the other hand, read as follows:

> "However, we believe that the work of Mr. Bantegui is worthwhile discussing since it reported lengthily on the

> procedures adopted by the National Economic Council in assessing the feasibility of the Board of Investment's Second Investment Priorities Plan (IPP) but at the same time, it fails to report on the results of their evaluation.
>
> One could rationalize on the probable reason for this since the paper was presented before a foreign audience. It would not be good to criticize one's own government agency if *the findings had been unfavorable.*" (Underscoring provided)

If Dr. Trinidad had taken just a little time to look over the NEC study which evaluated the Second Annual Investment Priorities Plan (IPP) of the BOI, he would have noted the following statement relating to the results of the aforementioned NEC study:

> "On the whole the coverage and scope of the priority sectors included in the second IPP of the BOI are fairly adequate viewed from the social rate of return to the economy and related considerations."

Thus the statement to the effect that "local sources reported that their (NEC's) findings were unfavorable" is not only incorrect but appears gratuitous.

It may be important to note, however, that the difference in views between NEC and BOI as indicated in the study is not on the desirability of the projects proposed by the BOl but in the size of the "basket" of investment projects proposed to be granted incentives under the second annual IPP. For instance, the overall foreign exchange requirements of the second IPP of the BOI which amounted to $1.096 billion ($1.025 billion for machinery and equipment and $71 million for raw materials) would require an annual procurement of foreign funds amount to $365 million from 1969 to 1972. This amount is substantially higher than the highest annual foreign loan authorization ($152 million) that the Philippines ever received in a single year.

2. *Inaccurate reporting of facts on methodology employed in construction of NEC input-output tables.*

In the first paragraph on p. 6 of his paper, Dr. Trinidad stated:

> "There is also the question of numerical accuracy of the figures presented in input-ouput table. We can also surmise that none of the figures used were obtained through actual or direct observation. Most likely, the data used were derived from the usual indirect estimating procedures applied on standard sets of primary data whenever they are available."

The above statement of Dr. Trinidad has no factual basis and should be corrected. Instead of just "surmising" on the methodology employed in the compilation of hasic data and the techniques used in the preparation of the input-output tables, a researcher of Dr. Trinidad's caliber should have inquired about the subject from the source agency — the NEC. Had he done that we would have been happy to accommodate him by making available to him all the materials that have relevance to the subject. For his information a whole chapter of the publication entitled "The 1961 Interindustry (Input-Output accounts of the Philippines" (see pages 5-16) was directed to the description of the concepts and estimation procedures employed in the preparation of said input-output accounts.

Dr. Trinidad further stated that:

> "It is doubtful whether the estimates (referring to those obtained through indirect estimation procedures) can be corrected by interviewing businessmen or executives even if they are indirectly involved in the different industry concerned although it may be helpful."

It might interest Dr. Trinidad to know that aside from the traditional sources of basic data obtained from usual statistical source agencies, a large volume of additional information were obtained directly from business firms through mailed inquiries and/or through interviews with people directly involved in actual operations of private business covered in the investigation. I am making available to Dr. Trinidad NEC records on various meetings held to review and assess data collected from individual firms as well as from other sources which were used in the preparation of the sectoral reports in input-output table of the Philippines. The names of the technical people directly involved in each firm's operations (not solely businessmen or executives as Dr. Trinidad claims), the dates of meetings and the subject matter discussed in the meetings which the NEC held in various aspects of the work are available and Dr. Trinidad can gain access to these records at any time he wishes to look into these records.

3. *Forecasting figures on demand*

Dr. Trinidad criticized the use of Imput-Output techniques in predicting future levels of output on the basic of forecast figures on final demand claiming, among others, that (a) coefficients of production are not in reality rigorously fixed and that all sorts of variations are possible and do occur, and (b) process changes in manufacturing activities "will distort any estimate of input and output based on an old set of stock flow coefficients." He further stated that the crucial question is not whether coefficients of production are in a descriptive sense rigorously fixed but whether treating them as if they were constants involves neglecting factors that are considered minor disturbances but which are actually major determining influence. He suggests instead the use of simpler techniques that would simply describe their effects through some kind of trend or cyclical behavior. Dr. Trinidad states: "'The use of much understood methods of economic time series analysis may be more practical than the sophisticated input-output tables derived form insufficient information." *However, what Dr. Trinidad does not realize is that projecting a past trend through usual methods employed in time series analysis assumes that the factors causing the change in the secular trend* *will continue to have the same influence in the future as in the past — an assumption which Dr. Trinidad refuses to accept.*

4. *The need for obtaining an overall and sectoral view of the economy*

In setting up industrial priorities, the objective is to determine what industries to develop and support to bring about an accelerated rate of economic progress and thereby generate maximum economic and social welfare for the people. Input-output techniques, through which the *determinants of growth* of the various sectors of the economy are ascertained more directly than any other method, provide a powerful tool in analyzing and assessing development alternative to serve as basis for designing a blueprint of short-term and long-term development. Input-Output analysis emphasizes the manner in which technological production behaviours of various industries affect those of other industries and thus provides a deeper insight into the industrial structure of the economy.

In seting up industrial projects it is important to ascertain the effects of the increased production generated in an industry or sector on other industries and sectors of the economy. Dr. Trinidad proposed the application of the so-called national economic profitably criterion, also called social rate of return, as basis for setting up industrial priorities for profits in the Philippines.

It might interest Dr. Trinidad to know that such approach had been adopted by the NEC in passing on industrial projects submitted for inclusion in past NEC four-year development plans, even as early as 1956. We, at the NEC used to call such priority ratings of industries as the IPR (Industrial Priority Rating) instead of the term the social rate of returns used by Dr. Trinidad in his paper. As I have indicated earlier in this discussion of Dr. Trinidad's paper, we have found such approach inadequate because the interindustry relations among projects considered for implementations are ignored.

The Philippine Statistician
Vol. XIX, Nos. 3 and 4, 1970

THE OBSERVATIONS OF GUNNAR MYRDAL ON THE GROSS NATIONAL PRODUCT AS A MEASURE OF ECONOMIC DEVELOPMENT AND THE INADEQUACY OF STATISTICS IN ASIAN COUNTRIES*

By

LEVY A. TRINIDAD**

I *Introduction*

The observations of Mr. Gunnar Myrdal as amplified in the recent One-Asia Assembly, sponsored by the Press Foundation of Asia, and in his three-volume work "Asian Drama" can actually be divided to bear on two (2 subjects, namely:

1) On the Gross National Product:
That the output or income per head of a country and its rate of change can never give anything more than a rough indication of underdevelopment and the speed of development.

2) On Inadequate Statistics:
That statistics in South Asian Countries have to be scrutinized most severely before being used; at best, they are highly uncertain and not specific; that in the analysis of the development problems of these countries, unreliable figures are often used which greatly impair the value of the conclusions made.

* Read at the Annual Conference of the Philippine Statistical Association on September 18, 1970. The author acknowledges the valuable help and suggestions given by Dr. Arturo Q. Trinidad and his research staff in the preparation of this paper.

** Chief, Office of the Agricultural Marketing News Service, Bureau of Agricultural Economics, DANR.

II. *Rationale Behind Myrdal's Observations*

A. On the Gross National Product:

1) Both development and underdevelopment are composite concepts and they can not be defined mainly in terms of single aggregrates or averages.

2) The fundamental assumption of identity between production and income within a geographic area does not hold in South Asian economies. While the discrepancy between "net domestic product" and "net national income" is relatively small for the industrialized countries, the same is not true in developing economies which are are most alien-dominated, with substantial remission of interest, profits, salaries and other factor payments to non-residents. Thus a serious descrepancy exists between "net domestic product" and the income actually received by local residents, and as long as the major areas of their economies remain under foreign domination, the income received by the indigenous residents would be less than the value of the "net domestic product".

3) There is a high degree of difficulty in giving value to output since some segments of the developing economies are non-monetized and without much of a link with any markets. Non-marketed or bartered output are valued under different and divergent ways, and therefore cannot be expected to reflect their true values. Thus, there is a serious doubt about the aggregation of these estimates into a single figure to represent "output". This aggregate value has more meaning in developing economies where most goods and services are sold in competitive markets.

4) The inaccuracy or inadequacy of the Gross National Product estimates can also be traced to the extreme frailty of available statistical materials.

In Appendix 2, section 7, pp. 1968-1870, Vol. III of Asian Drama, Myrdal further made two (2) observations on the rate of growth of the National product or income per head as an indicator of development:

First, the increase in national income per head can only serve as an indication of the movement of the entire social system within wide margines of uncertainty. The reasons are:

a. The basic interdependence of all conditions that makes it possible to conceive of them as constituting a social system;

d. The dominant importance of people's income in their levels of living;

c. Levels of living are important, at least in the longer run, even for attitudes, patterns of behavior and institutions;

d. Our knowledge that if the movement of the entire social system do not change or if it lags very much, this will show up by preventing productivity and incomes from rising substantially.

Second, "development" cannot be defined in terms of growth of national income per head, but has to be defined as the upward movement of the entire social system.

Mr. Myrdal also stated that the difficulty with the term, "development" is that we actually mean something broader such that it gives rise to contradictions between its accepted definition and the value implications the term carries. A change in national income per head can only be used as a rough and ready indicator of that more complex change in the whole social system than what we really want to record. It is a crude way of estimating development since the statistical basis in underdeveloped countries is inadequate. Income distribution, age and sex structure and leisure time are neglected for two reasons:

First, the desire of people for development actually includes the desire to improve many conditions which have independent, as well as instrumental values for them;

Second, the interdependence of all conditions is not such that secondary changes in response to an initial impulse from income are always reinforcing; and when they are, their size is not proportional in all sectors and for all groups and subgroups of conditions.

Mr. Myrdal also emphasized that the term "national income" or "product" is not in itself a concept for the following reasons:

1) That magnitudes like the national income per head are inadequate to account fully for our conception of development. It has, in principle, been allowed in at least one respect, namely, by general qualification of income distribution. But income distribution, the age and sex structure, leisure time and so on are neglected in the aggregation of income;
2) In a changing society, the weights attached to the heterogenous collection of goods and services that make up the "national income" depend on attitudes, behavior patterns and institutions, which in turn change as a result of development;
3) Not only do attitude, etc. change with development but the values attached to the system of tastes change as well due to the changes in other conditions in the system. Thus, although aggregate income is valued according to taste and/or valuation, any increase in taste or valuation will also change income; and to judge the performance of the system by standards that are partly its own creation is therefore circular.

B—*On Inadequate Statistics*

1) Development of adequate statistics in South Asian Countries should cover the full range of economic and social data and not merely those for direct use in current macroeconomic models.
2) That conceptual problems are largely caused by a Western orientation of the type of statistics collected and the

methodology applied. The use of the Western concept of unemployment in South Asian Countries, for example, is erroneous. While in Western countries, those unable to work are generally involuntarily idle, this concept is not applicable to South Asian countries.

3) Lack of evaluation skill on the part of indigenous economists and analysts, and the tendency on their part to accept published data without question act as a deterrent to statistical development in South Asian countries. This is due to a low literacy and a short supply of staff capable of acting as enumerators, coders, editors, etc.
4) That useful aggregative generalization cannot be based on sampling if the universe being sampled is too diverse.

III. *Comments on Mr. Myrdal's Observations*

A. On the Gross National Product

Economists are cognizant of the inherent limitations of the GNP as an index of economic development. But unlike Myrdal who is overwhelmingly pessimistic about the uses of estimates despite some conceptual and statistical problems involved in its measurements.

Simon Kuznets of the National Bureau of Economic Research, for example, has pioneered in the studies of national income wealth and the conceptual difficulties in their measurement. Although his views antedate the book *Asian Drama,* they can be taken as a reply to Myrdal.

Mr. Kuznet stated:

"Comparisons of economic growth are beset with difficulties. The most important arises out of the doxical structure of this intellectual operation. Comparison requires reduction to a common denominator. In the case of economic growth, unlike that of short-term changes, movements over time often involve structural shifts and qualitative transformation of society; and differences in space are tied in with radical difference in cultural and social heritage. In essence, therefore, we are trying to compare the noncomparable. Yet we must, and we do so by imposing common yardsticks upon diverse societies. We thus act in the belief that either the basic purposes are the same for all these societies, or if not, that a yardstick clearly related to some recognized set of purposes, presumably one that makes economic growth important, will permit the recording of results of the process and supply information that can be used for a valid analysis of determining factors."[1]

The problem of assessing economic development is inherent in modern economic growth which implies major structural changes and modifications in social and institutional conditions under which the greatly in creased product per pita is attained. It does not mean that the changing component of the structure cannot be reduced to a common denominator for the purpose of comparing the product of the economy of the United States with that of China, or the product of an advanced country today with its output a century ago.

It has always been admitted by economists that any attempt to reduce national products of differing compositions and produced under diverse social conditions to a comparable aggregate can only be done by making certain assumption on scope (inclusion and exclusion), netness and grossness basis of valuation and the subdivision by industry of origin, factor and type of income, type of use, domestic and foreign origin, etc. It is important, however, to point out the major assumptions on the basis of which conceptual problem are usually resolved in actual measurement.

On the discrepancy between Net Domestic Product and Net National Income, it can be said that this has progressively narrowed in some South Asian countries. While this observation may hold true in the 1930's, the situation has changed during the post war era with the emergence of widespread

[1] "Problems in Comparison of Economic Trends", Economic Growth — Brazil, India, Japan, Kuznets, Moore and Spengler, ed. (North Carolina 1955).

movements for independence. These movements brought about restrictions on repatriation of earnings to foreign countries. In some cases, nationalization edicts and taxation imposed on industries owned and dominated by foreigners have resulted in the displacements of aliens by native residents. The same is true with high government and corporate positions which were formerly dominated by alien colonizers. Al these transitions have minimized distortions in assuming domestic output to be identical with income. With regard to the comparison of aggregate figures over time, the problems can be solved by deflating for price changes. However, it is altogether a different issue if the price indexes used are inadequate.

It is also true that in actual application, the sets of concepts and weights employed are those developed in one or a few countries. The weights chosen are usually those of the more developed rather than the less developed economies. The reason is first, the greater availability of data in advanced countries. Second, economic growth can be better evaluated from the vantage point of both in time and in levels of economic growth actually attained. Here the end justifies the means since the less advanced economies aspire to the levels of the more advanced, not vice versa.

There is no question that the weights and concepts developed in advanced countries result in a bias that favors the advanced economies. This upward bias is due to non-marketed products which are proportionately greater in less developed countries. However, there are also offsetting factors which create an opposite bias. The use of weights of the advanced economics for valuation for instance, tends to give many raw materials and simpler goods in the underdeveloped countries higher values compared to the complex products of advanced countries. We cannot, therefore, conclude that the common basis of valuation of products produces a bias that favors uniformly the more advanced economies.

In conclusion, we can state that:

"The conceptual and other difficulties of measurement do not justify the refusal to measure and the substitution of a cavalier treatment of uncontrolled impressions (even if embodied in apparently precise mathematical models) for the strenuous task of empirical corroboration and testing. Despite the limitations resulting from a basic problem regarding inadequate underlying data and from concepts that are outmoded because of a serious cultural lag, much can be learned by a determined scrutiny of the data provided that one looks at them with significant questions in mind and is sufficiently familiar with the characteristics of both the data and the underlying processes. Whatever mistakes one makes in the process and there will be many — can at least be corrected by others; cumulative improvement and learning are possible so long as the data are mobilized to serve as a basis of one set of generalizations and as a check on another".[2]

The observations made by Mr. Myrdal regarding the inadequacy of statistics in South Asian countries seem to imply that the problems being met by them are beyond solution. One has to look back and confirm that the same problems had confronted the advanced countries of the West in the past. Fortunately, in the past 2, or 3 decades, methods of collecting data either through complete enumerations or sample surveys have undergone tremendous improvement and there is no reason why these refinements cannot be applied to South Asian countries although they have been developed in Western countries. The doubts of Mr. Myrdal may be due in part to the failure to make counter-checks to possess the credibility of data used in the Asian Drama. The work did not state whether collecting agencies were consulted with regard to the background of the aggregate data collected for it is possible that the cause of the biases might have been unearthed.

To cite a specific example, response errors might have been the reasons why the figures on unemployment in South Asian countries appear to be inconsistent. Another reason could be difference in definition and coverage.

[2] Six Lectures on Economic Growth (New York 1961), 15-18. by Simon.

One important point which Mr. Myrdal failed to mention is the usefulness and importance of sample surveys undertaken periodically in between census periods. The Philippines, through the PSSH, for example, has been continuously conducting post enumeration sample surveys to look into the statistical make-up of family income and consumption.

