AN ABSTRACT OF THE THESIS OF

BRUCE BYRON ELLIS for the M. S. in Agricultural Economics
(Name) (Degree) (Major)

Date thesis is presented November 27, 1962

Title DETERMINATION OF OPTIMUM RAW PRODUCT SAMPLING PROCEDURES WITH SPECIAL REFERENCE TO GREEN BEANS FOR PROCESSING

Abstract approved Redacted for privacy
Major Professor

Most raw product received at processing plants in Oregon is sampled to determine quality characteristics as a basis of grower payment. Raw product sampling procedures must provide the processor with a means of sampling which will be sufficiently precise to satisfy himself as well as his growers, and thus avoid misunderstandings relative to the true grade of product delivered.

How well a given sampling scheme will accurately provide this information depends upon the amount and nature of variability among characteristics of the product within a given load and the size and type of sample being drawn. Through the use of probability sampling theory it is possible to judge the precision of a sampling procedure by examining

the frequency distribution generated for the estimate if the procedure is applied repeatedly. Based upon the knowledge of the frequency distribution of the estimate, it is possible to determine the range within which, as an example, 95 per cent of all possible samples would fall.

The processor not only is interested in the precision and accuracy of his sampling procedures, but he also is concerned with costs of sampling. This study considers costs of sampling as well as precision in determining optimum sampling procedures which will minimize cost in achieving a given level of precision or maximize precision for a given cost.

Green beans for processing were used in the application of sampling theory. Green beans are delivered to the processing plants in tote bins loaded on trucks. Two-stage sampling procedures are used where a sample of totes is drawn from the truck and a sample of beans is drawn from each tote selected. Repeated sampling was used to estimate the variation existing among primary and secondary sampling units. Three sampling schemes were employed to estimate components of variation: In Sampling Scheme A two 10-pound secondary units were drawn from each of two primary units. Sample Scheme B consisted of drawing two 30-pound secondary units from each of two primary units. Sample Scheme C consisted of drawing a single continuous sample drawn from each of two primary units. The study was designed so that a comparison might be made of (1) the precision achieved by various numbers of primary and secondary sample units, (2) the effect of the size of the secondary units on among primary and among secondary unit variation, (3) the completely random and the random systematic (continuous) method of obtaining the secondary unit.

Costs involved in sampling raw product were estimated by economic-engineering techniques. Costs were divided into two components--costs associated with the drawing of the primary unit, and costs incurred in drawing and grading the secondary unit. Grading costs were a major portion of total estimated costs of sampling.

Variance and cost estimates were brought together for use in the final analysis as follows: (1) optimum sampling plans were determined for each of the three sampling schemes for selected levels of precision; (2) comparisons were made of the optimum for each of the three schemes to determine the scheme which provides selected levels of precision at the lowest costs; (3) generalizations were made with regard to other sampling schemes involving sub-samples of sizes other than those included in the study and (4) a comparison was made of the simple random and the continuous method of selecting the sub-sample.

Based on the range of precision and confidence considered in the study, sampling Scheme A appears to provide a least cost method of achieving given levels of precision and confidence.

DETERMINATION OF OPTIMUM RAW PRODUCT SAMPLING PROCEDURES WITH SPECIAL REFERENCE TO GREEN BEANS FOR PROCESSING

by

BRUCE BYRON ELLIS

A THESIS

submitted to

OREGON STATE UNIVERSITY

in partial fulfillment of the
requirements for the
degree of

MASTER OF SCIENCE

June 1963

APPROVED:

Redacted for privacy

Assistant Professor of Agricultural Economics

In Charge of Major

Redacted for privacy

Head of Department of Agricultural Economics

Redacted for privacy

Dean of Graduate School

Date thesis is presented November 27, 1962

Typed by Nancy Kerley

ACKNOWLEDGEMENTS

The writer wishes to express his gratitude to those who guided and aided him in the completion of this work. Without the valuable guidance and advice of Dr. Harvey M. Hutchings the manuscript could not have been completed. The late Professor George B. Davis contributed much as he supervised the inception of the project and the gathering of data until his appointment in Thailand. The technical supervision of Dr. Lyle D. Calvin was most valuable and contributed to the results in the thesis. Thanks are due to Mr. Frank H. J. Dickmann for his assistance in collecting the data and for his review of the manuscript.

Appreciation is also extended to Dr. William G. Brown, Mr. John A. Edwards and Mr. Charles Green for their review of the manuscript.

The fine cooperation of the fruit and vegetable processing industry enabled this study to be conducted.

TABLE OF CONTENTS

CHAPTER		PAGE
I	INTRODUCTION	1
II	THEORETICAL FRAMEWORK	12
	Probability Sampling	12
	Two-Stage Sampling	16
	Random Sampling	18
	Random Continuous Sampling	23
	Estimating Costs of Sampling	24
	Estimating Optimum Samples	25
III	PROCEDURES USED IN THE STUDY	28
	Experimental Sampling	28
	Estimating Costs of Sampling	33
	Determination of Optimum Sampling Methods	35
IV	RESULTS OF ANALYSIS	37
	Estimates of Variation	37
	Estimated Costs of Sampling	46
	Optimum Sampling	52
V	SUMMARY AND CONCLUSIONS	61
	BIBLIOGRAPHY	68
	APPENDIX	70

LIST OF FIGURES

FIGURE		PAGE
1	Normal Distribution	14
2	A Population of Raw Product Illustrating Primary and Secondary Sampling Units	16
3	Combination of Primary and Secondary Sampling Units Required to Achieve Selected Levels of Precision For Given Levels of Confidence and Sub-sample Sizes.	42
4	The Number of Totes Required to Achieve Selected Levels of Precision with 95 and 90 Per Cent Confidence Taking a Single Sixty Pound Continuously Drawn Sub-sample.	45
5	Total Costs Incurred in Drawing Samples Consisting of Various n and m in Using a Ten Pound Sub-sample, 1962.	51
6	Total Cost Incurred in Drawing Samples Consisting of Various n and m Using a Thirty Pound Sub-sample, 1962.	51
7	Total Cost Incurred in Drawing Samples Consisting of Various Numbers of Primary Units Using a Sixty Pound Sub-sample, 1962.	52
8	Costs of Achieving Various Levels of Precision at 95 and 90 Per Cent Confidence, 1962.	57