On the Western orientation of Asian economists and statisticians, the reason again can be traced to convenience. There is nothing wrong with adopting method proven to have been effective in the more advanced countries, although Asian statisticians may be capable of the same fault as the Asian Drama in that they failed to question the background of published statistics by not going to the compiling agencies concerned.

The use of statistical sampling is the best answer for the shortage of experienced enumerators, coders, editors, etc. They need not have big sampling ratios considering the scarce resources in the developing countries. But certainly, the more advanced and sophisticated techniques of analyzing sample data should be adopted by Asian economists and statisticians. We have also to correct the impression created by the Asian Drama that generalization cannot be derived from samples, if the universe is heterogeneous. It may be true that larger samples, or multi-sampling, have to be resorted to but it certainly does not mean that useful estimates regarding diverse universe cannot be derived within certain limits.

We can only agree with Mr. Myrdal with regard to his observations on the low supply of experienced personnel and the lack of familiarity with statistics. As Mr. Mahalanobis stated,[3] "the collection and scrutiny of primary data are considered dirty work (in Asian countries) and should not be done with one's own hands. Consequently, it should not be the subject of training nor one should acquire skill in doing it."

B. *On Inadequate Statistics*

Mr. Harry Oshima in his paper, "Asian Studies in Income and Wealth" has also argued that due to the inadequacies of

[3] P.C. Mahalanobis "Statistics as a Key Technology", American Statistics, April 1965, p. 43

the basic data and to the lack of orientation for purposes of developmental analysis, the present accounts of Asian countries are of limited value to the analytical needs of underdeveloped economies.[4]

In most Asian countries, estimates of income are based on labor force statistics from decennial censuses. For the non-farm sectors, this has been claimed as the best approach to estimating the total national income or the Gross National Product. However, as the years get further away from the census year, the labor force data become increasingly obsolete and accordingly, the reliability of the deflated series of total national income becomes questionable, more so because the price series used for deflation are usually limited to a few large urban areas.

In income, ratios used to multiply employment in small productive units are based on scattered and unsystematic surveys conducted in a limited number of areas. It is, therefore, expected that not much confidence can be placed on the money and the real values of income which they are supposed to measure. Since small units make up 80 per cent of the non-farm output in South Asian countries, these defects in the data covering employment, output, income and price statistics are very serious.

Regarding the national accounting of expenditure, estimates of private capital formation in most Asian Countries "may serve as an indication of the order of magnitude".[5] In some cases, they are even misleading as an indication of the order of magnitude. There is no adequate coverage of fixed capital formation on the farms, in the small industries and businesses. There is also no fixed coverage of inventory changes. Comprehensive data for the direct estimation of private consumption are lacking, and therefore, are only ob-

[4] Asian Studies on Income and Wealth, Papers presented at First Asian Conference on Income and Wealth, Hongkong, August 1960 International Asso., for Research in Income and Wealth, N.Y. Pub. House, 1965.

[5] American Statistical Association Journal, Paper written by William Abraham, p. 669.

tained as a residual by deducting government expenditures, net foreign investment and capital formation from Gross National Product which, in turn, is derived from income. Consequently, personal savings which is the difference between personal income and private consumption, cannot be taken seriously. Similarly, the major components of private consumption, food expenditures, clothing, etc., may not be reliable enough for purposes of computation of income elasticities. The basic difficulties lie in the lack of sufficient raw data.

In the most advanced countries, statistics are collected by public and private offices, in addition to censuses and regular surveys. In the developing countries of South Asia, it is difficult to obtain data from administrative and corporate sources. For instance, for some reasons, a large number of small firms do not bother to comply with the provision of the Social Security Law or the National Internal Revenue Code. Thus the data collected or kept by the government for tax purposes or by business firms for trade purposes are likely to be deficient for statistical purposes. Since the costs of undertaking censuses are high, most of the data needed for nation accounting purposes cannot be collected on a regular basis except during the decennial census. Of course, the alternative is to conduct sampling surveys and incur errors either through lack of response or through defective design.

In most South Asian countries, including the Philippines, probability sampling on households is being undertaken regularly and on the whole, the results seem to be very encouraging. Labor force statistics, agricultural and production data, household income and expenditure are regularly collected by sampling surveys. In the case of the Philippines through the joint efforts of the U.S. Agency for International Development and the National Economic Council of the Philippines, a comprehensive statistical program has been launched. The PSSH has regularly conducted sampling surveys of households and errors due to sampling have been reduced to levels low enough to make them adequate for analytical purposes. The main problem behind the surveys may still be traced to inadequate responses and in the case of response errors, they are likely to be due to the insufficient knowledge of respondents regarding the information sought or the lack of desire to give information which they consider personal and confidential. It is also difficult for respondents to remember the various sources and amounts of income received during the past months and the various components of their household expenditures for them to give fairly accurate answers. Since probability sampling surveys are the most economical media for getting data where productive activities are carried out by innumerable small units, the problem that should be solved by statisticians in Asian countries should bear on response errors and the means of reducing them.

CONCLUDING OBSERVATIONS

The criticisms of Mydral as expressed in the Asian Drama provide a new perspective to the existing gaps in the statistical systems of developing countries. This view point from an outsider of the Asian economic scene, although susceptible to valid rebuttal in some aspects, deserves more than passing considerations. It affords a greater stimulus to national introspection, to reexamine possibly and often frequently existing deficiencies which many Asians, not excluding Filipinos, in their anxious endeavors for progress, might fail to notice.

The Philippine Statistician
Vol. XIX, Nos. 3 and 4, 1970

MEASUREMENT OF DIGITAL PREFERENCE IN MALAYAN AGE STATISTICS

By
SAW SWEE-HOCK*

Among the many inaccuracies present in population census statistics, perhaps the most important one is that pertaining to mis-statement in ages which may be attributed to several factors such as ignorance of ages, negligence in computing the exact age, misunderstanding of the age question and deliberate misrepresentation. This paper is confined to an analysis of the type of mis-statement that leads to preferences for terminal digits. Unlike previous occasions when the age statistics, except for children under five years old, were published in quinary age groups only, the 1957 census age returns have been presented in single years of age, thus permitting a detailed study of the extent and nature of digital preference in the age statistics.[1]

A cursory inspection of the 1957 census tables depicting the population by single years of age is sufficient to reveal the marked preference for certain ages particularly those with terminal digits 0 and 5. For a more precise assessment of the errors as well as a comparison of these errors among the different sex-race components of the population, it is necessary to resort to some form of an index to measure the age inaccuracies. One such form is the measure known as Whipple's Index[2] which

* Prof. Saw Swee-Hock is Professor of Statistics and Head of the Department of Statistics at the University of Hongkong.

[1] H. Fell, 1957 *Population Census of the Federation of Malaya*, Report No. 14, Kuala Lumpur: Department of Statistics, 1960, Table 4, p.

[2] For an account of Whipple's Index, see United Nations, *Manual II: Methods of Appraisal of Quality of Basic Data for Population Estimates*, ST/SOA/Series A, Population Studies No. 23, New York, 1955, pp. 40-41.

can be used to gauge the extent of heaping at the most preferred terminal digits 0 and 5 within the convenient age range of 23 to 62.

TABLE I
MEASUREMENT OF DIGITAL PREFERENCE BY WHIPPLE'S INDEX IN 1957 MALAYAN CENSUS BY SEX AND RACE

Race		Whipple's Index	U.N. Grading
All Races:	Male	137.3	Rough
	Female	150.4	Rough
Malays:	Male	163.6	Rough
	Female	185.3	Very Rough
Chinese:	Male	106.6	Fairly Accurate
	Female	104.8	Highly Accurate
Indians:	Male	132.7	Rough
		141.7	Rough

The results of the calculations are set out in Table 1. The figures show that the degree of heaping is by far the most severe among the Malays, slightly less severe among the Indians, and very much less serious among the Chinese. According to the United Nations grading the age accuracy of the Malays may be classified as rough for the males and very rough for the females and that of the Chinese as fairly accurate for the males and highly accurate for the females. In the case of the Indians the age accuracy of both sexes may be considered as rough.[3]

For an inquiry into the extent of preference or dislike for each and every one of the ten terminal digits, it is essential to employ a more thorough measure such as Myers' Index.[4]

[3] United *Nations, Demographic Yearbook 1960*, New York, 1960. The full grading given in this publication is as follows:—

Less than 105	Highly accurate
105 - 109.9	Fairly accurate
110 - 124.9	Aproximate
125 - 174.9	Rough
More than 175	Very rough

[4] For an account of the theory of the Index, see R.J. Myers, "Errors and Biases in the Reporting of Ages in Census Data", *Actuarial Society of America' Transactions*, Vol. XLI, Part 2, No. 104, October 1940, pp. 395-415.

The results of the computations made over the age range 13-92 are presented in Table 2 in terms of the percentage deviation from 10 per cent, with the positive sign denoting preference and the negative sign dislike. Myer's Index is the sum of the deviation from 10 per cent, ignoring the signs, and may be taken as an overall measure of the general accuracy of age statement over the chosen age range. As the computed value of Myer's Index tends to 0, the greater is the accuracy of the age statement.

An examination of the results will bring out the somewhat marked similarity in the pattern of age mis-statement for both the Malays and the Indians. There does not seem to be any clear preference for even numbers to odd numbers, but it is apparent that 0, 5 and 8 are preferred and the other seven terminal digits are disliked. The order of preference for the three favoured digits are the same, but there is the difference that whereas preference for 0 and 5 is on similar intensity in the case of the Indians, 0 is preferred much more than 5 by the Malays. The order of dislike for the other seven digits are not similar except that the most disliked digit happens to be 1 in both cases. Another resemblance is that by comparison with the females, the males tend to exhibit a less pronounced preference or dislike as the case may be. However, it is to be noted that, for males as well as females, the age reporting among the Malays is not as faithful as that of the Indians.

The results of the Chinese should be regarded with some reservations since the magnitude of the errors (as calculated) is so small that it may be deduced that Myer's Index is not sensitive enough to measure it.[5] In fact, it is doubtful whether the calculated values prove anything except that either the Chinese have no digital preference error or, if they do have an error, it is very small indeed and what digits are preferred cannot be deduced. However, the results seem to be in agreement with the findings of Whipple's Index in that the normal

[5] See N.H. Carrier, "A Note on the Measurement of Digital Preference in Age Recordings", *Journal of the Institute of Actuaries*, Vol. 85, Part 1, No. 369, June 1959, pp. 71-85.

TABLE 2
MEASUREMENT OF DIGITAL PREFERENCE BY MYERS' INDEX IN 1957 MALAYAN CENSUS BY SEX AND RACE

Terminal Digit	All Races		Malays		Chinese		Indians	
	Male	Female	Male	Female	Male	Female	Male	Female
	Deviation from 10 per cent							
3	— 0.86	— 1.22	— 1.19	— 1.89	— 0.60	— 0.52	— 0.43	— 0.38
4	— 1.18	— 1.38	— 2.10	— 2.38	— 0.25	— 0.27	— 0.37	— 0.46
5	+ 2.22	+ 2.95	+ 3.81	+ 5.11	+ 0.16	+ 0.01	+ 2.86	+ 2.94
6	— 1.10	— 1.53	— 2.36	— 2.93	+ 0.41	+ 0.37	— 0.75	— 1.43
7	— 0.50	— 1.15	— 1.33	— 2.38	+ 0.49	+ 0.46	— 0.16	— 0.90
8	+ 0.80	+ 1.20	+ 1.37	+ 1.73	+ 0.27	+ 0.46	+ 0.67	+ 1.33
9	— 0.83	— 0.91	— 1.24	— 1.34	— 0.38	— 0.21	— 1.27	— 1.14
0	+ 4.54	+ 6.09	+ 7.95	+ 10.39	+ 0.87	+ 0.78	+ 2.91	+ 3.82
1	— 2.09	— 2.66	— 3.24	— 4.11	— 0.58	— 0.66	— 2.64	— 2.85
2	— 101	— 1.40	— 1.66	— 2.21	— 0.40	— 0.42	— 0.82	— 0.95
	15.11	20.49	26.25	34.47	4.40	4.16	12.88	16.20

feature of a much more pronounced age mis-statement among the females than males does not appear to prevail in the Chinese population.

The age statements of the Chinese are far less erroneous than those of the other two races, and are even superior to those of many Western countries, it would be misleading to reach for an answer in terms of the commonly known factors ascribed to age mis-statement; nor would it be quite true to say that "the Chinese have a fairly accurate knowledge of their ages".[6] The excellent quality of the Chinese age statistics is primarily the outcome of employing the animal-year method of collecting the Chinese age returns. Though the new method was designed to eliminate systematic over-statement of ages resulting from the peculiar Chinese way of counting ages, it also helped to reduce the frequency of digital preference when the supplementary information on the animal-year of birth was employed to calculate the Chinese ages at last birtday according to the Western method of reckoning.[7]

[6] H. Fell, *op. cit.*, p. 37. It is interesting to note that Fell, for some, unkown reasons, did not mention the connection between the introduction of the new method of collecting Cninese age statistics and the good quality of the Chinese age data compiled in the 1957 Census.

[7] For a detailed discussion of the principles underlying the new method using animal-year information and an appraisal of the method utilized in the 1957 Census, see Saw Swee-Hock, "Errors in Chinese Age Statistics," *Demography*, Vol. 4, No. 2, 1967, pp. 859-875.

The Philippine Statistician
Vol. XIX, Nos. 3 and 4, 1970

ON THE DUAL OF SOME BALANCED INCOMPLETE, DESIGNS

By

HARDEO SAHAI*
University of Kentucky

§ *Introduction*: Bose and Nair (1939) in their original paper on Partially Balanced Incomplete Block Design (PBIBD) discussed a variety of different methods of constructing PBIB designs. They also introduced the principle of duality between blocks and varieties and obtained some PBIB designs by inverting some balanced incomplete block designs (BIBD) of Yates with $b \neq v$.

Shrikhande (1952) has proved that the duals of the BIB designs with

(i) $v, b. r, k, \lambda = 1$

(ii) $v, b. r = k + 2, \lambda = 2$

are PBIB designs. He also gave some easy methods for writing down the designs thus obtained by using the technique of "method of differences" as developed by Bose (1939). This is an expository paper on the application of the dual of balanced incomplete designs in constructing some PBIB designs.

§ 1. Definition of Balanced Incomplete Block Design (BIBD):

An arrangement of v varieties in b blocks of k units each (k v) is known as BIBD if every variety occurs once and only once in r blocks and two varieties occur together in blocks.

* Revision of a term paper submitted in partial fulfillment of the requirements for the degree of Master of Science at the University of Chicago.

The Philippine Statistician
Vol. XIX, Nos. 3 and 4, 1970

ON THE DUAL OF SOME BALANCED INCOMPLETE, DESIGNS

By

HARDEO SAHAI*
University of Kentucky

§ *Introduction*: Bose and Nair (1939) in their original paper on Partially Balanced Incomplete Block Design (PBIBD) discussed a variety of different methods of constructing PBIB designs. They also introduced the principle of duality between blocks and varieties and obtained some PBIB designs by inverting some balanced incomplete block designs (BIBD) of Yates with $b \neq v$.

Shrikhande (1952) has proved that the duals of the BIB designs with

(i) v, b. r, k, $\lambda = 1$

(ii) v, b. $r = k + 2$, $\lambda = 2$

are PBIB designs. He also gave some easy methods for writing down the designs thus obtained by using the technique of "method of differences" as developed by Bose (1939). This is an expository paper on the application of the dual of balanced incomplete designs in constructing some PBIB designs.

§ 1. Definition of Balanced Incomplete Block Design (BIBD):

An arrangement of v varieties in b blocks of k units each (k v) is known as BIBD if every variety occurs once and only once in r blocks and two varieties occur together in blocks.

* Revision of a term paper submitted in partial fulfillment of the requirements for the degree of Master of Science at the University of Chicago.

associates with which we start. This number is denoted by p^i_{jk} clearly $p^i_{jk} = p^i_{kj}$

In the particular case when m = 1, we get BIBD.

The numbers v, b, r, k, $\lambda_1 \ldots, \lambda_m$; $n_1, n_2, \ldots, n_m$ may be called the parameters of the first kind, and the numbers p^i_{jk} (i, j, k = 1, 2, ... m) the parameters of the second kind. Thus there are 2m + 4 parameters of the first kind, and $\frac{m^2(m+1)}{2}$ parameters of the second kind. (Since $p^i_{jk} = p^i_{kj}$)

The following relationships amongst the parameters can be established

$$bk = vr \quad (2.1)$$

$$v - 1 = n_1 + n_2 + \ldots + n_m \quad (2.2)$$

$$r(k-1) = n_1 \quad {}_1 + \ldots + n_m \quad {}_m \quad (2.3)$$

$$1/2v(n_1 + \ldots + n_m) = v(v-1) \quad (2.4)$$

$$\sum_{k=1}^{m} p^i_{jk} = n_{j-1} \text{ or } n_j \text{ according as } i = j \text{ or } i \neq j \quad (2.5)$$

$$n_i\, p^i_{jk} = n_j\, p^j_{ik} \quad (2.6)$$

Given the parameters of the first kind there are only $m(m^2-1)/6$ parameters of the second kind. When m = 1 i.e. when our design is BIBD, there are no independent parameters of the second kind.

3. Definition of the Dual of a BIBD:

Consider a BIBD with parameters v, b, r, k and λ. Let the varieties and blocks of the design be denoted by $T_1, \ldots, T_v$ and $B_1, B_2, \ldots, B_b$. Let the block B_i contain the treatments $T_{i1}, T_{i2}, \ldots, T_{ik}$. Then a new design not necessarily balanced, can be obtained from this design by letting the varieties and blocks of the original design become the blocks and varieties respectively of the new design in which the variety number i occurs in blocks numbered $B_{i1}, B_{i2}, \ldots, B_{ik}$ (i = 1, 2, ..., b).

Hence for the new design v = b, b = v, r = k, k = r. This design is called the dual of the original design.

For illustration consider the design with parameters v = 6, b = 10, r = 5, k = 3, λ = 2.

It is given by the following 10 blocks.

(1, 2, 3)
(2, 3, 4)
(3, 4, 5)
(4, 5, 1)
(5, 1, 2)
(1, 3, 6)
(2, 4, 6)
(3, 5, 6)
(4, 1, 6)
(5, 2, 6)

The design dual to this given by the following 6 blocks.

(1, 4, 5, 6, 9)
(1, 2, 5, 7, 10)
(1, 2, 3, 6, 8)
(2, 3, 4, 7, 9)
(3, 4, 5, 8, 10)
(6, 7, 8, 9, 10)

The first block of the dual design contains varieties 1, 4, 5, 6 and 9 since the variety number 1 in the original design occured in blocks numbered 1, 4, 5, 6 and 9. The other blocks of the dual design are similarly obtained and in particular the last block of the dual design corresponds to variety numbered 6 which occurred in blocks numbered 6, 7, 8,9 and 10.

The designs with b = v are self dual so that we will not get new design from them. There are many cases where no dual design of any sort exists.

Now we will discuss the above example in terms of the incidence matrix of the design.

Let n_{ij} denote the number of times the ith variety occurs in the jth block. $(i = 1, \ldots v, j = 1, 2, \ldots, b)$ Then

$$N = (n_{ij})$$

will be called the incidence matrix of the design.

Given the incidence matrix of the design it is easy to write down the layout of the design. In any block only those varieties occur corresponding to which there is unity in the incidence matrix.

In the present example, the incidence matrix of the design is

$$N = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

It is obvious that the incidence matrix of the dual design is the transpose of the incidence matrix of the original design.

Thus the incidence matrix of the dual design is

$$\begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 \end{bmatrix}$$

4. Some More Definitions:

4.1. Symmetrical Designs: The design with $b = v$, $r = k$ is called symmetrical design.

4.2. Complementary Design: The design obtained by taking in any block the varieties which do not occur in that block, is called complementary design. The parameters of the complementary design are

$$v^* = v,\ b^* = b,\ r = b\text{-}r,\ k^* = v\text{-}k \text{ and } \lambda^* = b\text{-}2r + \lambda.$$

If N is the incidence matrix of the original design then E-N is the incidence matrix of the complementary design where E is a v x b matrix with all its elements equal to unity.

4.3. Derived Design: In a Symmetrical BIBD (SBIBD) by omitting a block and retaining in the remaining blocks only the varieties contained in the omitted block, we obtain a new design which is a BIBD with parameters

$$v^* = k,\ b^* = b\text{-}1,\ r^* = r\text{-}1,\ k^* = \lambda \text{ and } \lambda^* = \lambda -1$$

This new design is called the derived design.

4.4. Residual Design: In a symmetrical BIBD by omitting a block and all the varieties contained in this block from remaining blocks, we obtain a new design which will be a BIBD with parameters

$$v^* = v\text{-}k,\ b^* = b\text{-}1,\ r^* = r,\ k^* = k-\lambda \text{ and } \lambda^* = \lambda-1.$$

The new design is called the residual design.

5. Dual of a BIBD with $\lambda = 1$

Theorem 5.1. The dual of a BIBD with parameters $v = (cd - c+1)$, $b = (cd - c+1)\ c/d$, $r = c$, $k = d$ and $\lambda = 1$ is a PBIBD with parameters (5.1)

$$v = (cd - c+1)\ c/d,\ b = (cd - c+1),\ r = d,\ k = c$$

$$n_1 = d(c-1),\ n_2 = (c\text{-}d)\ (d-1)\ (c-1)/d,\ \lambda_1 = 1,\ \lambda_2 = 0$$

and

$$\begin{bmatrix} p^1_{jk} \end{bmatrix} = \begin{bmatrix} (c-2)+(d-1)^2 & (d-1)(c-d \\ (d-1)(c-d) & \dfrac{(d-1)(c-d)(c-d-1)}{d} \end{bmatrix}$$

$$\begin{bmatrix} p^{(2)}_{jk} \end{bmatrix} = \begin{bmatrix} d^2 & d(c-d-1) \\ d(c-d-1) & (c-d)^2 + 2(d-1) - \frac{c(c-1)}{d} \end{bmatrix}$$

Lemma 5.1. In a BIBD with parameters v, b, r, k, λ, if we take any initial block then remaining blocks fall into groups of ϕ (n) blocks having n varieties common with the initial block (n = 0, 1, 2, ..., k), and the following relations hold.

$$\sum_{n=o}^{k} \phi(n) = b - 1 \qquad \text{(i)}$$

$$\sum_{n=o}^{k} n\,\phi(n) = k(r-1) \qquad \text{(ii)}$$

$$\sum_{n=o}^{k} n(n-1)\,\phi(n) = k(k-1)(\lambda-1) \qquad \text{(iii)}$$

Proof: Since with respect to any initial block the remaining b − 1 fall into groups of ϕ (0), ϕ (1), ..., ϕ (k) we have $\phi(0) + \phi(1) + \ldots, + \phi(k) = b - 1$

Now consider any particular block θ. It contains k varieties. Each of these varieties occurs in r − 1 other blocks. Hence any particular block is a member of k(r − 1) pairs. But ϕ (n) blocks (which have n varieties in common with the initial block), each yield n pairs. Hence we have

$$o\,\phi(o) + 1\,\phi(1) + \ldots + k\,\phi(k) = k(r-1).$$

For proof of relation (iii) see Hussain (1948).

Proof of Theorem 5.1: Let A be any block of the design (5.1). Since $\lambda = 1$, no other block of the design can have more than one variety in common with A.

Let m_o be the number of blocks A_o each of which has no variety in common with A and m_1 the number of blocks A_1 each having one treatment in common with it. Then using the results of Lemma 5.1 we have

$$m_o + m_1 = (cd - c + 1)\,\frac{c-1}{d}$$

and $m_1 = d(c-1)$

Hence $m_o = (c-d)(c-1)(d-1)/d$

In the dual design, with respect to any given variety, the other varieties can be divided into two sets of size n_1 and n_2 respectively where $n_1 = d(c-1)$ and $n_2 = (c-d)(c-1)(d-)/d$ such that the varieties of the first set occur λ_1 $(=1)$ times and those of the second set occur λ_2 $(=0)$ times with the given variety.

Let a non-initial variety (i.e. not occuring in A) occur y_o times in A_o and y_1 times in A_1.

Then $y_o + y_1 = c$

and remembering that it must occur once with every initial variety (since $\lambda_1 = 1$), we get

$o.y_o + 1.y_1 = d$ i.e. $y_1 = d$

Hence $y_o = c\text{-}d$

Let B be a block of A_o. Let there be z_o blocks B_O and z_1 blocks B_1 in A_o. Then

$z_o + z_1 = m_o - 1 = (c\text{-}d)(c\text{-}1)(d\text{-}1)/d\text{-}1$

Since B has no variety in common with A, it has d non-initial varieties. In the remaining $m_o - 1$ blocks of A_o these d non-initial varieties occur $y_o - 1 = c - d - 1$ times and the $c-d-1$ blocks of A_o containing a given treatment of B must be different from the $c-d-1$ blocks corresponding to any other treatment of B, we must have

$z_1 = d(c-d-1)$

Hence $z_o = (c-d)^2 + 2\ (d-1) - c(c-1)/d$

Let C be a block of A_1. Let there be u_o blocks C_o and u_1 blocks in A_o. Then

$u_o + u_1 = m_o = (c\text{-}d)(c\text{-}1)(d\text{-}1)/d$

Again the block C consists of 1 initial variety and $c-1$ non-initial varieties with reference to the standard A. In the set of m_o blocks of A_o this initial variety occurs zero times and the $c - 1$ non-initial varieties occur $c-d$ times so that

$$o.u_o + 1.u_1 = o.1 + (c-d)(c-1)$$

i.e. $u_1 = (c-1)(c-d)$

Hence $u_o = (c-d)(d-1)(c-d-1)/d$

For the dual design the blocks A and B become 2-associates and the number of varieties which are 1-associates of the variety corresponding to A and simultaneously 2-associates of the variety corresponding to B is the number of blocks B_1 each of which has no variety in common with A and simultaneously one variety in common with B. Thus $p^2_{12} = z_1 = d(c-d-1) = p^2_{21}$. Similarly the number of varieties which are simultaneously 2-associates of these two varieties of the dual design is obviously the number z_o of blocks B_o each of which has no treatment in common with both A and B. Thus

$$p^2_{22} = (c-d)^2 + 2(d-1) - c(c-1)/d.$$

Further the blocks A and C become 1-associates and we easily see that

$$p^1_{12} = u_1 = (c-1)(c-d) = p^1_{21}$$

$$p^1_{22} = u_o = (c-d)(d-1)(c-d-1)/d$$

$$p^1_{11} = (c-2) + (d-1)^2$$

Thus the dual design is a PBIBD with just two types of associates.

Corollary 5.1. The dual of a BIBD with parameters

$$v = n,\ b = \frac{n(n-1)}{2},\ r = n\text{-}1,\ k = 2,\ \lambda = 1$$

is a PBIBD with parameters.

$$v = n\frac{(n-1)}{2},\ b = n,\ k = n-1,\ r = 2,\ \lambda_1 = 1,\ \lambda_2 = 0$$

$$n_1 = 2(n-2),\ n_2 = \frac{(n-2)(n-3)}{2}$$

$$\left[\ p^1_{ij}\ \right] = \left[\begin{array}{cc} n-2 & n-3 \\ n-3 & \frac{(n-3)(n-4)}{2} \end{array}\right]$$

$$\left[\ p^2_{ij}\ \right] = \left[\begin{array}{cc} 4 & 2(n-4) \\ 2(n-4) & \frac{(n-4)(n-5)}{2} \end{array}\right]$$

Proof: Put $c = n-1$, $d = 2$ in Theorem 5.1 and result follows immediately.

Corollary 5.2. By omitting all the blocks containing any particular variety from the SBIBD with parameters

$$v = b = p^2 - p + 1,\ r = k = p,\ \lambda = 1$$

we get a PBIBD with parameters

$$v = p(p-1),\ b = (p-1)^2$$

$$r = p-1,\ k = p,\ \lambda_1 = 1,\ \lambda_2 = 0$$

$$n_1 = (p-1)^2,\ n_2 = p-2$$

$$\left[\ p^2_{jk}\ \right] = \left[\begin{array}{cc} (p-1)(p-2) & p-2 \\ p-2 & 0 \end{array}\right]$$

$$\left[\ p^1_{jk}\ \right] = \left[\begin{array}{cc} (p-1)^2 & 0 \\ 0 & p-3 \end{array}\right]$$

Proof: Consider the residual design of the given BIBD which will be a BIBD with parameters

$v = (p-1)^2$, $b = p(p-1)$, $r = p$, $k = p-1$ and $\lambda = 1$

The dual of the residual design can be obtained by omitting all the blocks containing any particular variety. (This variety corresponds to the block omitted in obtaining the residual design.) Putting $c = p$ and $d = p-1$ in theorem 5.1, the result follows.

6. The Method of Differences:

Most of the designs obtained by the above theorem can be easily constructed by the method of differences. The method of differences has been extensively discussed and used by Bose (1939). We shall give here the application of this method to construction of PBIBD obtained as a dual to the BIBD.

A set of elements is said to form a modul M, when there exists a law of composition viz. The addition denoted by +, satisfying the following axioms.

(i) To any two elements a and b of M, there exists a unique element s of M defined by $a + b = s$

(ii) $a + b = b + a$

(iii) $a + (b+c) = (a+b) + c$

(iv) To any two elements a and b of M there exists an element x belonging to M satisfying $a + x = b$

On the basis of these axioms we can prove that the element x in (iv) is unique. Also there exists a unique element o, with the property that c being any element of M, $c + o = c$. If $c + d = o$, we denote d by -c $a + (-c)$ may be denoted by $a - c$. The element x in (iv) is then equal to b-a, and may be said to be the difference of b and a.

Consider a modul M containing exactly n elements. To each element $u^{(i)}$ of the modul let there correspond one variety which may be denoted by $u^{(i)}$ itself. Let a block contain K elements say

$u^{(1)}, u^{(2)}, \ldots, u^{(k)}$

The k(k-1) elements obtained as $u^{(i)} - u^{(j)}$ (i, j = 1, 2, . . . k;

$i \neq j$) are called differences arising from a block. For example let there be 19 varieties corresponding to the elements of the modul of residue classes (mod. 19). Then there arise 12 differences from the block (o, 1, 2, 8), viz o − 1 = 18, o − 2 = 17, 0 − 8 = 11, 1 − o = 1, 1 − 2 = 18, 1 − 8 = 12, 2 — 0 = 2, 2 − 1 = 1, 2 − 8 = 13, 8 − o = 8, 8 − 1 = 7, 8 − 2 = 6.

Let n_l denote the number of varieties in the l^{th} block. Then if among the

$$n_1(n_1 - 1) + n_2(n_2 - 1) + \ldots + n_t(n_t - 1)$$

block differences arising from the t blocks, every element of M is repeated times, we say that in the t blocks the differences are symmetrically repeated, each occuring λ times.

For example consider the modul of residue classes (mod 9), and to every element of the modul let there correspond one variety. Then in the two blocks (o, 1, 2, 4) and (o, 3, 4, 7) the differences are symmetrically repeated. Since the first block gives rise to 12 differences 8, 7, 5, 1, 8, 6, 2, 1, 7, 4, 3, 2 and the second block gives rise to the 12 differences 6, 5, 2, 3, 8, 5, 4, 1, 6, 7, 4, 3; and among these 24 differences each of the elements 1, 2, 3, 4, 5, 6, 7, 8 occurs exactly thrice.

We now state a special case of fundamental theorem (Bose 1939) which gives a method of construction of BIBD given a set of initial difference sets.

Theorem 6.1. Let M be a modul containing the n elements

$$u^{(o)}, u^{(1)}, \ldots, u^{(n-1)}.$$

To any element $u^{(i)}$ let there correspond one variety. Let it be possible to find a set of t blocks $B_1, B_2, \ldots, B_t$ satisfying the following conditions:

(i) Every block contains exactly k varieties. (The varieties contained in the same block being different from one another.)

(ii) The differences arising from the t blocks are symmetrically repeated, each occuring λ times.

If θ be any element of M, then from each block B_l ($l = 1, 2, \ldots, t$), we can form another block $B_{l,\theta}$ by taking corresponding to every variety $u^{(i)}$ in B_1, the variety $u^{(j)}$ in B_1, where

$$u^{(j)} = u^{(i)} + \theta$$

then the n.t blocks $B_{1,\theta}$ $(1 = 1, 2, \ldots, t; \theta = u^{(o)}, \ldots, u^{(n-1)})$ provide us with a balanced incomplete block design with parameters

$$v = n, b = n.t., r = kt, k, \lambda.$$

Theorem 6.2. Let $((a_{1,1}), \ldots, (a_{1,k})); ((a_{2,1}), \ldots, (a_{2,k})); \ldots, ((a_{t,1}), \ldots, (a_{t,k}))$, where each (a_{ij}) is one of the numbers 1, 2, ..., v, form a set of t initial difference sets (mod = v) for the BIBD with v = v, b = tv, r = tk, k= k, $\lambda = 1$. Then the difference set

$((a_{1,1})_1, \ldots, (a_{1,k})_1; \ldots; (a_{t,1})_t, \ldots, (a_{t,r})_t)$ developed (mod v) keeping the lower suffixes fixed, gives the PBIBD dual to the above design.