LIST OF TABLES

TABLE		PAGE
1	Analysis of Variance of Sample Estimates.	20
2	Sampling Schemes Used in Estimating Variation.	29
3	Size Classifications for Green Beans and the Values in Dollars for Each Grade, 1962.	32
4	Mean Squares and Estimates of Variation for 10, 30, and 60 Pound Samples.	38
5	Precision Levels Resulting from Various Combinations of Primary and Secondary Sampling Units for Given Levels of Confidence and Sub-sample Size.	44
6	Labor Standards and Costs for Drawing and Grading Green Bean Samples Using Two-Stage Sampling Techniques and Selected Sizes of Sub-samples.	47
7	Equipment Standards of Performance and Costs for Drawing and Grading Green Bean Samples, Using Two-Stage Sampling Techniques and Selected Size of Sub-sample.	49
8	Estimated Costs of Obtaining a Primary Sampling Unit and Drawing and Grading Secondary Sample Units of Size 10, 30, and 60 Pounds.	50
9	Optimum Sample Size and Cost for 10, 30, and 60 Pound Secondary Sample Units at 95 and 90 Per Cent Confidence.	55

LIST OF APPENDIX TABLES

TABLE		PAGE
1	Raw Product Value Estimates in Dollars Per Ton. Derived from Repeated Sampling of Ten Pound Samples by Plant and Year.	70
2	Raw Product Value Estimates in Dollars Per Ton, Derived from Repeated Sampling of Thirty Pound Samples by Plant and Year.	71
3	Raw Product Value Estimates in Dollars Per Ton, Derived from Repeated Sampling of Sixty Pound Samples by Plant and Year.	72
4	Analysis of Variance by Plant and Year for Ten, Thirty, and Sixty Pound Sub-samples, 1961-62.	72
5	Estimated Costs for Equipment Used in Sampling Green Beans, 1962.	73
6	Job and Job Descriptions for Labor Standards for Green Bean Sampling, 1962.	74

DETERMINATION OF OPTIMUM RAW PRODUCT SAMPLING PROCEDURES WITH SPECIAL REFERENCE TO GREEN BEANS FOR PROCESSING

CHAPTER 1

INTRODUCTION

Field run fruits and vegetables vary with respect to size, shape, color, texture, state of ripeness and other attributes of quality. These characteristics vary within and among farms, fields, truckloads and even boxes on a single truck. Quality-attributes provide a basis for grading the product, determining its value, and establishing the level of grower payments for most fruits and vegetables grown for processing in Oregon. Determination of the quality and value of product in a precise and accurate manner is of importance to both growers and processors.

Two means of making this determination are presently available to growers and processors in Oregon. One method is through the use of a system that retains the identity of separate loads received from individual growers until they have passed through the culling, grading and inspection stages of the processing plant, so that the product falling into each grade classification can be weighed and credited to the proper grower. This method is costly in that it requires individual loads of product to be processed separately and results in breaks in flow of product between these individual loads, thus causing idle time for equipment and personnel of the plant while lots are being changed. The method is accurate because it permits grower payments to be based on the grading of the entire lot of product.

The other method of grade determination, and the one most commonly practiced in the fruit and vegetable processing industry today, is that of sample grading. This is the process in which only a portion of any given load is graded to determine quality and grades for the entire load of product. The sample is usually drawn before the product has entered the processing plant. Therefore, it eliminates the need to keep growers' lots separate during processing.

The objective of both of these methods of grading is to provide a basis for accurate payment to growers for their product. Under the sample grading method the ability of the sample to reflect the actual grades of product in the load becomes extremely important. How well a given sampling scheme will accurately provide this information depends upon the amount and nature of variability or variation among the characteristics of the product within a given load and the size and type of sample being drawn.

When the product of a given load is homogeneous and uniform in its quality characteristics, a very small sample is sufficient to accurately estimate the characteristics of the product from which it was drawn. In such a case the way in which the sample is drawn from the load also is of little importance because any sample will accurately reflect the quality of the load. Unfortunately, considerable variability in quality characteristics usually exists for most fruits and vegetables grown for processing and under these circumstances the method of drawing the sample and the quantity of product to be drawn is of importance. These considerations become more important, of course, as variation in raw product increases.

Raw product sampling must provide the processor with a means of sampling which will be sufficiently precise to satisfy himself as well as his growers and, thus, avoid misunderstandings relative to the true grade of product delivered. A measurement of this precision is made possible through the use of probability sampling theory.[1/] This theory unfortunately does not yield a measurement of the error in any given single sample estimate. This cannot be measured unless the true value for the entire load is known and this generally is not the case or there would be no need for sampling. It does, however, make it possible to judge the precision of the sampling procedure by examining the frequency distribution which is generated for the estimate if the sampling procedure is applied over and over again to the same load of product (3, p. 6). Based upon the knowledge of the frequency distribution of the estimate, it is possible to determine the range within which, as an example, 95 per cent of all possible samples would fall and thus indicate something about the frequency of an event occurring. This theory can be applied to measure the precision for a given size of sample drawn by probability sampling. Probability sampling theory as it applies to this problem is discussed more fully in Chapter 2. It should be pointed out here, however, that the theory applies only to methods of sampling in which every unit has a known

---

1/ Precision refers to the clustering of the estimates about the mean of all possible samples while accuracy refers to the ability of the mean of all possible samples to estimate the true population mean. Accuracy refers to lack of bias.

chance of being selected. Judgment or purposive sampling, which is the technique whereby units are drawn because they appear "typical" or convenient to draw is not amenable to the development of a theory (3, p. 7). Precision cannot be determined for this type of sampling.

There appears to be a general lack of understanding on the part of the processing industry regarding the theory and application of probability sampling procedures. This is evidenced by the wide variety of sampling procedures currently being practiced in the industry today. A preliminary survey of sampling procedures currently used in fruit and vegetable processing plants in Oregon indicated, for example, that strawberry samples for common size loads ranged from 2 to 18 hallocs. In some cases sample size was increased as load size increased, while in others, the sample size per load was held constant. Similar results were found in the sampling of other fruits and vegetables. Samples drawn by many processing plants were not selected in such a manner that every unit had a known chance of being drawn. The level of precision achieved by these judgment samples is not subject to measurement and therefore is not known by growers and processors.

A knowledge of the level of precision and degree of confidence within which the value of the product can be estimated will result in greater mutual respect and understanding on the part of both growers and processors. Utilizing probability sampling procedures and theory, processing plants would be able to determine the amount of error in their sampling methods. One purpose of this study is to present probability sampling theory and techniques as they apply to the sampling of raw product. These can be used in implementing procedures for sampling raw product at levels of precision acceptable to both the grower and the processor. Even more important the study will provide the methodology for evaluating existing sampling procedures should their validity be questioned.