Proof. The design obtained by developing the latter difference set will be dual to the BIBD given by the set of m difference sets if any two blocks of the dual design have just one variety in common. It is obvious that this condition is satisfied in virtue of the original set giving rise to the BIBD.

As an illustration consider the design v = 13, b = 26, r = 6, k = 3, $\lambda = 1$. In the two inittial blocks (1, 3, 9) and (2, 6, 5) the differences are symmetrically repeated each occuring one time. Thus the above design can be obtained by cyclically developing the two difference sets (1, 3, 9) and (2, 6, 5) mod 13. The set (1, 3, 9) gives rise to 13 blocks by adding the numbers o, 1, 2, ..., 12 respectively and reducing any number which exceeds 13 by subtracting 13 from it. Thus the blocks obtained by adding 10 to the difference set (1, 3, 9) gives (11, 13, 19) which when reduced mod 13 is (11, 13, 6). Hence the blocks obtained from the set (1, 3, 9) are (1, 3, 9), (2, 4, 10), (3, 5, 11), (4, 6, 12), (5, 7, 13), (6, 8, 1), (7, 9, 2), (8, 10, 3), (9, 11, 4), (10, 12, 5), (11, 13, 6), (12, 1, 7) and (13, 2, 8).

Similarly the set (2, 6, 5) gives rise to another set of 13 blocks as follows: (2, 6, 5), (3, 7, 6), (4, 8, 7), (5, 9, 8), (6, 10, 9) (7, 11, 10), (8, 12, 11), (9, 13, 12), (10, 1, 13), (11, 2, 1) (12, 3, 2), (13, 4, 3), (1, 5, 4).

These 26 blocks are the blocks of the above balanced incomplete block design.

The dual design is given by the difference set $(1_1, 3_1, 9_1; 2_2, 6_2, 5_2)$ which is to be developed cyclically mod 13, keeping the lower suffixes fixed.

Thus the 13 blocks obtained from this set are:

$(1_1, 3_1, 9_1; 2_2, 6_2, 5_2)$, $(2_1, 4_1, 10_1; 3_2, 7_2, 6_2)$
$(3_1, 5_1, 11_1; 4_2, 8_2, 7_2)$, $(4_1, 6_1, 12_1; 5_2, 9_2, 8_2)$
$(5_1, 7_1, 13_1; 6_2, 10_2, 9_2)$. $(6_1, 8_1, 1_1; 7_2, 11_2, 10_2)$
$(7_1, 9_1, 2_1; 8_2, 12_2, 11_2)$, $(8_1, 10_1, 3_1; 9_2, 13_2\ 12_2)$
$(9_1, 11_1, 41_1; 10_2, 1_2, 13_2)$, $(10_1, 12_1, 5_1; 11_2, 2_2, 1_2)$
$(11_1, 13_1, 6_1; 12_2, 3_2, 2_2)$, $(12_1, 1_1, 7_1; 13_2, 4_2, 3_2)$
and $(13_1, 2_1, 8_1; 1_2, 5_2, 4_2)$

We can identify the 26 different symbols occuring here with the 26 varieties of the dual design in the following manner.

Put $x_1 = x$ and $x_2 = 13 + x$ for $x = 1, 2, \ldots, 13$. Thus the symbol 9_1 stands for the variety numbered 9 whereas 9_2 stands for variety numbers 22. With this identification the above 13 sets give the 13 blocks of the dual design.

In general the problem of identification can be explained as follows.

If we have a difference set mod p in which there are q suffixes, we get in all pq different symbols after cyclic development. Put $x_y = x + (y-1)p$, $y = 1, 2, \ldots, q$.

Thus we have varieties numbered 1, 2, ..., pq corresponding to these different symbols.

7. The dual of BIBD with $r = p$, $k = p-2$, $\lambda = 2$.

Theorem 7.1. The dual of a BIBD with parameters

$$v = \frac{(p-1)(p-2)}{2}, b = \frac{p(p-1)}{2}, r = p, k = p-2, \lambda = 2$$

is a PBIBD with parameters

$$v = \frac{p(p-1)}{2},\ b = \frac{(p-1)(p-2)}{2},\ r = p - 2,\ k = p$$

$$\lambda_1 = 1,\ \lambda_2 = 2,\ n_1 = 2(p-2),\ n_2 = \frac{(p-2)(p-3)}{2}$$

$$\begin{bmatrix} p^{(1)}_{ij} \end{bmatrix} = \begin{bmatrix} p-2 & p-3 \\ p-3 & \frac{(p-3)(p-4)}{2} \end{bmatrix}$$

$$\begin{bmatrix} p^{(2)}_{ij} \end{bmatrix} = \begin{bmatrix} 4 & 2(p-4) \\ 2(p-4) & \frac{(p-4)(p-5)}{2} \end{bmatrix}$$

Proof: Let A be an initial block of the design. Since $\cdot = 2$, no other blocks of the design can have more than two varieties in common with A. Let m_o be the number of blocks A_o each of which has no variety in common with A, m_1 the number of blocks A_1 each of which has one variety in common with A and m_2 the number of blocks A_2 each of which has two varieties in common with it.

Then using the results of Lemma 5.1 we have the following relations.

$$m_o + m_1 + m_2 = b - 1 = \frac{p(p-1)}{2} - 1$$

$$m_1 + m_2 = k(r-1) = (p-2)(p-1)$$

$$2m_2 = k(k-1)(\lambda-1) = (p-2)(p-3)$$

From which we have

$$m_o = o,\ m_1 = 2(p-2),\ m_2 = \frac{(p-2)(p-3)}{2}$$

In the dual design therefore with respect to any given variety, the remaining varieties can be divided into two groups of size n_1 and n_2 respectively where

$$n_1 = 2(p-2),\ n_2 = \frac{(p-2)(p-3)}{2}$$

such that the varieties of the first group occur $\lambda_1(=1)$ time and those of the second group occur $\lambda_2(=2)$ times with the given variety.

Let an initial variety occur x_1 times in the set A_1 and x_2 times in the set A_2. There are $\frac{(p-2)(p-3)}{2}$ blocks containing as many pairs of the initial varieties. Every pair has occured once in A and so can occur at most once in A_2. Thus all the pairs of initial varieties in A_2 are distinct. This shows that every initial variety occurs p-3 times in A_2, once in A and therefore just twice in A_1 i.e. $x_1 = 2$ and $x_2 = p-3$.

Let a non-initial variety occur y_1 times in the set A_1 and y_2 times in the set A_2. Then

$$y_1 + y_2 = p$$

and remembering that it must occur twice with every initial variety, we get

$$y_1 + 2y_2 = 2(p-2).$$

Hence $y_1 = 4$, $y_2 = p-4$

Let B be a block of A_1. Let there be z_1 blocks and z_2 blocks B_2 in A_1. Then

$$z_1 + z_2 = m_1 - 1 = 2(p-2) -1$$

Now block B consists of one initial variety and (p-3) non-initial varieties with reference to the standard A. In the remaining 2(p-2) -1 blocks of A_1 this initial variety occurs once and the non-initial varieties occur thrice. So we have $z_1 + 2z_2 = 1 + 3(p-3)$.

Hence $z_1 = p-2$, $z_2 = p-3$

Let B' be a block of A_2. Let there be u_1 blocks B'_1 and u_2 blocks B'_2 in A_1. Then

$$u_1 + u_2 = m_1 = 2(p-2)$$

Again the block B' consists of 2 initial-varieties and p-4 non-initial varieties, with reference to the standard A. In the set of $m_1 = 2(p-2)$ blocks of A_1, the initial varieties occur twice and the non-initial varieties occur four times. So we have

$$u_1 + 2u_2 = 2x_2 + 4(p-4)$$

$$u_1 = 4, \; u_2 = 2(p-2)-4 = 2(p-4)$$

For the dual design the blocks A and B become 1-associates and the number of varieties which are 1-associates of the treatment corresponding to A and simultaneously 2-associates of the treatment corresponding to B is the number of blocks B_2 each of which has one variety in common with A and simultaneously 2 varieties in common with B. Thus

$$p^1_{12} = z_2 = p-3 = p^1_{21} .$$

Similarly the number of varieties which are simultaneously 1-associates of these two varieties of the dual design is obviously the number z_1 of blocks B_1 each of which has one variety in common with both A and B. Thus $p^1_{11} = z_1 = p-2$.

Further the blocks A and B′ become 2 associates and we easily see that

$$p^2_{12} = u_2 = 2(p-4) = p^2_{12}$$

$$p^2_{11} = u_1 = 4$$

$$p^2_{22} = \frac{(p-4)(p-5)}{2}$$

Thus the dual design is a PBIB with just two-types of associates.

Corollary 7.1. The above PBIBD can also be obtained by omitting all the blocks containing any particular variety from the SBIBD with parameters

$$v = b = \frac{p^2 - p + 2}{2}, \; r = k = p, \; \lambda = 2$$

Proof: Consider the residual design of the above PBIBD which will be a BIBD with parameters

$$v = \frac{(p-1)(p-2)}{2}, \; b = \frac{p(p-1)}{2}$$

$$r = p, \; k = p-2 \text{ and } \lambda = 2$$

The dual of the residual design can be obtained by omitting all the blocks containing any particular variety. (This variety corresponds to the block omitted in obtaining residual design.) By ϕ above theorem then this dual is a PBIBD of Theorem 7.1.

8. Some More Results on Duals:

In this section we give some more results concerning duals. The reader is referred to Agrawal (1963).

Lemma 8.1. If N is the incidence matrix of a PBIBD with parameters v, b, r, k, λ_i n_i, (p^i_{jk}), i, j, k = 1; 2, ..., m and E_2 is a v x b matrix with all elements equal to unity then $N^* = E_{vb} - N$ is the incidence matrix of a PBIBD with parameters

$$v^* = v,\ b^* = b,\ r^* = b\text{-}r,\ k^* = v\text{-}k$$

$$\lambda^*_i = b\text{-}2r + \lambda_i,\ n^*_1 = n_i,\ (p^{i*}_{jk}) = (p^i_{jk})$$

Proof: Consider $N^*N^{*\prime} = (E_{vb} - N)\ (E_{vb} - N)'$

$$= (E_{vb} - N)(E_{vb} - N')$$

$$= bE_{vv} - E_{vb}N' - N_{vb}E_{bv} + NN'$$

$$= bE_{vv} - E_{vb}N' - N_{vb}E_{bv} + NN'$$

$$= bE_{vv} - rE_{vv} - rE_{vv} + NN'$$

$$= (b - 2r)\ E_{vv} + NN'$$

The matrix NN′ consists of r in the principal diagonal and off-diagonal elements are one of the numbers $_i$'s.

Now remembering that matrix $N^*N^{*\prime}$ will consist of r^* in the principal diagonal and off-diagonal elements will be one of the numbers $_i{}^*$'s, we see on equating diagonal and off-diagonal elements from both sides that

$$r^* = b\text{-}2r + y = b\text{-}r$$

$$\lambda^*_i = b\text{-}2r + \lambda_i$$

The other relationships follow immediately.

Theorem 8.1. If (i) N is the incidence matrix of a SBIBD with parameters v, k, and L_1 is the incidence matrix of its derived design (ii) L'_1 is the incidence matrix of a PBIBD with parameters

$$v^* = v\text{-}1,\ b^* = k,\ r^* = \lambda,\ k^* = k\text{-}1$$

$$\lambda_i^*, n_i^*, (p_{jk}^{i*}),\ i, j, k = 1, 2, \ldots, m$$

then the dual of the residual design N_1 of N is a PBIBD with parameters

$$v^{**} = v\text{-}1,\ b^* = v\text{-}k,\ r^{**} = k\text{-}\lambda,\ k^{**} = k$$

$$\lambda_i^{**} = \lambda\text{-}\lambda_i,\ n_i^{**} = n_i^*,\ (p_{jk}^{i**}) = (p_{jk}^{i*})\ i, j, k$$

$$= 1, 2, \ldots, m.$$

Proof: Remembering the definition of L_1 and N_1, we can, without any loss of generality represent N by

$$N = \begin{bmatrix} E_{k,1} & \vdots & L_{1\,k,v-1} \\ \cdots & & \cdots \\ {}_{v-k,1} & \vdots & N_{1\,v-k,\,v-1} \end{bmatrix}$$

Then

$$N'N = \begin{bmatrix} E_{1,k} & \vdots & L_{1\,1,v-k} \\ \cdots & & \cdots \\ L'_{1\,v-1,k} & \vdots & N'_{1\,v-1,v-k} \end{bmatrix} \begin{bmatrix} E_{k,1} & \vdots & L_{1\,v-k,1} \\ \cdots & & \cdots \\ L_{v,k,1} & \vdots & N_{1\,k,v-1} \end{bmatrix}$$

$$= \begin{bmatrix} k\ E_{11} & \vdots & E_{1,v-1} \\ \cdots & & \cdots \\ E_{v-1,1} & \vdots & \underset{v-1,\ v-1}{L'_1 L_1} + \underset{v-1,\ v-1}{N'_1 N_1} \end{bmatrix}$$

Now N'N consists of k in the principal diagonal and off-diagonal elements consist of λ.

i.e. $N'N = (k-\lambda) I_v + \lambda E_{vv}$

$N_1' N_1 + L_1' L_1 = (k-\lambda) I_{v-1} + \lambda E_{v-1, v-1}$

i.e. $N_1' N_1 = (k-\lambda) I_{v-1} + \lambda E_{v-1, v-1} - L_1' L_1$

$L_1' L_1 = L_1' (L_1')'$ consists of r^* – in the principal diagonal and off-diagonal elements consist of one of the numbers λ_i^* , and matrix $N_1' N_1$ will consist of r^{**} in the principal diagonal and off-diagonal element will be one of the numbers λ_i^{**}.

Now equating diagonal and off-diagonal elements from both sides we get

$r^{**} = k - \lambda + \lambda - \lambda = k - \lambda$

$\lambda_i^{**} = \lambda - \lambda_i^*$.

The other relationships follow immediately.

Corollary 8.1. It follows from Lemma 8.1 that duals of the complementary designs $(E_{k,v-1} - L_1)$ and $(E_{v-k,v-1} - N_1)$ are PBIB designs with the same association scheme.

Corollary 8.2. If the SBIBD N has $\lambda = 2$, then the derived BIBD has $\lambda = 1$. The dual of the BIBD with $\lambda = 1$ has been proved to be a PBIBD (Theorem 5.1). Hence the duals of N_1, $(E_{k,v-1} - L_1)$ and $(E_{v-k,v-1} - N_1)$ are PBIBD with the same association scheme.

Corollary 8.3. If the SBIBD has $\lambda = 1$, the the residual design has $\lambda = 1$ and its dual is a PBIBD (Theorem 5.1) Applying Lemma 8.1 and Theorem 8.1 we see that duals of L_1, $(E_{k,v-1} - L_1)$ and $(E_{v-k,v-1})$ are PBIBD with the same association scheme.

Examples. (i) Consider the orthogonal series of BIB designs with parameters $v = s^2 + s + 1 = b$, $r = s + 1 = k$,

$\lambda = 1$. As the residual design is a BIBD with parameters

$$v = s^2, b = s^2 = s, r = s + 1, k = s, \lambda = 1,$$

so its dual is a PBIBD (Theorem 5.1).

Applying Lemma 8.1 and Theorem 8.1 the duals of the derived design, its complementary design and the complementary design of the residual design are PBIBD's with the same association scheme.

(ii) Consider the SBIBD with parameters

$$v = b = p^2 - p + 2 \quad , r = k = p \text{ and } \lambda = 2$$

The derived design of the above series is a BIBD with $\lambda = 1$, hence the duals of the residual design, the derived design, and the duals of the complementary of the residual and derived BIBD's are again PBIBD's. The residual design is

$$v^* = \frac{(p-1)(p-2}{2}, \quad b^* = \frac{p(p-1),}{2} \; r^* = p$$

$$k^* = p-2 \text{ and } \lambda^* = 2$$

compare with Corollary 7.1).

SUMMARY

In this paper it is shown that the Balanced Incomplete Block Design with parameters

(i) $v, b, r, k, \lambda = 1$

and

(ii) $v = \frac{(p-1)(p-2),}{2} \; b = \frac{p(p-1)}{2}$

$r = p, k = p-2, \lambda = 2$

can be dualised to give Partially Balanced Incomplete Block Designs with only two types of associates. Easy methods for constructing these designs and some more results concerning such duals are also given.

ACKNOWLEDGMENT

Thanks go to Professor Morris L. Eaton for his advice on the preparation of the Master paper on which the present paper is based.

REFERENCES

1. Agrawal, H. (1963), "On the duals of BIB designs." Calcutta Statistical Association Bulletin, 12, 104-105.

2. Bose, R. C. (1939). "On the construction of balanced incomplete block designs." Annals of Eugenics 9, 353-399.

3. Bose, R. C. and Nair, K. R. (1939). "Partially balanced incomplete block designs." Sankhyā 4, 337-373

4. Chakrabarti, M. C. (1962). Mathematics of design and analysis of experiments, Asia Publishing House, Bombay.

5. Fisher, R. A. and Yates, F. (1949). Statistical Tables Hafner Publishing Company, New York.

6. Hussain, Q. M. (1948). "Structure of some incomplete block designs." Sankhyā 8, 381-383.

7. Mann, H. B. (1949). Analysis and design of experiments, Dover Publications, Inc., New York.

The Philippine Statistician
Vol. XIX, Nos. 3 and 4, 1970

A SOCIO-ECONOMIC PROFILE OF THE PHILIPPINE ELECTORATE

By

CRISTINA P. PAREL*

The data on which this paper is based were obtained from the Opinion Poll Survey of the Philippine electorate conducted by the U.P. Statistical Center before the November, 1969 national elections. Such surveys were conducted by the Center mainly to: (1) ascertain which sampling scheme is best suited for a survey of Philippine voters; and (2) determine what method is most efficient in gathering information on voting preferences of the electorate. However, some questions extarneous to the main objectives were included in the last survey to elicit more information about them; viz. some socio-economic characteristics. This paper gives a summary of these characteristics and makes an attempt to determine and analyze their relationships.

The last opinion poll survey used a two-stage sampling scheme with an over-all sampling fraction of 1/500. The lists of electoral precincts of the Commission on Elections were used as a frame.

The mailed questionnaire was used in gathering information from the chosen respondents.

I. *Demographic Characteristics*

A. *Sex*

There are more male voters than female voters. Only a little more than 40 per cent are female voters in both the

* Director and Professor of Statistics, U.P. Statistical Center.

Manila and non-Manila areas. This could be an indication that the males are still more interested in politics than the female; that the women still consider politics as the field of the men, and that it should be left only to the men. Some housewives feel that their husbands can very well take care of representing the family in political affairs without their participation.

TABLE I. SEX DISTRIBUTION OF PHILIPPINE REGISTERED VOTERS (NOVEMBER, 1969) (IN PER CENT)

SEX	Philippines	Greater Manila Area	Non-Manila Area
Total	100.0	100.0	100.0
Male	56.0	54.2	56.4
Female	43.2	41.7	43.6
Not Reported*	0.8	4.1	–

* Those who returned questionnaires but did not answer the question.

B. *Age*

The voters in the Greater Manila area seem to be younger than those in the non-Manila area. About 30 per cent of the voters in the Manila area were below 30 years old in the last national elections, while only about 24 per cent are below this age in the non-Manila area. Perhaps, one reason for this is that the elections fall within a school year and many students of voting ages do not go home to their respective provinces to cast their votes. A great number also of those working in the Manila area from the provinces do not go back to their respective provinces to vote. Many of those who stay, register and cast their votes in the Manila area.

There is slightly a higher proportion of voters above 50 years old in the Greater Manila area than in the non-Manila area (19 per cent versus 17 per cent.).

As observed in Table 2, information on age is still one of the hardest to obtain from the respondents. More than 10 per cent of those who returned their accomplished questionnaires did not give their ages.

TABLE. 2. AGE DISTRIBUTION OF PHILIPPINE VOTERS (NOVEMBER, 1969) (In Per Cent)

AGE GROUP	Philippines	Greater Manila Area	Non-Manila Area
Total	99.9*	99.8*	99.9*
21 - 24	9.6	12.1	9.0
25 - 29	15.1	16.8	14.7
30 - 34	13.5	15.3	13.1
35 - 39	13.1	13.3	13.0
40 - 44	11.5	8.9	12.1
45 - 49	9.3	8.9	9.4
50 & Above	17.4	18.6	17.2
Not Reported	10.4	5.9	11.4

* The total is not exactly 100.0% due to rounding.

C. *Civil Status*

About seven out of ten of the Philippine electors are married (70.7 per cent.). Only about 15 per cent are single.

TABLE 3. DISTRIBUTION OF PHILIPPINE VOTERS BY CIVIL STATUS (NOVEMBER, 1969) (In Per Cent)

CIVIL STATUS	Philippines	Greater Manila Area	Non-Manila Area
Total	99.9*	99.9*	100.0
Single	16.7	23.6	15.1
Married	70.7	67.9	71.4
Widowed	3.4	3.2	3.4
Separated	1.0	0.4	1.2
Not Reported	8.1	4.8	8.9

* The total does not add up to 100.0% due to rounding.

There is a higher proportion of married voters in the non-Manila area than in the Manila area (71.4% versus 67.9%.).

D. *Educational Level*

More than half of the Philippine electorate have only high school education and below. About 34 per cent have reached and gone beyond, college. In the non-Manila area about 34 per cent have only "intermediate and below" education and only about 30 per cent have gone to college. In the Greater Manila area, only about 14 per cent have "intermediate and below" education and more than 50 per cent have gone to college. This low level of education of some of non-Manila voters may be a reason for the "emotional" following of "attractive" and "personable" candidates. The high proportion of voters who have gone to college in the Greater Manila area is possibly one explanation why it has been observed in many past elections that popular candidates in the non-Manila area do not generally appeal to the Greater Manila voters, and that voters generally are not easily swayed by promises and the grandiose eloquence of some candidates.

TABLE 4. DISTRIBUTION OF PHILIPPINE VOTERS BY EDUCATIONAL LEVEL (NOVEMBER, 1969)
(In Per Cent)

EDUCATIONAL LEVEL	Philippines	Greater Manila Area	Non-Manila Area
Total	100.1*	100.1*	100.0
Primary	7.8	1.7	9.2
Intermediate	22.7	12.3	25.1
High School	26.8	26.8	26.8
College	27.3	42.1	23.8
Post-College	6.6	10.2	5.8
Note Reported	8.9	7.0	9.3

* The total is not exactly 100.0% due to rounding.

II. A. *Employment*

Out of the sampled voters who returned their questionnaires, about one-third (33 per cent) are unemployed. More than one-third of those who returned their accomplished questionnaires in the non-Manila areas are unemployed; the proportion is much less in the Greater Manila area. About one out of five reported that they are self-employed. There is a higher proportion of government employees than private agencies-employees among the voters (16.4% versus 11.7%). However, in the Greater Manila area, there are many more private agencies-employees than government employees (24.8% versus 14.4%.).

More than 10 per cent of those who returned their questionaires did not give the information on employment status.

TABLE 5. PER CENT DISTRIBUTION OF PHILIPPINE VOTERS (NOVEMBER, 1969)

EMPLOYMENT STATUS	Philippines	Greater Manila Area	Non-Manila Area
Total	100.0	99.9*	100.0
Gov. Employees	16.4	14.4	16.9
Priv. Employees	11.7	24.8	8.6
Self-Employees	20.7	18.9	21.1
Pensioners	5.4	9.5	4.5
Unemployed	33.0	22.1	35.5
Not Reported	12.8	10.7	13.4

* The total is not exactly 100.0% due to rounding.

B. *Income*

Like information on age, information on income is very difficult to obtain from the respondents. This is shown by the more than 30 per cent of non-response to the question on income among those who returned their questionnaires.

About 46 per cent of the Philippine voters who returned their questionnaires reported an income of less than ₱3,600 per annum; about 48 per cent in the non-Manila areas and about 38 per cent in the Greater Manila area. Only about 7% reported an income of above ₱5,000 per annum. However, in the Greater Manila area, about 16 per cent reported an income of above ₱5,000 per annum. Voters in the non-Manila area with the same income constitute only about 5 per cent.

TABLE 6. PER CENT DISTRIBUTION OF PHILIPPINE VOTERS BY INCOME (NOVEMBER, 1969)

Income per year	Philippines	Greater Manila Area	Non-Manila Area
Total	100.0	100.0	100.0
Below ₱2,00	25.5	16.2	31.4
₱2,000 to less than ₱3,600	17.3	21.5	16.4
₱3,600 to less than ₱5,000	7.8	14.4	6.3
Above ₱5,000	7.4	15.8	5.3
Not Reported	39.0	32.1	40.6

III. *Analysis of Income in Relation to Other Characteristics*

Income distribution is a very essential factor in social stratification. To understand more fully income differences of the Philippine voters, association studies are made in this paper.

A. *Income and Sex*

The male members of society are generally considered as the stronger of the two sexes. In a household, the male members are looked up to as the "breadwinners" and the "hope" of the family. But from the information gathered from the last opinion poll of Philippine voters, there are no evidences that income is associated with sex. The men voters seem not to be earning more than their female counterparts.

TABLE 7-A. INCOME BY SEX FOR 21 YEARS OLD AND OVER, FOR NINE REGIONS

SEX	Income Level				Total
	Below ₱2,000	₱2,000 – Less ₱3,600	₱3,600 – ₱5,000	Above ₱5,000	
Male	526	273	109	105	1,013
Female	370	193	70	46	679
Total	896	466	179	151	1,692

$\chi^2 = 6.78$ (not significant at $a = .05$)

TABLE 7-B. INCOME BY SEX FOR 21 YEARS OLD AND OVER FOR THE FIRST REGION

SEX	Income Level				Total
	₱2,000 Below	₱2,000 – Less ₱3,600	₱3,600 – ₱5,000	₱5,000 Above	
Male	56	96	54	66	272
Femade	46	43	32	31	152
Total	102	139	86	97	424

$\chi^2 = 5.95$ (significant at $a = .05$)

TABLE 7-C. INCOME BY SEX FOR 21 YEARS OLD AND OVER FOR ALL TEN REGIONS

SEX	Income Level				Total
	Below ₱2,000	₱2,000 – Less ₱3,600	₱3,600 – ₱5,000	Above ₱5,000	
Male	582	369	163	171	1,285
Female	416	236	102	77	831
Total	998	605	265	248	2,116

$\chi^2 = 9.55$ (Not significant at $a = .05$)

B. *Income and Education*

Observations seem to indicate that persons with more schooling tend to earn more money than those with less schooling. More schooling particularly, at the secondary and college levels, generally improves the productivity of an individual. This observation seems tto be confirmed in this study. Among the Philippine voters, as one goes higher on the "educational ladder", the income becomes bigger. However, there are some with relatively little education with higher income, and a number with relatively higher education with low income. This is possibly due to differences in abilities, motivations and efforts of the individuals and many other factors that impinge on the observed association between income and education.

TABLE 8-A. INCOME BY EDUCATIONAL ATTAINMENT FOR 21 YEARS OLD AND OVER, FOR ALL TEN REGIONS

EDNCATIONAL ATTAINMENT	Income Level								Total	
	Below ₱2,000		₱2,000 - Less ₱3,600		₱3,600 – ₱5,000		Above ₱5,000			
	No.	%	No.	%	No.	%	No.	%	No.	%
Intermediate and below	443	77.99	75	13.20	22	3.87	28	4.93	568	100.0
High School	334	59.12	144	25.49	54	9.56	33	5.84	565	100.0
College	167	23.49	274	38.54	152	21.38	118	16.60	711	100.0
Post College	17	8.54	83	41.71	39	19.60	60	30.15	199	100.0
Total	961		576		267		239		2,043	

$x^2 = 563.82$ (Not significant at $\alpha = .05$)

TABLE 8-B. INCOME BY EDUCATIONAL ATTAINMENT FOR 21 YEARS OLD AND OVER, FOR NINE REGIONS EXCLUDING THE FIRST REGION

EDUCATIONAL ATTAINMET	Income Level								Total	
	Below ₱2,000		₱2,000 – Less ₱3,600		₱3,600 – ₱5,000		Above ₱5,000			
	No.	%	No.	%	No.	%	No.	%	No.	%
Primary	121	83.95	13	8.96	2	1.38	9	6.21	145	100.0
Intermediate	304	79.58	52	13.61	14	3.66	12	3.14	382	100.0
High School	293	64.53	104	22.90	37	8.14	20	4.40	454	100.0
College	128	25.60	204	40.80	99	19.80	69	13.80	500	100.0
Post-College	13	9.35	72	51.79	24	17.26	30	21.58	139	100.0
Total	859		445		176		140		1,620	

$x^2 = 460.33223$ (Not significant at $a = .05$)

TABLE 8-C. INCOME BY EDUCATIONAL ATTAINMENT FOR 21 YEARS OLD AND OVER FOR THE FIRST REGION

EDUCATIONAL ATTAINMENT	Income Level								Total	
	Below ₱2,000		₱2,000 – Less ₱3,600		₱3,600 – ₱5,000		Above ₱5,000			
	No.	%	No.	%	No.	%	No.	%	No.	%
Intermediate and below	18	43.90	10	24.39	6	14.63	7	17.07	41	100.0
High School	41	36.94	40	36.04	17	15.22	13	11.71	111	100.0
College	39	18.48	70	33.18	53	25.12	49	23.22	211	100.0
Post-College	4	6.67	11	18.33	15	25.00	30	50.00	60	100.0
Total	102		131		91		99		423	

$x^2 = 59.32155$ (Not significant at $a = .05$)

TABLE 9-A. INCOME BY AGE FOR 21 YEARS OLD AND OVER (PHILIPPINES) (NOVEMBER, 1969)

AGE GROUP	Income Level								Total	
	Below ₱2,000		₱2,000 – Less ₱3,600		₱3,600 – ₱5,000		Above ₱5,000			
	No.	%	No.	%	No.	%	No.	%	No.	%
21 – 24	97	55.11	52	29.54	18	10.23	9	5.11	176	100.0
25 – 29	136	42.24	115	35.71	36	11.18	35	10.87	322	100.0
30 – 34	154	49.84	83	26.86	42	13.59	30	9.71	309	100.0
35 – 39	148	48.84	83	27.39	38	12.54	34	11.22	303	100.0
40 – 44	131	44.71	91	31.06	34	11.60	37	12.63	293	100.0
45 – 49	97	41.99	77	33.33	28	12.12	29	12.55	231	100.0
50 & Above	211	47.95	89	20.23	69	15.68	71	16.14	440	100.0
Total	974		590		265		245	.	2,074	

$x^2 = 46.74$ (Not significant at $a = .05$)

TABLE 9-B. INCOME BY AGE FOR 21 YEARS OLD AND OVER (NON-MANILA AREAS) (NOVEMBER, 1969)

AGE GROUP	Income Level				Total
	Below ₱2,000	₱2,000 – Less ₱3,600	₱3,600 – ₱5,000	Above ₱5,000	
21 – 24	81	31	6	5	123
25 – 29	119	85	24	21	249
30 – 34	130	66	26	13	235
35 – 39	137	61	25	21	244
40 – 44	118	80	30	25	253
45 – 49	91	58	21	16	186
50 & Above	192	72	39	41	344
Total	868	453	171	142	1,634

$x^2 = 38.35$ (Not significant at $a = .05$)

TABLE 9-C. INCOME BY AGE FOR 21 YEARS OLD AND OVER (GREATER MANILA AREA) (NOVEMBER, 1969)

AGE GROUP	Income Level				Total
	Below ₱2,000	₱2,000 - Less ₱3,600	₱3,600 – ₱5,000	Above ₱5,000	
21 – 24	16	21	12	4	53
25 – 29	17	30	12	14	73
30 – 34	24	17	16	17	74
35 – 39	11	22	13	13	59
40 – 44	13	11	4	12	40
45 – 49	6	19	7	13	45
50 & Above	19	17	30	30	96
Total	106	137	94	103	440

$x^2 = 39.68$ (Not significant at $\alpha = .05$)

TABLE 10-A. INCOME BY CIVIL STAATUS FOR 21 YEARS OLD AND OVER (NON-MANILA AREA) (NOVEMBER, 1969)

CIVIL STATUS	Income Level				Total
	Below ₱2,000	₱2,000 - Less ₱3,600	₱3,600 – ₱5,000	Above ₱5,000	
Single	153	97	27	15	292
Married	698	356	140	128	1,322
Total	851	453	167	143	1,614

$x^2 = 9.37$ (Not significant at $a = .05$)

TABLE 10-B. INCOME BY CIVIL STATUS FOR 21 YEARS OLD AND OVER (MANILA AREA) (NOVEMBER, 1969)

CIVIL STATUS	Income Level				Total
	Below ₱2,000	₱2,000 - Less ₱3,600	₱3,600 – ₱5,000	Above ₱5,000	
Single	25	46	21	20	112
Married	75	85	68	78	306
Widow/Separated	2	7	4	1	14
Total	102	138	93	99	432

$x^2 = 11.79$ (Not significant at $a = .05$)

C. *Income and Age*

Analysis of the information gathered from Philippine voters seem to indicate that income is associated with age; as one grows older, the income tends to increase. However, if the age groups are examined by educational level, there seems to be no association between age and income; that is, if level of education is "controlled", no association is indicated. It is possible that the proportions of registered voters relative to the population in the age groups vary considerably and that the proportion is higher in older age groups than in the lower age groups.