The processor not only is interested in the precision and accuracy of his sampling procedures, but he also is concerned with the costs of sampling. As was indicated earlier, the most precise method for determining the value of product can be achieved by grading the entire lot. This procedure is not necessary if sampling techniques employed at a lower cost will yield the desired precision. This study therefore will consider cost of sampling as well as level of precision in determining optimum sampling procedures. More specifically it will deal with means of determining optimum sampling procedures which will minimize cost in achieving a given level of precision or maximize precision for a given cost.

Probability sampling techniques need not be confined by processors to the sampling of raw product. It may be possible with only slight modification to utilize the same techniques at other points in the processing plant. The information derived might be used in quality control of the end product or in improving the efficiency of plant operations. For example, sampling might be done in the plant to determine if sufficient cullage or reject material was being removed prior to final processing.

Probability sampling procedures may be used for a variety of raw products. Although the amount of variability that exists within the raw product is different for each commodity, procedures used to estimate variability and determine the optimum number of samples are similar. Resources available for this study do not permit a determination of optimum sampling plans for a wide variety of fruits and vegetables. As an alternative, one commodity has been selected and used as a vehicle to carry out the more general objectives of the study. Green beans are used to demonstrate the employment of probability sampling theory to estimate precision of the sampling procedure and to develop optimum sampling plans.

Green beans have been selected for study because of (1) the importance of the green bean industry to the economy of Oregon and (2) the lack of uniformity in the employment of sampling procedures among green bean processing plants. Oregon leads the nation in the production and processing of green beans. In 1961 an estimated 24 per cent of the green beans for processing produced in the United States was grown in Oregon. This product returned to growers an estimated $13,363,000 in 1961 (18, p. 18-19).

The survey of sampling procedures of the fruit and vegetable processing industry in Oregon mentioned earlier, revealed that there are no standard methods for selecting or drawing the sample among the eighteen plants processing green beans. In all cases raw product arrived at the plant in large boxes, called tote bins, loaded on a truck. The tote bins are about 46" x 46" x 46" (inside dimensions) and hold approximately 1,000 pounds of green beans. It is equipped with a pallet arrangement so that it can be handled by a forklift truck. There were usually several tote bins per truck load with the average being about nine. In most cases there was no defined procedure for selecting the tote from which the sample was to be drawn. The decision was therefore left to the individual drawing the sample as to what tote would be drawn so that frequently the tote from a given position in the load was sampled repeatedly.

Only portions of the contents of the tote bin or bins selected were used as the sample. The sub-samples drawn from each tote selected were combined to make up the load sample. Eleven plants drew the sub-sample by hand from near the top of the tote box. In some plants beans from the tote box to be sampled were dumped into an empty tote box and the sub-samples extracted during the process. At two of the plants samples were drawn continuously by mechanical means as the beans entered the plant for processing.

The processing plants also vary with regard to the weight of samples drawn. Eight plants increased the size of samples drawn as the load size increased. The remainder drew a uniform total sample weight regardless of the load size. Ten plants drew increased numbers of sub-samples as load size increased. Of these, some reduced the weight of sub-samples drawn so that total weight of the load sample remained the same while others maintained the same sub-sample weight.

That sampling practices vary so widely can at least partially be attributed to the fact that the amount of variation that exists in raw product is unknown. The variability in product grades exists among tote boxes in the load and within each tote box. As a result a sample drawn from one segment of a tote box will yield a different estimate of value than one drawn from another portion in that or any other tote box in the load. The amount of variability among and within totes determines the amount of sampling necessary to achieve a satisfactory level of precision. Therefore, along with the estimated costs of sampling, a knowledge of the amount of variation is a prerequisite in determining an optimum sampling technique.

Utilizing information regarding variability and costs, the sampling technique that will give the least cost method of achieving a fixed level of precision or the greatest precision for a fixed cost is determined as the optimum sampling technique.

The specific objectives of the study with special application to the sampling of green beans are as follows:

1. To determine the variability among sampling units utilizing different sample weights and techniques.
2. To determine the cost of selecting and grading different numbers and weights of samples.

3. To determine an optimum sampling technique.

No attempt is made to develop or set grade standards for green beans in this study. Grade standards are of interest here only in that it is necessary to use some criterion for estimating variability in quality characteristics of green beans. United States Department of Agriculture standards have been used for this purpose. Any other set of grades could be used in successfully accomplishing the purposes of the study (8, p. 1-7).

CHAPTER 2

THEORETICAL FRAMEWORK

Probability Sampling

Most raw product sampling performed by processing plants in Oregon is done in a manner that might be termed "purposive selection." This method of sampling involves drawing a sample because it is convenient or thought to be representative. It also may be done in what is only thought to be a random manner such as "drawing a unit from a different place on the load every time." The sample is drawn at the discretion of the individual doing the sampling so that a known and stated method of selecting the sample does not always exist. Although estimates may be derived from a sample drawn in this manner, the frequency distribution for the estimates of all possible samples cannot be calculated as in probability sampling theory (3, p. 7).

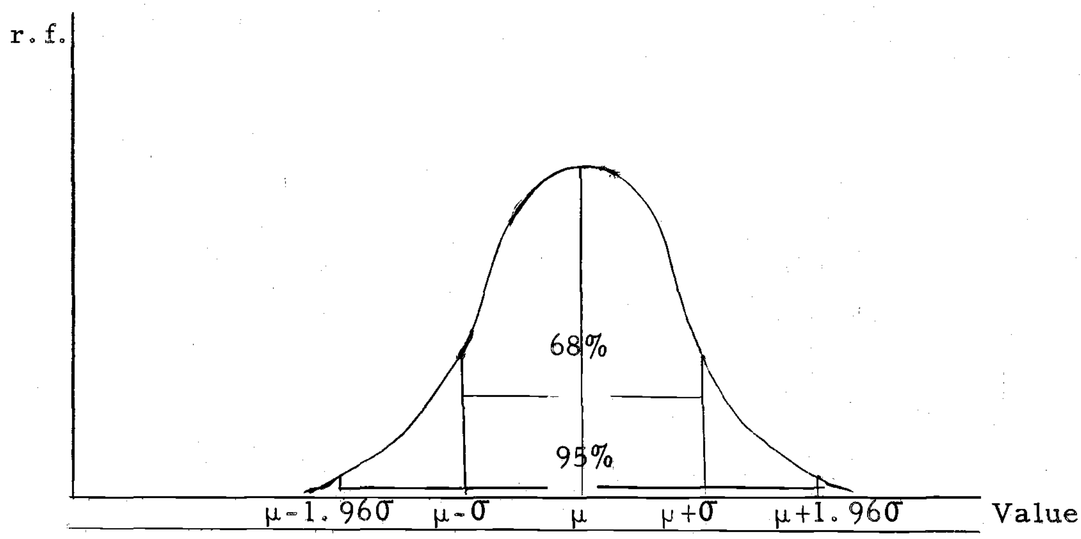
The frequency distribution of all possible samples can be calculated only when certain procedures are used in sampling (3, p. 6). These procedures are those of probability sampling and are as follows:

1. There is a set of samples comprising the population and the number of possible samples is known.
2. The probability of selecting any given sample is known and is determined by the method used in drawing the sample. In simple random sampling the probability of selecting any given sample is equal to one divided by the number of samples. Thus, all samples have an equal chance of being drawn in simple random sampling.
3. The sample is drawn to yield a statistic for which the method of computation is known and stated. The statistic may be the mean, total, or some other value derived from the sample.

When a sampling procedure with the above qualities is applied repeatedly to the population the frequency distribution of the sample estimate can be calculated so that a sampling theory can be applied (3, p. 6-7).

The estimates of value derived from a large number of samples of the same size taken from a raw product population are assumed to follow the normal distribution (12, p. 6, 33). This frequency distribution is described by a bell-shaped curve with the values clustered around the central point which is the mean (μ). The spread of the normal distribution is measured by the standard deviation (σ).

Figure 1. NORMAL DISTRIBUTION.



of the values which comprise it.[1/] The area under the normal curve is equal to 1.0 indicating that the curve describes estimates of all possible samples from a given population. The probability of a given event or events occurring can be determined by comparing the sub-area inscribing the event or events with the total area. The resulting proportion is equal to the probability of occurrence. For example, 68 per cent of the values fall within plus or minus one standard deviation of the mean. Ninety-five

---

1/ The standard deviation is a measure of the variability and is the square root of the variance.

per cent of the values comprising the normal distribution lie within plus or minus 1.96 standard deviations of the mean. Tables have been computed giving the area under the normal curve (12, p. 517).

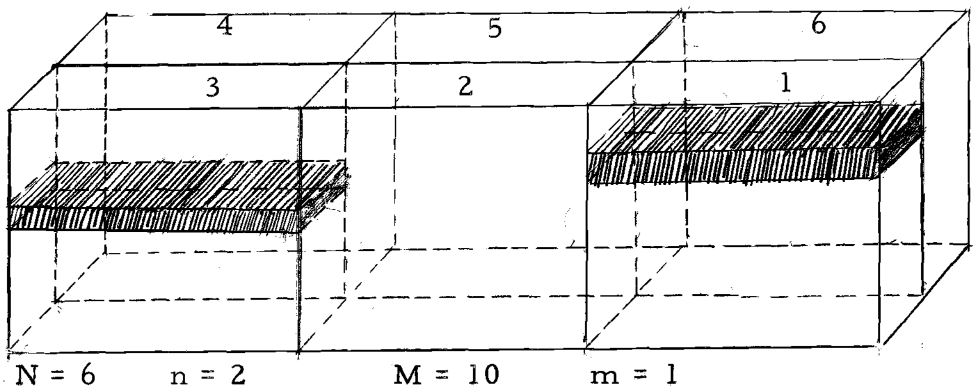
If the standard deviation of the sample estimates is known, the precision of the sample statistic can be stated as a confidence interval. The confidence interval is a range of values within which the true value is expected to lie with a given confidence. For if 95 per cent of the estimates of the normal distribution fall within plus or minus 1.96 standard deviations of the true mean then conversely the true mean will lie within plus or minus 1.96 standard deviations of the sample statistic with 95 per cent confidence (12, p. 14-21).

The standard deviation of the distribution of estimates derived from repeated sampling must be known to state the precision of the sample estimates. It is, therefore, necessary to determine the standard deviation for any raw product for which optimum sampling procedures are considered. It can be determined as will be shown later by repeated sampling of the population. The estimate of the standard deviation (σ) of all possible estimates is the standard deviation of the mean $\sqrt{V(\bar{\bar{y}})}$ of the sample estimates.

Two Stage Sampling

In estimating the standard deviation of the sample estimates it is necessary to set up a procedure for sampling. The nature of the population being sampled influences the way in which the samples will be drawn. Green beans are delivered in tote boxes on trucks and because it is not necessary to sample all produce only one or several totes may be drawn from the truck. Again it may not be necessary to sample all raw product from the box so only a portion of the 1,000 pounds of green beans is drawn. Samples drawn in this manner are two stage samples. The sample from the totes in a load is the first stage or primary unit and the sample of product from the tote is the second stage or secondary unit. In Figure 2 there are six primary sampling units (N) from which two are drawn as a sample (n).

Figure 2. A POPULATION OF RAW PRODUCT ILLUSTRATING PRIMARY AND SECONDARY SAMPLING UNITS.



Within each primary unit there are ten secondary sample units (M) of which one is drawn as a sample (m). The truck load of N = 6 is the population from which the sampling is to take place.

There are several methods by which the sampling units may be selected. Three of the basic and more commonly used methods are simple random sampling, systematic sampling and stratified sampling. These methods have been developed to achieve maximum precision of the sample statistic when applied to particular populations (3, p. 5).

In stratified sampling the population is divided into groups from which the samples are drawn. Stratified sampling seeks to increase precision by subdividing a heterogenous population into sub-populations that are homogenous (3, p. 65). Systematic sampling is defined by drawing units at regular intervals throughout the population. There are particular advantages in the ease of drawing the samples in a systematic manner. It also can under certain circumstances, yield a more precise estimate (3, p. 160-188). Simple random sampling is employed when units from the population are drawn at random. This method of sampling is employed when each item has equal chance of selection. These methods of sampling can be used and combined in two-stage sampling (3, p. 231-232; 9, p. 1-87).

Because of the assumed random nature of the population the simple random sampling procedure was one of two methods used in selecting appropriate sample units. Some processing plants use a mechanical device that draws the secondary sample of green beans continuously from the primary sample unit. This technique is essentially systematic in nature and also was utilized in the study. The sampling procedures applicable therefore, are the completely random selection of primary and secondary units and the random selection of the primary unit combined with systematic selection of the secondary unit.