D. *Income and Civil Status*

It is a general impression that married people have more motivation and incentive to earn more than the unmarried ones. But the results of the poll show that income does not seem to be associated with civil status. However, among the Philippine electorate, there is a higher proportion of married voters earning ₱5,000 and above per annum than among the unmarried ones.

TABLE 11-A. PER CENT DISTRIBUTION OF PHILIPPINE VOTERS BY INCOME, AGE AND EDUCATIONAL ATTAINMENT (NOVEMBER, 1969)

AGE AND EDUCATIONAL ATTAINMENT	Total		Below ₱2,000		₱2,000 – Less ₱3,600		₱3,600 – ₱5,000		Above ₱5,000	
	No.	%	No.	%	No.	%	No.	%	No.	%
Intermediate										
21 – 24	28	5.1	23	5.3	2	2.8	2	9.1	1	4.0
25 – 29	55	10.0	43	10.0	8	11.4	2	9.1	2	8.0
30 – 34	65	11.8	59	13.6	6	8.6	—	—	—	—
35 – 39	73	13.3	62	14.4	7	10.0	2	9.1	2	8.0
40 – 44	81	14.8	56	13.0	14	20.0	4	18.2	7	28.0
45 – 49	83	15.1	63	14.6	11	15.7	5	22.7	4	16.0
50 & Above	164	29.9	126	29.2	22	31.4	7	31.8	9	36.0
Subtotal —	549	100.0	432	100.1	70	99.9	22	100.0	25	100.0
High School										
21 – 24	42	7.7	33	10.2	8	5.9	—	—	1	3.1
25 – 29	71	13.0	46	14.2	20	14.7	3	5.7	2	6.2
30 – 34	88	16.1	56	17.2	20	14.7	7	13.2	5	15.6
35 – 39	94	17.2	57	17.5	22	16.2	10	18.9	5	15.6
40 – 44	81	14.8	48	14.8	18	13.2	8	15.1	7	21.9
45 – 49	58	10.6	28	8.6	19	14.0	7	13.2	4	12.5
50 & Above	112	20.5	57	17.5	29	21.3	18	34.0	8	25.0
Subtotal	546	99.9	325	100.0	136	100.0	53	100.1	32	99.9

TABLE 11-A. PER CENT DISTRIBUTION OF PHILIPPINE VOTERS BY INCOME, AGE AN EDUCATIONAL .... (CONT'D)

AGE AND EDUCATIONAL ATTAINMENT	Total		Below ₱2,000		Less ₱3,600 ₱2,000 –		₱3,600 – ₱5,000		₱5,000 Above	
	No.	%	No.	%	No.	%	No.	%	No.	%
College & Post College										
21 – 24	99	11.1	37	20.4	41	11.7	15	8.1	6	3.4
25 – 29	187	20.9	38	21.0	81	23.1	32	17.2	36	20.3
30 – 34	141	15.8	34	18.8	49	18.0	34	18.3	24	13.6
35 – 39	124	13.8	23	12.7	51	14.5	25	13.4	25	14.1
40 – 44	121	13.5	21	11.6	54	15.4	24	12.9	22	12.4
45 – 49	81	9.0	4	2.2	44	12.5	16	8.6	17	9.6
50 & Above	142	15.9	24	13.3	31	8.8	40	21.5	47	26.6
Subtotal —	895	100.0	181	100.0	351	100.0	186	100.0	177	100.0
Total	1,990		938		557		261		234	

Non-response $= 1{,}522;\ \frac{1{,}522}{3{,}512} = 43.3\%$

TABLE 12-B. DISTRIBUTION OF GREATER MANILA VOTERS BY INCOME, AGE, AND EDUCATIONAL ATTAINMENT (NOVEMBER, 1969)

AGE AND EDUCATIONAL ATTAINMENT	Total		Below ₱2,000		₱2,000 – Less ₱3,600		₱3,600 – ₱5,000		₱5,000 Above	
	No.	%	No.	%	No.	%	No.	%	No.	%
Intermediate										
21 – 24	2	5.1	—	—	—	—	1	20.00	1	16.7
25 – 29	1	2.6	1	5.6	—	—	—	—	—	—
30 – 34	6	15.4	5	29.8	1	10.0	—	—	—	—
35 – 39	1	2.6	—	—	1	10.0	—	—	—	—
40 – 44	5	12.8	1	5.6	2	20.0	1	20.0	1	16.7
45 – 49	10	25.6	4	22.2	4	40.0	1	20.0	1	16.7
50 & Above	14	35.9	7	38.9	2	20.0	2	40.0	3	50.0
Subtotal —	39	100.0	18	100.1	10	100.0	5	100.0	6	100.1
High School										
21 – 24	8	7.3	5	11.9	3	7.9	—	—	—	—
25 – 29	13	11.8	7	16.7	6	15.8	—	—	—	—
30 – 34	18	16.4	11	26.2	3	7.9	2	11.8	2	15.4
35 – 39	18	16.4	6	14.3	8	21.0	3	17.6	1	7.7
40 – 44	10	9.1	5	11.9	2	5.3	1	5.9	2	15.4
45 – 49	14	12.7	1	2.9	7	18.4	2	11.8	4	30.8
50 & Above	29	26.4	7	16.7	9	23.7	9	52.9	4	30.8
Subtotal —	110	100.1	42	100.1	38	100.0	17	100.0	13	100 1

TABLE 12-B. DISTRIBUTION OF GREATER MANILA VOTERS BY INCOME, AGE, AND EDUCATIONAL ATTAINMENT (NOVEMBER, 1969) (continued)

AGE AND EDUCATIONAL ATTAINMENT	Total		Below ₱2,000		₱2,000 – Less ₱3,600		₱3,600 – ₱5,000		Above ₱5,000	
	No.	%	No.	%	No.	%	No.	%	No.	%
College & Post College										
21 – 24	43	16.0	11	26.2	18	22.5	11	16.2	3	3.8
25 – 29	52	19.3	8	19.0	20	25.0	11	16.2	13	16.4
30 – 34	48	17.8	7	16.7	13	16.2	14	20.6	14	17.7
35 – 39	35	13.0	5	11.9	10	12.5	9	13.2	11	13.9
40 – 44	24	8.9	7	16.7	7	8.8	2	2.9	8	10.1
45 – 49	17	6.3	1	2.4	6	7.5	3	4.4	7	8.9
50 & Above	50	18.6	3	7.1	6	7.5	18	26.5	23	29.1
Subtotal —	269	99.9	42	100.1	80	100.0	68	100.0	79	99.9
Total	418		102		128		90		98	

Non-response = 242; percentage of total = 36.7%

TABLE 13-C. PER CENT DISTRIBUTION OF NON-MANILA VOTERS BY INCOME, AGE, AND EDUCATIONAL ATTAINMENT (NOVEMBER, 1969)

AGE AND EDUCATIONAL ATTAINMENT	Total		Below ₱2,000		₱2,000 – Less ₱3,600		₱3,600 – ₱5,000		₱5,000 Above	
	No.	%	No.	%	No.	%	No.	%	No.	%
Intermediate										
21 – 24	26	5.1	23	5.6	2	3.3	1	5.9	—	—
25 – 29	54	10.6	42	10.1	8	13.3	2	11.8	2	10.5
30 – 34	59	11.6	54	13.0	5	8.3	—	—	—	—
35 – 39	72	14.1	62	15.0	6	10.0	2	11.8	2	10.5
40 – 44	76	14.9	55	13.3	12	20.0	3	17.6	6	31.6
45 – 49	73	14.3	59	14.2	7	11.7	4	23.5	3	15.8
50 & Above	150	29.4	119	28.7	20	33.3	5	29.4	6	31.6
Subtotal —	510	100.1*	414	99.9*	60	99.9*	17	100.0	19	100.0
High School										
21 – 24	34	7.8	28	9.9	5	5.1	—	—	1	5.3
25 – 29	58	13.3	39	13.8	14	14.3	3	8.3	2	10.5
30 – 34	70	16.1	45	15.9	17	17.3	5	13.9	3	15.8
35 – 39	76	17.4	51	18.0	14	14.3	7	19.4	4	21.0
40 – 44	71	16.3	43	15.2	16	16.3	7	19.4	5	26.3
45 – 49	44	10.1	27	9.5	12	12.2	5	13.9	—	—
50 & Above	83	19.0	50	17.7	20	20.4	9	25.0	4	21.0
Subtotal —	436	100.0	283	100.0	98	99.9*	36	99.9*	19	99.0*

TABLE 13-C. PER CENT DISTRIBUTION OF NON-MANILA VOTERS BY INCOME, AGE, AND ... (CONTINUED)

AGE AND EDUCATIONAL ATTAINMENT	Total		Below ₱2,000		₱2,000 – Less ₱3,600		₱3,600 – 5,000		₱5,000 Above	
	No.	%	No.	%	No.	%	No.	%	No.	%
College & Post College										
21 – 24	56	8.9	26	18.7	23	8.5	4	3.4	3	3.1
25 – 29	135	21.6	30	21.6	61	22.5	21	17.8	23	23.5
30 – 34	93	14.9	27	19.4	36	13.3	20	16.9	10	10.2
35 – 39	89	14.2	18	12.9	41	15.1	16	13.6	14	14.3
40 – 44	97	15.5	14	10.1	47	17.3	22	18.6	14	14.3
45 – 49	64	10.2	3	2.2	38	14.0	13	11.0	10	10.2
50 & Above	92	14.7	21	15.1	25	9.2	22	18.6	24	24.5
Subtotal —	626	100.0	139	100.0	271	99.9*	118	99.9*	98	100.1*
Total	1,572**		836		429		171		136	

* *The total is not exactly 100.0% due to rounding.*

** *Not all of the returned questionnaires had complete responses on income, age and educational level; hence, the total shown here is only about 50 per cent of those who returned the questionnaires.*

IV. *Limitations of the Findings*

Much care was exercised at all stages of the survey, from planning thru data collection up to the processing and tabulation stages, but as it is true in all surveys, "errors" arising from the respondents could not be avoided. Out of the about 40,000 questionnaires mailed out in Manila and the provinces, only about 18 per cent were returned accomplished, and out of those returned, still a number of them did not have complete responses. In view of this low returns, caution should be taken in interpreting the findings, especially since the survey was not designed primarily to measure these socio-economic characteristics.

However, after tabulating the returns, it appeared that those who accomplished and returned the questionnaires came from all sectors of the population, as to educational level, employment status, age and sex. It could not be ascertained, though, whether or not a sector was not over or under represented relative to the other sectors. In making this report, therefore, it is a assumed that the returns constitute a random subsample of the original sample.

The Philippine Statistician
Vol. XIX, Nos. 3 and 4, 1970

1970 DIRECTORY OF MEMBERSHIP

BOARD OF DIRECTORS
1970

CESAR M. LORENZO
President

WALFREDO F. FOJAS
First Vice President

AIDA R. SALDUA
Second Vice President

ANGELES R. BUENAVENTURA
Secretary-Treasurer

MEMBERS

MERCEDES B. CONCEPCION
TALA P. LAZATIN
TITO A. MIJARES
CRISTINA P. PAREL
SIMEON R. VENTURA

EX-OFFICIO MEMBERS

BERNARDINO G. BANTEGUI
PAZ B. CULABUTAN
MANUEL O. HIZON
BURTON T. OÑATE
BERNARDINO A. PEREZ
EXEQUIEL S. SEVILLA
ENRIQUE T. VIRATA

LIFE MEMBERS

MR. POLICARPIO AROMIN
Bauang, La Union

MISS SOPHYA M. BALICKA
USOM/Karachi
U.S. Embassy
Karachi, Pakistan

MISS CONSUELO CRUZ BANAG
869 Teresa, Ermita
Manila D-406

DR. AMADO A. CASTRO
School of Economics
University of the Philippines
Diliman, Quezon City D-505

MISS PAZ B. CULABUTAN
EDP Center
Central Bank of the Philippines
NCR Building, UN Avenue
Manila

DR. MEREDITH B. GIVENS
Milbrook Heights
314 Hill Drive, State College
Pennsylvania U.S.A.

ATTY. LEON MA. GONZALES
630 S. Kingsley Drive
Los Angeles, California 90005
U.S.A.

MRS. LUZ S. HERNANDEZ
Department of Labor
Phoenix Building
Recoletos St.
Manila D-406

MR. MAX LACROIX
96, Rue D'arlon
Bruxelles 4
Belgium

DR. BENITO LEGARDA, JR.
Department of Economic Research
Central Bank of the Phil'ppines
Aduana, Intramuros, Manila D-406

PROF. CESAR M. LORENZO
394 Guevarra Avenue
San Juan, Rizal

DR. K. C. S. PILLAI
214 Forest Hill Drive
West Lafayette, Indiana 47906
U.S.A.

DR. LUIS R. SALVOSA
Philippine International Life Insurance Company
San Vicente, Manila D-404

MR. EXEQUIEL S. SEVILLA
National Life Insurance Co.
P.O. Box 2056, Manila

COL. SIMEON R. VENTURA
Headquarters Phil. Constabulary
Comptroller Division
Camp Crame, Quezon City D-507

MISS CONSUELO VIRATA
Department of Economic Research
Central Bank of the Philippines
Aduana, Intramuros, Manila D-406

DR. ENRIQUE T. VIRATA
University of the Philippines
Diliman, Quezon City D-505

INSTITUTIONAL MEMBERS AND THEIR REPRESENTATIVES

BANKS

1. CENTRAL BANK OF THE PHILIPPINES

   Miss Paz B. Culabutan
   Mr. Robustiano Espiritu
   Mrs. Aida R. Saldua
   Mr. Edgardo P. Zialcita

2. CHINA BANKING CORPORATION

   Mr. L. L. Pan

3. DEVELOPMENT BANK OF THE PHILIPPINES

   Mrs. Purita Floro
   Mrs. Alicia L. Reyes

4. PHILIPPINE NATIONAL BANK

   Miss Teresita S. Abaya
   Mr. Alfredo Ligon, Jr.