Random Sampling

The number of primary units drawn as a random sample from the load is equal to n. The number of secondary units drawn as a random sample from a primary unit is equal to m. The estimate derived from a secondary sampling unit can be identified as an observation y_{ij}, where i refers to the number of the primary unit from which the secondary unit was selected and is some number, 1. . . n and j refers to the number of the secondary unit drawn from the ith primary unit and is some number 1. . . m. The observation y_{ij} is assumed to be composed of the following components

$$y_{ij} = \bar{\bar{Y}} + u_i + w_{ij} \tag{2.1}$$

where $\bar{\bar{Y}}$ is the general mean of all secondary sampling units, u_i is the error component associated with the primary unit, with mean equal to zero and variance equal to S_u^2. The component u_i is constant for all observations in a primary unit. The component w_{ij} is the error component associated with the secondary unit within a primary unit, with mean equal to zero and variance equal to S_w^2 (3, p. 217, 231; 9, p. 152, 328). All observations taken in sampling a load can be expressed as in equation 2.1 above. The amount of variation existing among secondary sampling unit observations is expressed as the variance of the error components u_i and w_{ij}.

Expressed formally, the variance of the sample mean $V(\bar{\bar{y}})$ of the secondary sampling units is

$$V(\bar{\bar{y}}) = \frac{S_u^2}{n} + \frac{S_w^2}{nm} \tag{2.2}$$

where S_u^2 and S_w^2 are defined as above, n is the number of primary units taken and nm is the number of secondary units drawn from a load. The variation of estimates derived from repeated sampling is therefore dependent upon the amount of variation existing among primary and secondary units and the number of primary and secondary units taken. Thus, it is necessary to estimate S_u^2 and S_w^2 to determine the variance of the mean $V(\bar{\bar{y}})$. This is done by the analysis of variance which is shown in Table 1 (3, p. 219).

Table 1

ANALYSIS OF VARIANCE OF SAMPLE ESTIMATES

Variance	df	Mean Square	Estimates of
Among primary units	n-1	$s_b^2 = \frac{\sum(\bar{y}_i - \bar{\bar{y}})^2}{n-1}$	$S_w^2 + m\,S_u^2$
Among secondary units	n(m-1)	$s_w^2 = \frac{\sum(y_{ij} - \bar{y}_i)}{n(m-1)}$	S_w^2

In Table 1 the term $\bar{y}_i$ as used in the computation of the mean squares is the mean of the secondary units in the ith primary unit. The general mean for all observations is $(\bar{\bar{y}})$.

The component of variation of among secondary units (S_w^2) is estimated directly as shown in Table 1 as the among secondary unit mean square. The component of variation of among primary units (S_u^2) is estimated by subtracting from the among primary unit mean square (S_b^2) the estimate of S_w^2 and dividing the result by (m) the number of secondary units drawn from a primary unit (3, p. 219).

The estimates of the components of variation determined by the analysis of variance are used in equation 2.2 to determine the variation of the sample mean. The variance of the sample mean resulting from increasing or decreasing the number of primary and secondary units taken can then be determined by varying n and m in equation 2.2 (5, p. 137; 3, p. 222-223). There are two ways in which the variance of the mean of the sample estimate $V(\bar{\bar{y}})$ can be increased or reduced--by changing the number of primary and secondary units taken, or by changing the weight of the secondary unit. A change in the weight of the secondary unit is a change in the amount of raw product taken as a secondary unit and affects the estimates of variation among primary and secondary units.

In the above discussion it was assumed that the total number of primary units and secondary units within a primary unit were large in respect to the number actually taken as a sample. If this were not the case and the fraction n/N and m/M were greater than 5 per cent the finite population correction would be applied. To do this the relationship of the population and sample sizes would need to be introduced into equation 2.2. This is shown in equation 2.3 as

$$V(\bar{\bar{y}}) = \frac{N-n}{N} \frac{S_u^2}{n} + \frac{NM-nm}{NM} \frac{S_w^2}{nm} \qquad (2.3)$$

The effect of the finite population correction is to reduce the variance of the sample mean when the sampling procedure is applied to a small population (3, p. 17).

The square root of the variance of the sample mean is the standard deviation of the sample mean and can be used to make a statement of precision as a confidence interval. As was previously stated the sample estimates were assumed to follow the normal distribution. An estimate of the standard deviation permits statements of probability based on knowledge of the area under the normal curve. Confidence in a sample estimate arises from this probability so that precision may be stated as

$$d = Z \sqrt{V(\bar{\bar{y}})}$$

where d is the confidence interval or stated precision of the estimate. Z is the value under the normal curve corresponding to the per cent of confidence to be achieved and $\sqrt{V(\bar{\bar{y}})}$ is the standard deviation of the sample mean. The confidence interval may be given in terms of: (1) an absolute value in units in which the estimate is stated or (2) a percentage of the unit value of the estimate.

The statement of precision provides a means of comparing sample procedures for a given level of confidence. The effect on precision of varying the number of primary and secondary units taken and the weight of the secondary unit can be stated. These are compared in considering various sampling techniques.

Random Continuous Sampling

The second method of sampling considered in this study is the continuous method of drawing the secondary unit and is essentially a systematic sample applied at the second stage. Assuming the product is in a random order the estimate derived from the sample drawn in this manner can be used as in the case of a sample drawn in a random manner (3, p. 168-172). However, when but one secondary unit is drawn, as is true of the continuous sampling method used by the processing industry, the among secondary unit variation S_w^2 cannot be estimated. To determine an estimate of the variance of the mean $V(\bar{\bar{y}})$ the among primary unit mean square is estimated and divided by the size of the primary sample (n) drawn (3, p. 225). The determination of the primary unit mean square follows from the analysis of variance in Table 1 as shown for completely random sampling.

Precision can be stated as a confidence interval as in the previous case. As the number of secondary units drawn by this method is limited to one the stated level of prevision can be affected only by changing the number of primary units drawn or changing the weight of the secondary unit.

Estimating Costs of Sampling

Costs of sampling were synthesized from economic engineering studies made in a number of bean processing plants. Time and production studies were made of the particular jobs and machine operations involved in sampling. Labor and equipment standards of performance were developed and were used in estimating the physical requirements of sampling. Interviews with management and supervisory personnel provided further information regarding the allocation of labor and equipment time to the sampling operation and in providing applicable wage rates and factor costs.

The number of workers and items of equipment utilized in the sampling operation are relatively few and as a result the process of estimating sampling costs is rather simple and straightforward. Only a brief description of the type of costs estimates used in the study are presented here. The theory and methodology of economic-engineering studies is explained by French _et al._ (8, p. 573-721).

Costs estimated by the method described above were expressed as

$$C = C_u n + C_w nm \tag{2.5}$$

where total cost C is composed of C_u--costs associated with the selection and drawing of the primary units (n), and C_w--costs of obtaining the secondary units (m).

Estimating Optimum Samples

To determine the optimum number of primary and secondary units to be taken in a sample, it is necessary to use one of two criterion: (1) achieve the least possible cost with a given level of precision or (2) achieve the highest possible level of precision for a given total cost. Either of the above criterion can be used in determining an optimum technique. The choice of which one to use depends upon which is considered the more critical--precision or cost.

In order to derive a function that will solve for the optimum sampling procedure it is necessary to combine the variance and cost functions. In two-stage sampling the variance function is

$$V(\bar{\bar{y}}) = \frac{S_u^2}{n} + \frac{S_w^2}{nm} \qquad (2.6)$$

The general cost function is

$$C = C_u n + C_w nm \qquad (2.7)$$

The number of primary and secondary units drawn in the sample is then chosen to minimize

$$V(\bar{\bar{y}}) + \lambda(C - C_u n + C_w nm) \qquad (2.8)$$

By substituting in the function of the variance of the mean for $V(\bar{\bar{y}})$, differentiating with respect to n and m and substituting like quantities the function for the optimum number of m is found to be (3, p. 225-226)

$$m_{(opt)} = \frac{S_w}{S_u} \sqrt{\frac{C_u}{C_w}} \qquad (2.9)$$

After $m_{(opt)}$ has been determined it is substituted into the variance function 2.6 or the cost function 2.7, depending upon which of the two quantities is fixed, to find the optimum number of n.

In estimating the optimum n and m from sample data the estimates of among primary and among secondary unit mean square as determined in the analysis of variance, Table 1, can be used. The estimate of the $m_{(opt)}$ is

$$m_{(opt)} = \frac{S_w \sqrt{m}}{\sqrt{S_b^2 - S_w^2}} \sqrt{\frac{C_u}{C_w}} \qquad (2.10)$$

The value found for m may not be a whole integer in which case the rounding must be done judiciously since simple rounding is not applicable here. If $m_{(opt)}$ is between two integers m and m + 1, m + 1 should be used if $m^2_{(opt)}$ is more than m(m + 1) (3, p. 226). The optimum number of secondary units will be large if the size of the among secondary unit variation is large compared to the among primary unit variation. Similarly the optimum number of secondary units (m) will be large if the cost of drawing the primary unit is high relative to the cost of drawing the secondary unit. Again the optimum number of primary units is found by substituting $m_{(opt)}$ in equation 2.6 or 2.7 depending upon whether variance or cost is the fixed quantity.

CHAPTER 3

PROCEDURES USED IN THE STUDY

Experimental Sampling

To determine optimum raw product sampling techniques it is first necessary to estimate the variation existing within the raw product. Repeated sampling was used to estimate the variation existing among primary and secondary sampling units in green beans. The sampling study was designed so that a comparison might be made of: (1) the precision achieved by various numbers of primary and secondary sampling units; (2) the effect of the weight of the secondary units on among primary and among secondary unit variation, and (3) the completely random and the random systematic (continuous) method of obtaining the secondary sampling unit. Because of limited resources only two sizes of secondary units were studied in the completely random technique, and but one size of secondary unit was drawn by the continuous method.

The design of the study showing the weight of the secondary units drawn and the number of primary and secondary units taken is shown in Table 2. There were three sampling schemes employed to estimate the components of variation.

Sample Scheme A consisted of drawing two 10-pound sub-samples from each of two totes. Two 30-pound sub-samples were drawn from each of two totes in Sample Scheme B. In Sample Scheme C a single 60-pound continuous sub-sample was drawn from each of two totes.

Table 2

SAMPLING SCHEMES USED IN ESTIMATING VARIATION

Sample Scheme	No. of Loads Sampled	No. of Totes Per Load	No. of Secondary Units per Tote	Weight of Secondary Unit (pounds)
A	38	2	2	10
B	53	2	2	30
C	32	2	1	60

The samples were drawn at selected processing plants in the Willamette Valley during the 1961 and 1962 processing seasons. The processing plants permitted the use of their facilities to draw the samples.

Random sampling was used to select the primary units drawn in all three sampling schemes. In Sample Schemes A and B the sub-samples were also selected by random sampling. The use of a random numbers table was employed to achieve a random selection so that each tote and sub-sample would have equal likelihood of being selected. In sampling Scheme C the beans in the continuously selected secondary unit were drawn by a mechanical device and were assumed to be randomly selected.

Selection of the sample of totes was made either while the beans were on the truck or after the beans had been taken from the truck and were stored prior to processing. Totes in the load in either case were numbered in a precise prearranged order so that the particular digit drawn from the random numbers table indicated the tote to be sampled. After the totes had been selected they were marked, so that identification for sampling could be made, and set aside awaiting sampling. The totes drawn from the load were dumped and the contents passed over a conveyor belt into an empty tote box. During this operation the sub-samples were drawn. The beans not drawn in the sub-sample were then processed with the remainder of the load.

In drawing the 10 and 30 pound sub-samples the totes were assumed to be divided into tenths. The tenth from which a sub-sample was to be drawn was selected by random numbers and the sub-sample drawn by hand from within that tenth.

This ensured random selection of the sub-sample. As the samples were drawn from the appropriate section or sections they were marked for proper identification. In drawing the continuous sub-sample the beans were drawn from the tote as they passed into the processing line.

Grading of the samples involved removal of defective or cull beans by hand and sizing by machine of the remainder of the product. Beans drawn in the sampling experiment were graded and sized under the supervision of an Oregon State Department of Agriculture inspector. United States Department of Agriculture Standards for Snap Beans for Processing were used as the basis of making grade determination (18, p. 1-7). Cull beans removed by hand were those that were broken, damaged or had rust or other defects.

After removal of the culls, the beans were sorted into various size classifications according to diameter, by a machine designed for the purpose and set to industry standards. There were four different size classifications as shown in Table 3. Beans smaller than 10/64 inches in diameter were removed as immature beans (18, p. 2). Samples taken for the study were graded at a cooperating plant that permitted the use of its grading facilities. All samples were sized through the same machine to insure consistent results.