BUSINESS AND INDUSTRY

1. FILIPINAS CEMENT CORPORATION

   Mr. Maximiano Fantillo
   Mr. Rodrigo Jose
   Mr. Teofilo Raqueño

2. MANILA ELECTRIC COMPANY

   Mr. Jose M. Abreu
   Mr. Magsikap B. Mole
   Mr. Victor C. Risma

3. MENZI AND COMPANY, INC.

   Mr. Numeriano B. Aquino
   Mr. Loreto V. Alzate

4. NATIONAL INVESTMENT AND DEVELOPMENT CORPORATION

   Mr. Arturo G. Trinidad

5. PHILIPPINE REFINING COMPANY
   Mr. N. H. Timbol
   Mr. R. Zamora

6. PROCTER AND GAMBLE PHIL. MANUFACTURING CORPORATION
   Mrs. Carmencita N. Esteban
   Mr. Pedro A. Reyes

7. SYCIP, GORRES, VELAYO AND COMPANY
   Mrs. Gloria B. Mangoba
   Mr. Miguel Paredes

8. VICTORIAS MILLING CO., INC.
   Mr. Jose Sto. Domingo
   Mr. Teodorico R. Escober

INSURANCE

1. GOVERNMENT SERVICE INSURANCE SYSTEM
   Dr. Manuel O. Hizon
   Mrs. Tala P. Lazatin

2. THE INSULAR LIFE ASSURANCE CO., LTD.
   Mr. Walfredo Fojas
   Mr. Rolando Ello

3. NATIONAL LIFE INSURANCE COMPANY OF THE PHILIPPINES
   Mr. Exequiel S. Sevilla
   Mr. D. L. Francisco
   Mr. A. M. Nietes

4. PHILIPPINE AMERICAN LIFE INSURANCE COMPANY
   Mr. Alexander T. Brooks
   Miss Eligia G. Fernando

5. PROVIDENT INSURANCE COMPANY OF THE PHILIPPINES

6. SOCIAL SECURITY SYSTEM
   Mr. Agustin Mercado
   Mr. Pedro Sen

INDIVIDUAL MEMBERS

1. ABAD, AMANDO R.
   1623 J. Zamora St.
   Paco, Manila D-401

2. ABAYA, TERESITA S.
   EDP Center
   Philippine National Bank
   Escolta, Manila D-404

3. ABALOS, LAGRIMAS V.
   Bureau of the Census and Statistics
   J. P. Laurel St., Manila D-405

4. ABRIGO, TEODORICO
   San Miguel Corporation
   Market Research Department
   J. P. Laurel St., Manila D-405

5. ADRIANO, ROSA V.
   Bureau of the Census and Statistics
   P. O. Box 779, Manila

6. AFRICA, TOMAS P.
   434 Mendoza, Quiapo
   Manila

7. AGANON, VIRGILIO
   61 Main Street
   Cubao, Quezon City D-501

8. ALBA, MANUEL S.
   College of Business Administration
   University of the Philippines
   Diliman, Quezon City

9. ALCALA, CARMENIA R.
   Edward J. Nell Co.
   P. O. Box 612, Manila

10. ALCANTARA, ERNESTO S.
    Bureau of the Census and Statistics
    Laoag City Regional Branch
    Laoag City B-309

11. ALICBUSAN, ROSALINDA C.
    IRRI, College
    Laguna E-109

12. ALVIAR, SANTIAGO M.
    U.P. College of Agriculture
    Los Baños, Laguna E-109

13. AMEN, ALICIA M.
    Department of Mathematics
    Ateneo de Manila University
    Loyola Heights, Quezon City

14. ANIÑON, QUIRICO D.
Bureau of the Census and Statistics
P.O. Box 779, Manila

15. ARAGONES, FE CLARA N.
Division of City Schools
City Hall, Manila D-406

16. ARBOLEDA, HEIDI R.
76 South 18, Diliman
Quezon City D-505

17. ARICA, FLORENCIO A., JR.
Congressional Econ Planning Office
Congress of the Philippines
Manila

18. ARMEÑA, NELLIE D.
Bureau of the Census and Statistics
P.O. Box 779, Manila

19. BACALLA, ARTURO R.
Bureau of the Census and Statistics
P.O. Box 779, Manila

20. BAGUI, BRIGIDO O.
Bureau of the Census and Statistics
San Jose Regional Branch
San Jose, Occidental Mindoro F-207

21. BAILEY K. VERNON
c/o WHO Regional Office
P.O. Box 2932, Manila

22. BALITON, RODOLFO Q.
Bureau of the Census and Statistics
P.O. Box 779, Manila

23. BALMACEDA, ESTELA, A.
Department of Applied Mathematics
U.P. College of Agriculture
Los Baños, Laguna E-109

24. BANASIHAN, GERARDO
Philippine Sugar Institute
Philsugin Building
Diliman, Quezon City D-505

25. BANTEGUI, BERNARDINO G.
Office of Statistical Coordination
and Standards
National Economic Council
Padre Faura, Manila D-406

26. BARKER, RANDOLPH
U.P. College of Agriculture
College, Laguna E-109

27. BARTOLOME, MARCELO M.
Bureau of Agricultural Economics
San Fernando, La Union B-515

28. BARTOLOME, PEDRO V.
Philamlife Building
Tarlac Branch
Tarlac, Tarlac C-417

29. BELLEZA, INES G.
Mathematics Department
University of the Philippines
Diliman, Quezon City D-505

30. BELTRAN, DIOSDADO G.
Munitions Division
OASDS, DND
Camp Aguinaldo
Quezon City D-502

31. BERENGUEL, AURORA P.
46 San Lorenzo St.
Capitol Subdivision
Pasig, Rizal D-721

32. BENGSON, ROMULO Q.
Mathematics Department
Ateneo de Manila University
P.O. Box 154, Manila

33. BENGSON, ARTURO
103 R. Lagmay St.
San Juan, Rizal D-725

34. BERGSTRESSER, ROBERTO
c/o Actuarial Society of
the Philippines
P.O. Box 2712, Manila

35. BERNAL, ROSALINDA
Everett Steamship Corporation
290 Atlanta, Port Area
Manila D-402

36. BERNARDO, RENATO C.
Aliaga, Nueva Ecija A-601

37. BIAG, JUAN C.
Bureau of the Census and Statistics
P.O. Box 779, Manila

38. BLANCAS, REMEDIOS S.
University of the East
Claro M. Recto St.
Manila D-403

39. BLAQUERA, ROSARIO LL.
U.P. Statistical Center
P.O. Box 479, Manila

40. BORJA, ARTURO S.
Bureau of the Census and Statistics
P.O. Box 779, Manila

41. BRENNAN, CAROLINA
Mathematics Department
University of the Philippines
Diliman, Quezon City D-505

42. BRINGAS, NOLI
1111 M. de la Fuente
Sampaloc, Manila D-403

43. BUENAVENTURA, ANGELES R.
2111 Dakota St., Malate
Manila D-406

44. BUENAVENTURA, CORAZON R.
Office of Statistical Coordinartion
and Standards
National Economic Council
Padre Faura, Manila D-406

45. BUENO, SENEN C.
c/o Special Services Corporation
Insular Life Buiding
Ayala Avenue, Makati, Rizal D-711

46. CALABIO, WILFREDO T.
227-C Comdt. Garcia
Pasay City D-720

47. CALADO, RENE R.
184 Alabat, Balut
Tondo, Manila D-402

48. CALAYAG, BENJAMIN S.
Department of Agriculture
and Natural Resources
Casman Bldg., Quezon City D-505

49. CALDITO, GLORIA C.
U.P. Statistical Center
P.O. Box 479, Manila

50. CALISTON, VICENTE A.
Bureau of the Census and Statistics
Bacolod City K-501

51. CALIXTO, JULIA
Philippine Womens' University
Taft Avenue, Manila D-406

52. CAMPBELL, WALLACE G., S.J.
Department of Mathematics
Ateneo de Manila University
Loyola Heights, Quezon City

53. CAPAREDA, ELPIDIO A.
Bureau of the Census and Statistics
P.O. Box 779, Manila

54. CARILLO, MATILDE D.
69 Kundiman St.
San Francisco del Monte
Quezon City

55. CAROLL, JOHN S. J.
Institute of Social Order
509 Padre Faura, Ermita
Manila D-406

56. CASTRO, PROSPERO M.
SRDP, OSCAS
National Economic Council
Padre Faura, Manila D-406

57. CASTILLO, JOSE V.
Bureau of Agricultural Economics
Department of Agriculture and
Natural Resources
Casman Building, Quezon City D-505

58. CASTRO, ISAGANI DE
181 Pinatubo St.
Mandaluyong, Rizal D-713

59. CAYABYAB, MERLIN C.
Bureau of the Census and Statistics
P.O. Box 779, Manila

60. CAYANGA, EDNA M.
Bancom Development Corp.
CBTC Building, Ayala Avenue
Makati, Rizal

61. CAYME, NICOLAS T.
Bureau of the Census and Statistics
Dumaguete City, Negros Oriental
D-409

62. CEJALVO, FLOR V.
Mathematics Department
University of the Philippines
Diliman, Quezon City D-505

63. CHAN, MA. TERESA O.
128 Kanlaon St.
Quezon City

64. CHENG, JOHN WILLIAM
Philippine Bank of Communications
P.O. Box 934
Manila

65. CONCEPCION, MERCEDES B.
Population Institute
University of the Philippines
P.O. Box 479, Manila

66. CRUZ, ALFONSO R.
893-F Santolan, Cubao,
Quezon City D-501

67. CRUZ, ILDEFONSO T.
Institute of Hygiene
University of the Philippines
625 Herran, Manila

68. CRUZ, MILAGROS O. DE LA
10 Subic, San Francisco del Monte
Quezon City D-503

69. CRUZ, SANTIAGO F. DE LA
P.O. Box 1245, Manila

70. CUNANAN, JOAQUIN
Joaquin Cunanan & Co.
P.O. Box 2288, Manila

71. DAVID, ISIDORO P.
U.P. College of Agriculture
Los Baños, Laguna E-109

72. DERA, JESUS T.
103 Buagan, Caloocan City
D-706

73. DIAZ, LUIS S.
L. C. Diaz and Associates
2nd Floor Union Cement Bldg.
Port Area, Manila

74. DILLAGUE, VIRGINIA L.
Bureau of the Census and Statistics
P.O. Box 779, Manila

75. DIZON, MA. CARMENCITA
478 Gen. Luna Street
Marikina, Rizal

76. DY, BONIFACIO
461 U. N. Avenue
Ermita, Manila D-406

77. ENCARNACION, JOSE JR.
School of Economics
University of the Philippines
Quezon City D-505

78. ENGLE, LIWAYWAY M.
IRRI, College, Laguna E-109

79. ENGRACIA, LUISA T.
Bureau of the Census and Statistics
P.O. Box 779, Manila

80. ESGUERRA, RAYMUNDO S.
Mindanao Development Authority
C. M. Recto Avenue
Davao City

81. ESPIRITU, FABIAN G.
Agricultural Development Council for Rizal
Office of the Governor
Pasig, Rizal D-721

82. ESQUIVEL, JOSEFINA F.
Bureau of Plant Industry
Manila D-406

83. ESQUIVEL. ROMEO D.
2528 Bato St.
Gagalangin, Tondo
Manila D-402

84. EUSTAQUIO, MILAGROS J.
184 N. Domingo, San Juan, Rizal
D-724

85. EXMUNDO, SOLOMON R.
27-A Seattle, Cubao
Quezon City D-501

86. FERNANDEZ, CARLOS P.
Fernandez Hermanos, Inc.
109 Juan Luna, Manila D-406

87. FERNANDEZ, JOSE B. JR.
Far East Bank & Trust Co.
Trade Center Building
Aduana, Intramuros, Manila

88. FERRAZZINI, ANNIE M.
E. R. Squibb and Sons, Phil. Corporation
P.O. Box 1618, Manila

89. FERRER, PILAR L.
1569 Dapitan Street
Manila D-403

90. FLIEGER, WILHELM, S.V.D.
Population Institute
University of the Philippines
P.O. Box 1280, Manila

91. FLORES, BONIFACIO L.
Merck Sharp & Dohme (Phils) Inc.
P.O. Box 1280, Manila

92. FLORES, LYDIA H.
Mathematics Department
University of the Philippines
Diliman, Quezon City D-505

93. FLORES, TOMAS W.
Wage and Position Classification Office
Classification Office
Budget Commission
Malacañang Compound, Manila

94. FLORES, MERCEDES G.
67 San Francisco St.
Bo. Kapitolyo
Pasig, Rizal

95. FLORESCA, EMMANUEL T.
IRRI, College, Laguna E-109

96. FONACIER, JOSEFINA C.
Mathematics Department
University of the Philippines
Diliman, Quezon City

97. FRANCIA, JOSE MA., S.J.
Ateneo de Manila University
P.O. Box 154, Manila

98. FRANCIA, RAFAELITA C.
4 Batay, Cubao, Quezon City

99. FRANCISCO, CORAZON
25 Seattle, Quezon City D-501

100. GABRILLO, PROCOPIO
Bureau of the Census and Statistics
Cabanatuan City Branch
Cabanatuan City A-603

101. GACUSAN, NELIA Z.
T-161 C, Area 2
U.P. Campus, Diliman
Quezon City D-505

102. GADI, REYNALDO T.
55 Matimtiman St.
Teachers' Village
Quezon City D-505

103. GALANG, EULOGIO G.
Office of the Assistant Secretary for Services
Department of National Defense
Fort Aguinaldo
Quezon City D-502

104. GAON, BENJAMIN V.
Agricultural Economics
U.P. College of Agriculture
College, Laguna E-109

105. GARCIA, AGAPITO M.
General Operations Analysis Supervisor
Philamlife, U. N. Avenue
Manila D-406

106. GOMEZ, KWANCHAI ASSApUL
IRRI, College, Laguna E-109

107 GOMEZ. WILFREDO C.
IRRI, College, Laguna E-109

108 GONZALES. FRANCISCO V., JR.
Erlanger and Galinger
Araza Building, Paseo de Roxas
Makati, Rizal

109 GONZALES, MANUEL M.
Filoil Marketing Corporation
984 Taft Avenue, Manila D-406

110. GOPEZ, EDUARDO C.
#9 Hercules St.
Bel Air Village
Makati, Rizal D-711

111. GOTIA, CECILIA D.
U.E. College of Liberal Arts and Sciences
Claro M. Recto Avenue
Manila D-403

112. GRAU, CESAREO H.
Philippine American Life Insurance Company
P.O. Box 1152, Manila

113. GULLA, TEOFILO F.
Bureau of the Census and Statistics
Dagupan City C-318

114. GUMABONG, RODOLFO P.
Bureau of the Census and Statistics
Iloilo Branch, Iloilo City K-423

115. GUTIERREZ, JOSE S.
The Statistical Center
University of the Philippines
P.O. Box 479, Manila

116. GUZMAN, ADOLFO M. DE
Mathematics Department
Ateneo de Manila University
Ateneo de Manila University
P.O. Box 154, Manila

117. GUZMAN, AMELIA G. DE
2512 Nacar Street
San Andres Subdivision
Manila D-406

118. GUZMAN, GENEROSO G. DE
The Statistical Center
University of the Philippines
P.O. Box 479, Manila

119. HENARES, ROSARIO
66 Banahaw St., Cubao
Quezon City D-501

120. HENDERSHOT, GERRY E.
Sociology and Anthrophology Dept.
Vanderbilt University
Nashville, Tennessee
U.S.A.

121. HERBER, TEODORICO
Dept. of Economic Research
Central Bank of the Phils.
Aduana St., Intramuros
Manila D-406

122. HORIGAN, MARTHA R.
Great Pacific Life Assurance
Corporation
Yuchengco Building
484 Rosario, Binondo
Manila D-405

123. IBAÑEZ, DOROTHY
Bureau of the Census and Statistics
P.O. Box 779, Manila

124. ILANO, ALBERTO R.
College of Business Administration
University of the Philippines
Diliman, Quezon City D-505

125. INCIONG, RODRIGO
20 Ynchausti Subdivision
Malabon, Rizal D-712

126. ITCHON, GABRIEL Y.
Department of Economic Research
Central Bank of the Phils.
Intramuros, Manila D-406

127. JACILDO, LUIS
Economic Census Division
Bureau of the Census and Statistics
J. P. Laurel Sr., Manila D-405

128. JOSE, ROSARIO D.
2740 Zamora St.
Pasay City D-720

129. JOSE, FLORDELIZA J.
2448 Bato Street
Gagalangin, Tondo, Manila

130. JUAN, MABINI L.
Office of the Actuary
Social Security System
East Avenue, Quezon City
D-505

131. KEH, ALFONSO
359 Plaza Sta. Cruz
Manila D-404

132. LADEMORA, ELISEO M., JR.
800 Herran St., Malate
Manila D-406

133. LAING, JOHN
Population Institute
University of the Philippines
P.O. Box 479, Manila

134. LASMARIAS, VICENTE B.
Department of Applied Mathematics
UPCA, College, Laguna E-109

135. LEGASPI, CATALINO P.
Bureau of the Census and Statistics
Iligan City M-104

136. LEONOR, CONCEPCION
College of Science
University of Sto. Tomas
España, Manila D-403

137. LIAO, CHRISTINA Y. M.
Apt. No. 10, U.P.C.O.
Staff Housing
Los Baños, Laguna E-109

138. LIM, WILLIAM O.
557 Nueva Street
Escolta, Manila D-404

139. LITTAUA, ROBERTO R.
Elizalde and Company, Inc.
P.O. Box 121, Manila

140. LOMOTAN, CESAR J.
Assistant to the Governor
Central Bank of the Philippines
Aduana, Intramuros, Manila D-406

141. LONTOC, PAULINA L.
College of Liberal Arts
and Sciences
University of the East
Claro M. Recto Avenue
Manila

142. LYNCH, FRANK, S.J.
Institute of Philippine Culture
Ateneo de Manila University
P.O. Box 154, Manila

143. MADIGAN, FRANCIS C., S.J.
Research Institute in
Mindanao Culture
Xavier University, Cagayan de Oro
City L-305

144. MALVAR, AMANDO
837 Craig, Sampaloc
Manila D-403

145. MAKANAS, ELPIDIO D.
National Income Branch
Office of the Statistical
Coordination and Standards
National Economic Council
Padre Faura St., Manila