Table 3

SIZE CLASSIFICATIONS FOR GREEN BEANS AND THE VALUE IN DOLLARS FOR EACH GRADE, 1962

Company's Grade	U. S. Sieve Size	Diameter of Bean	Price per Ton (dollars)
No. 1	1, 2, 3	under 21/64"	170.00
No. 2	4	21/64 - 24/64"	155.00
No. 3	5	24/64-27/64"	110.00
No. 4	6	27/64-29/64"	70.00
Culls			0
Immature			0

After grade and size determination was made the weights of each size classification were recorded for each sample. The weights were then converted into a per cent of total weight for each size classification. Based upon a uniform weight base the percentages were converted into dollar values for each grade of beans. The total dollar value of the load calculated in this manner became a single sample estimate to be used in the analysis of variance computations. These sample estimates were assumed to be normally distributed about the mean of all possible sample estimates (9, p. 33).

Estimating Costs of Sampling

As was indicated earlier, sampling costs were synthesized from economic-engineering studies made in a number of cooperating bean processing plants in the Willamette Valley of Oregon.

Estimated costs of labor required for sampling green beans are based on (1) labor standards for each job involved in sampling and (2) typical wage rates being paid by the industry for those jobs. In the development of labor standards, time and production studies were made to determine the amount of time required per sample for each job. Total time requirements per sample were then converted to numbers of samples per man-hour. Labor standards calculated in this manner represent the output a worker is able to achieve under conditions of continuous operation and a reasonably efficient use of his time. Labor standards were then used to compute the number of workers or fractions of workers required for each job. Current wage rates plus fringe benefits were used in converting the physical labor requirements to costs.

It was assumed that a worker's time could be utilized in raw product sampling as he was needed. In other words, only that portion of a worker's time which was spent in sampling green beans was chargeable to the sampling operation.

The numbers of items of equipment required for various operations were estimated using equipment standards developed from time and production studies, manufacturers specifications, interviews with plant personnel, and secondary data sources. These physical equipment requirements were first converted to investment costs through the use of current equipment replacement costs (f. o. b. equipment manufacturer), transportation costs to the plants, and installation costs. Annual costs of equipment, both fixed and variable, were then estimated. Fixed equipment costs include an allowance for depreciation, taxes, insurance, interest on investment, and fixed repairs and maintenance. Variable costs include the estimated costs of variable repairs and maintenance, power and fuel.

Total season's equipment costs were computed on the basis of estimated number of hours of use per season. All or a portion of the costs were then charged to the sampling operation based on the amount of time each item of equipment was used in the sampling of green beans.

Estimated sampling costs for this study include only labor and equipment costs. No attempt has been made here to allocate to sampling such indirect costs as building, management and supervisory labor.

Total sampling costs estimated in the manner described above were divided into the two components of equation (2.7). These components were (1) costs attributable to the drawing of the primary unit and (2) costs of drawing and grading the secondary sample unit. Grading costs depend on the number of pounds being graded and are a function of the number of totes as well as the number and weight of the sub-samples drawn in the sample.

Determination of Optimum Sampling Methods

Estimates of the variance existing in the raw product and costs incurred in drawing various numbers of samples of different weights, estimated as outlined above, can be used in the determination of an optimum sampling procedure. The optimum sample size is calculated so as to yield a given level of precision at the least possible cost or conversely the greatest precision for a fixed cost.

CHAPTER 4

RESULTS OF ANALYSIS

Estimates of Variation

Resources and time available for the study did not permit the drawing of samples from a large number of plants. It also was not possible to repeat the sampling program for more than two operating seasons. The assumption must be made, therefore, that data pertaining to the variance estimates of the three sampling schemes from the plants and years actually studied are representative of all bean processing plants and all years. Based on this assumption the pooled analysis of variance for years and plants included in the study for each sampling scheme are shown in Table 4. The original data and estimates of variation from which the pooled variance estimates were obtained are shown in Appendix Tables 1, 2, 3 and 4. The pooled estimates of the variance components for the ten, thirty and sixty pound samples indicate the amount of variation that exists at each level of the sampling method (Table 4). As is illustrated in Table 4 the estimate of among tote variation (S_u^2) for the pooled ten pound sub-sample (4.56)

Table 4

MEAN SQUARES AND ESTIMATES OF VARIATION FOR 10, 30, AND 60 POUND SAMPLES.

Estimate of	10 Lb. Sub-sample	30 Lb. Sub-sample	60 Lb. Sub-sample
Primary unit mean squares	33.02	29.18	14.42
Secondary unit mean squares	23.89	15.94	
S_u^2	4.56	6.62	14.42
S_w^2	23.89	15.94	

is smaller than the component of among sub-sample variation (23.89). A similar relationship exists in the 30 pound samples as the estimate of S_u^2 (6.62) is smaller than that of S_w^2 (15.94). These estimates indicate that the among tote variation is a smaller component of total variation than the among sub-sample variation. Also, for the 30 pound sampling scheme the estimate of S_w^2 is not as large relative to the estimate of S_u^2 as is the case in the scheme using the 10 pound sub-sample. Although the result is anticipated--as a larger sub-sample is taken it is expected to yield a lower estimate of among sub-sample variation--the decrease in the estimated variation is not proportional to the increase in the size of sub-sample taken.

The mean square estimate of among primary unit variation is smaller for the 60 pound sampling scheme than for the 10 and 30 pound schemes. As only one sub-sample was taken in scheme C there was no secondary unit mean square.

The estimates of variation for each size of sub-sample when inserted into equation 2.2 yield the variance of the population of sample estimates for each sampling scheme. For example, using the components of variation as computed from the 10 pound sub-samples the variance of the sample mean is

$$V(\bar{\bar{y}}) = \frac{4.56}{n} + \frac{23.89}{nm}$$

The square root of the variation $V(\bar{\bar{y}})$ is the estimate of the standard deviation of the sample estimates developed by repeated sampling of the population. The finite population correction was not used in the computation of the variance of the sample means as the effect of applying it is small. For example, assuming in the above illustration a small load size of six primary units and a sample of two totes and three sub-samples per tote the estimate of the variance of the mean without the finite population correction is

$$V(\bar{\bar{y}}) = \frac{4.56}{2} + \frac{23.89}{6} = 6.22$$

With the finite population correction the estimate of variance of the mean is

$$V(\bar{\bar{y}}) = \frac{6-2}{6} \frac{(4.56)}{2} + \frac{600-6}{600} \frac{(23.89)}{6} = 5.28$$

Then the estimate of the standard deviation without the finite population correction is 2.49 and with the finite population correction it is 2.30. In most cases the load size in reality would be larger than the one used in this example, thus further minimizing the difference in the estimates.