146. MAGALIT, HENRY F.
Department of Applied Mathemaitcs
College of Agriculture
College, Laguna

147. MAGARO. EFIGENIA
Social Security System
East Avenue, Quezon City D-505

148. MAGTAJAS, PEDRITO P.
Bureau of the Census and Statistics
Cagayan de Oro City L-305

149. MALANG, MAGUID B.
Bureau of Agricultural Economics
Davao City

150. MANALO, ANTERO S.
IRRI, Los Baños, Laguna
E-109

151. MAPA, FELINA G.
Mathematics Department
University of the Philippines
Diliman, Quezon City D-505

152. MARGALLO, MODESTA, U.
1142 Estrada, Manila D-406

153. MARIANO, ROBERTO S.
School of Economics
University of the Philippines
Diliman, Quezon City

154. MARQUEZ, NELIA R.
Bureau of the Census and Statistics
Dña. Soledad Bldg .
De Dios St., Sta. Mesa
Manila D-408

155. MATA, NICOLAS D., JR.
Bureau of the Census and Statistics
P.O. Box 779, Manila

156 MAULIT, DIMAS A.
Presidential Economic Staff
Arlegui St., Manila

157. MCPHELIN, MICHAEL, S.J.
Department of Economics
Ateneo de Manila University
P.O. Box 154, Manila

158. MENDOZA, ARTEMIO
The Insular Life Assurance Co.
P.O. Box 128, Manila

159. MENDOZA, CASIANO B.
9609 Josephine St.
Baltao, Parañaque, Rizal

160. MERCADO, JOSE C.
Statistical Division, GHQ
AFP, Fort Aguinaldo
Quezon City D-502

161. MERCADO, JULIAN
Central Bank of the Phils.
Aduana, Intramuros, Manila
D-406

162. MIJARES, TITO A.
Bureau of the Census and
Statistics
Dña. Soledad Building
Sta. Mesa, Manila D-403

163. MONTESINO, RAYMUNDO
1131 Maria Cristina St.
Sampaloc, Manila

164. MORTEL, DOLORES M.
Bureau of the Census and
Statistics
P.O. Box 779, Manila

165. MORTON, JOSEPH E.
Troy Towers
380 Mountain Road
Union City, New Jersey 07087
U.S.A.

166. MUÑOZ, ANGELITA G.
1055 T. Ayala St.
Ermita, Manila D-406

167. MURPHY, EDMUND M.
1 B Springfield Place
420 E. Woodland Avenue
Springfield, Pennysylavania 19064
U.S.A.

168. NATIVIDAD, CARIDAD M.
Department of Mathematics
University of the Philippines
Diliman, Quezon City D-505

169. NAZARET, FRANCISCO V.
Demography and Social
Statistics Division
Bureau of the Census and
Statistics
J. P. Laurel Sr., Manila D-405

170. NOLLEDO, MANUEL M.
Bureau of the Census and
Statistics
Puerto Princesa, Palawan
D-617

171. OCAYA, JOSEPHINE A.
Bureau of the Census and
Statistics
P.O. Box 779, Manila

172. OLAGUER, EDUARDO B.
33 Dr. Lazcano St.
Quezon City D-501

173. OÑATE, BURTON T.
Asian Development Bank
Oledan Building
Makati, Rizal D-711

174. OPEÑA, LILIA C.
2339-E Singalong St.
Malate, Manila D-406

175. ORENSE, MARCELO M.
Mathematics Department
Ateneo de Manila University
Loyola Heights, Quezon City

176. PACHECO, ANTONIO R.
Lagao, General Santos
South Cotobato

177. PANGANIBAN, ELENITA C.
Department of Savings and
Loan Association
Central Bank
Beneficial Life Building
Solano, Manila

178. PARAZO, CONSUELO M.
Bureau of the Census and Statistics
P.O. Box 779, Manila

179. PAREL, CRISTINA P.
The Statistical Center
University of the Philippines
P.O. Box 479, Manila

180. PARKER, BENEDICTO A.
Department of Animal Husbandry
U.P. College of Agriculture
College, Laguna E-109

181. PARRAS, JOSEFINA P.
Department of Applied Mathematics
UPCA, College, Laguna E-109

182. PAULINO, LEONARDO A.
Bureau of Agricultural and Natural Resources
Casman Building, Quezon City
D-505

183. PAZ, ANA MARIA DE LA
1235 Trece de Agosto
Paco, Manila

184. PAZ, PANFILO P. DE LA
Office of the Insurance Commissioner
Gen. Luna, Intramuros,
Manila D-406

185. PEREDO, BENJAMIN D.
Bureau of the Census and Statistics
3rd Floor, Dña. Soledad Bldg.
Sta. Mesa, Manila D-403

186. PEREZ, ANTONIO G.
#39 Kamias Road
Quezon City D-502

187. PEREZ, BERNARDINO A.
Office of Statistical Coordination and Standards
National Economic Council
Padre Faura, Manila D-406

188. PEREZ, ELISEO
Bureau of the Census and Statistics
P.O. Box 779, Manila

189. PETATE, MAXIMO P.
Bureau of the Census and Statistics
Tacloban City Branch
Tacloban City

190. PIRON, GEORGES
Department of Economics
De la Salle College
Taft Avenue, Manila

191. PRATT, WILLIAM F.
Population Research and Training Center
University of Chicago
1413 East 60th St.
Chicago, U.S.A.

192. PUYAT, GIL J.
Philippine Senate
Manila D-406

193. QUIAOIT, FILEMON D.
996 Sisa St.
Sampaloc, Manila D-403

194. QUICHO, GERARDO
Engr. Training Center
Camp Tinio, Bañgad
Cabanatuan City A-603

195. RAMISCAL, EFREN R.
Department of Mathematics
Ateneo de Manila University
Loyola Heights
Quezon City D-502

196. RAMOS, EDUARDO R.
9023 Aranga, San Antonio Village
Makati, Rizal D-711

197. RAMOS, ERNESTO R.
2309 Radium St.
San Andres, Manila
D-406

198. RAMOS, MARIANO B. DE
Department of Applied Mathematics
UPCA, College, Laguna E-109

199. RAMOS, VIOLETA L.
2449 Agata, San Andres Subd.
Manila D-406

200. RAÑOA, EULOGIO F.
Mineral Economics and Information Division
Bureau of Mines
Herran, Manila D-406

201. REGUDO, ADRIANA
Population Institute
University of the Philippines
P.O. Box 479, Manila

202. RETIRADO, LETICIA E.
Bureau of the Census and Statistics
P.O. Box 779, Manila

203. REYES, AZUCENA
142 Retiro, Sta. Mesa Heights
Quezon City D-502

204. REYES, FE
Mathematics Department
University of the Philippines
Diliman, Quezon City D-505

205. REYES, PEREGRINO S.
Office of the Statistical Coordination and Standards
NEC
Padre Faura, Manila D-406

206. REYES, RODOLFO D.
IRRI, College
Laguna E-109

207. RICO, HONESTO B.
Asian Development Bank
Metropolitan Bank Building
Makati, Rizal D-711

208. ROA, FEDERICO, SR.
105 A. Melchor
Loyola Heights
Quezon City D-502

209. ROLDAN, GLORIOSO
Bureau of the Census and Statistics
P.O. Box 779, Manila

210. ROSAL, LETICIA R.
Bureau of the Census and Statistics
P.O. Box 779, Manila

211. ROSAURO, PANFILO J.
Mindanao State University
Marawi City M-206

212. ROSETE, TIMOTEO
Division of Surveys
Bureau of the Census and Statistics
J. P. Laurel Sr. St.
Manila D-405

213. ROSIMO, ELPIDIO T.
Bureau of the Census and Statistics
P.O. Box 779, Manila

214. ROXAS, CONSUELO A.
Bureau of the Census and Statistics
P.O. Box 779, Manila

215. RUCIO, ADELA R.
Bureau of the Census and Statistics
P.O. Box 779, Manila
Rizal

216. RUCIO, CARMENCITA R.
21 Aquamarine St.
SSS Village, Marikina
Rizal

217. RUTTAN, VERNON W.
Dept. of Agricultural Economics
University of Minnesota
St. Paul, Minnesota 55101
U.S.A.

218. SABINO, GUADALUPE C.
2164 Leveriza St.
Malate, Manila D-406

210. SADABA, CAMILO S.
Government Service Insurance System
Arroceros, Manila D-406

220. SAGURAN, FELIX M.
Bureau of the Census and Statistics
P.O. Box 779, Manila

221. SALVOSA, SALVADOR B.
Pan Malayan Management and Investment Corporation
Yuchengco Building
484 Rosario, Binondo, Manila

222. SAMSON, ALEJANDRO M.
119 Scout Lozano St.
Quezon City D-502

223. SAMSON, PABLO Q., JR.
Office of Statistical Coordination and Standards
National Economics Council
Padre Faura, Manila D-406

224. SANCHEZ, CECILIA A.
Board of Investments
Amalgamated Building
Ayala Avenue, Makati
Rizal D-711

225. SANTOS, EMILIO A.
Department of Economic Research
Central Bank of the Philippines
Aduana St., Intramuros
Manila

226. SANTOS, FRANCISCO C.
Bureau of the Census and Statistics
San Fernando, Pampanga
C-217

227. SANTOS, ONOFRE L.
1027 Oroquieta St.
Sta. Cruz, Manila D-404

228. SARDIDO, MOISES L.
IRRI, College, Laguna
E-109

229. SARREAL, ROBERTO
1433 Laguna, Sta. Cruz,
Manila D-404

230. SIA, LUISA
UP Statistical Center
P.O. Box 479, Manila

231. SELDA, GERVACIO G., JR.
111Naranhita St.
Project 2, Quezon City

232. SEVILLA, CORAZON L.
22 3rd Avenue, Cubao
Quezon City

233. SIMBULAN, CESAR G.
Philippine American Life Insurance Company
Philamlife Building
Manila D-406

234. SINGGIH, UMAR
P.O. Box 2719, Manila

235. SINSIOCO, CEFERINO S.
186 M. Clara St.
4th Avenue, Grace Park
Caloocan City

236. SISON, MA. LOURDES LILIA G.
National Life Insurance Co.
Regina Building, Escolta, Manila
D-404

237. SMITH, DOLORES S.
Department of Statistics
U-120
University of Connecticut
Storrs, Connecticut 06268
U.S.A.

238. SMITH, PETER COLIN
UP Population Institute
PO Box 479, Manila

239. SOBERANO, EDITA M.
College of Education
University of the Philippines
Diliman, Quezon City

240. SONGCO, ALBERTO S.
45 Silencio cor. Concepcion St.
Araneta Subdivision
Quezon City D-502

241. STA. INES, SANTIAGO
Philippine Military Academy
Fort del Pilar, Baguio City

242. STINNER, WILLIAM
Population Institute
P.O. Box 479, Manila

243. SUGUITAN, LOURDES
Bureau of the Census
and Statistics
P.O. Box 779, Manila

244. SUMAGUI, JUAN O.
Office of the Statistical
Coordination and Standards
National Economic Council
Padre Faura, Manila D-406

245. SUTJIHNO
IRRI, College, Laguna E-109

246. TALEON, JESUSA T.
Department of Applied
Mathematics
UPCA College, Laguna E-109

247. TALLO, RODOLFO B.
1123-D Consuelo St.
Paco, Manila D-401

248. TAN, EVELYN L.
U.P. Counseling and Testing
Center
Diliman, Quezon City D-505

249. TAN, FLORENTINA A.
2626 Dian St., Manila

250. TAN, JOSE G.
486 Del Monte Avenue
Quezon City D-503

251. TAN, RUDY
The Statistical Center
P.O. Box 479, Manila

252. TAYCO, HERMINIA J.
Tariff Commission
Quezon Blvd. Extension
Quezon City D-502

253. TEODORO, PEDRO E.
Philprom Incorporated
640 Vito Cruz, Manila D-406

254. TIAOQUI, ERLINDA V.
Beneficial Life Insurance Co.
Trade Center Building
Aduana, Intramuros, Manila

255. TIENZO, BENJAMIN P.
Board of Investments
Amalgamated Building
Ayala Avenue, Makati
Rizal D-711

256. TIENZO, IRENEA V.
Labor Statistics Services
Department of Labor
Phoenix Building, Intramuros
Manila

257. TRINIDAD, IRENEA V.
BAECON, DANR
Casman Building, Quezon Blvd.
Quezon City

258. TRINIDAD, RUBEN F.
Joint Legislative Executive
Tax Commission, Manila

259. UY, ALFREDO S.
Alpha Mutual Life Insurance Co.
Plaza Sta. Cruz,
Manila D-404

260. VALBUENA, JUSTO B.
Forecasting Center
Weather Bureau
Manila D-402

261. VALDEPEÑAS, VICENTE B., JR.
Department of Economics
Ateneo de Manila University
Quezon City D-502

262. VALENZUELA, VICTOR C.
Institute of Hygiene
University of the Philippines
Herran, Manila

263. VARONA, ALBINO P.
10 DANR St., Vasra Village
Project 6, Quezon City

264. VIDAL, MELBA J.
Asian Soical Institute
1518 Leon Guinto, Malate
Manila D-406

265. VIERNES, CECILIA G.
Bureau of the Census
and Statistics
P.O. Box 779, Manila

266. VILLANUEVA, BONIFACIO D.
Bureau of Agricultural
Economics
Butuan City L-104

267. VILLANUEVA, RODRIGO S.
Bureau of the Census and
Statistics
Tuguegarao Branch
Tuguegarao, Cagayan A-329

268. VILLARUZ, FELICIDAD
Government Service
Insurance System
Arroceros, Manila

269. VILLAVICENCIO, BENITO
Asia Research Organization, Inc.
884 Vito Cruz, Manila D-406

270. VIRATA, NESTOR A.
Executive Management Group, Inc.
Metropolitan Bank Building
144 Ayala Avenue, Makati, Rizal
D-711

271. VIZCARRA, ERNESTO R.
U.P. College of Dentistry
Padre Faura, Manila D-406

272. WARREN, NARCISO V.
Bureau of the Census
and Statistics
P.O. Box 779, Manila

273. WATTS, IAN EDWARD MEIN
United Nations Development
Programme
P.O. Box 1864, Manila

274. YAP CHAI, JANE
1141 Florida, Ermita
Manila D-406

275. YÑIGUEZ, AMADOR D.
Department of Applied
Mathematics
U.P. College of Agriculture
College, Laguna E-109

276. YOINGCO, ANGEL
Joint Legislative Executive
Tax Commission
Intramuros, Manila D-406

277. ZABALLA, FIDELA
Colgate Palmolive Phils. Inc.
Makati, Rizal D-711

278. ZALAMEA, CESAR G.
Philamlife Co.
Philamlife Building
U.N. Avenue, Manila
D-406

Republic of the Philippines
Department of Public Works and Communication

BUREAU OF POSTS
Manila

SWORN STATEMENT
(Required by Act 2580)

The undersigned CRISTINA P. PAREL, editor of THE PHILIPPINE STATISTICIAN published bi-annually 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: CRISTINA P. PAREL	U.P. Statistical Center
Managing Editor: JOSE S. GUTIERREZ	U.P. Statistical Center
Owner: Philippine Statistical Association, Inc.	— do —
Publisher: — do —	— do —
Printer: U.S.T. Press	España, Manila
Office of Publication: Rizal Hall, Padre Faura St. Manila	

In case of publication other than daily, total number of copies printed and circulate of the last issue dated , 197

1. Sent to paid subscribers 475
2. Sent to others than paid subscribers 50

Total 525

(Sgd.) CRISTINA P. PAREL
Editor

SUBSCRIBED AND SWORN to before me this 14th day of November, 1970, at Makati Rizal, the affiant exhibiting his Residence Certificate No. A-4293792 issued at Quezon City on January 13, 1970.

Doc. No. 66
Page No. 15
Book No. III
Series of 1970

(Sgd.) ROMULO S. de ERAÑO
Notary Public
Until December 31, 1970