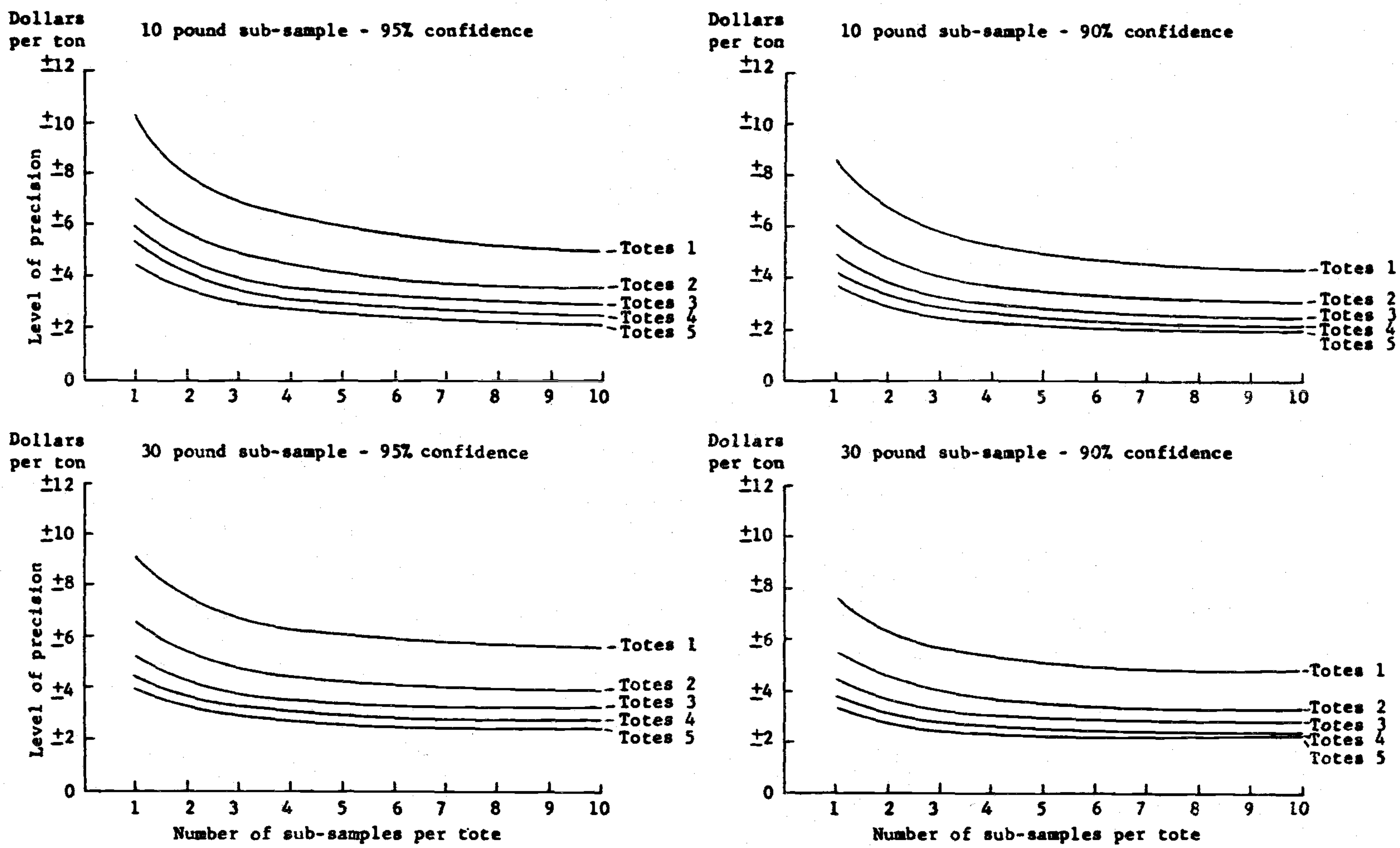
It follows from the variance function that the estimated standard deviation depends upon the size of sample drawn. As larger numbers of totes and sub-samples are drawn the estimated standard deviation will decrease. This means that the estimated values from repeated sampling will tend to draw closer to the mean.

The level of precision can be stated as an interval within which the true value of product lies with a given degree of confidence as in equation 2.5. This interval is expressed in terms of dollars per ton, enclosing the true value of the beans delivered. The levels of precision achieved by drawing various numbers of totes and sub-samples for the 10 and 30 pound sub-samples have been computed at the 95 and 90 per cent confidence levels. These are illustrated in Figure 3.[1/] Upon selection of a dollar interval such as ±$5.00 within which the true value per ton for the load is expected to fall, the number of n and m that would achieve this desired interval and confidence level can be determined.

Figure 3 illustrates that for a given sampling scheme the level of precision can be increased by increasing either the number of tote bins or the number of sub-samples drawn. The decrease in the length of the estimated confidence interval is greater when the number of totes drawn is increased than that achieved when the number of sub-samples drawn is increased by an equal number of units. The indication is that increasing the number of totes drawn is a more effective method of increasing the precision of the estimate than increasing the number of secondary units. In either case the precision gained decreases as the number of tote bins or sub-samples drawn increases.

[1/] The precision achieved by the 60 pound sub-sampling method is considered separately because of the continuous method of drawing the single sub-sample.

Figure 3: Combination of totes and sub-samples required to achieve selected levels of precision for given levels of confidence and sub-sample sizes. 1/



1/ The precision level is stated as a dollar interval within which the true value of a ton of green beans will lie with a given level of confidence. This assumes an unbiased sampling method.

The precision level achieved by drawing a 30 pound sample is greater than that of drawing a 10 pound sample (Figure 3) assuming the same number of primary and secondary units are drawn. The increase gained in this manner is not great, however. These relationships are illustrated in tabular form in Table 5 where the precision levels achieved by various combinations of totes and sub-samples are shown for sampling Schemes A and B.

The levels of precision achieved by drawing various numbers of tote bins in the 60 pound continuous sampling technique are shown in Figure 4. As the number of sub-samples per tote taken in this method was limited to one, the precision utilizing this sampling scheme can be increased only by increasing the number of primary units drawn. A comparison of the precision level achieved in each of the three sampling schemes by drawing a given number of tote bins and but one sub-sample per tote shows that sampling scheme C with the 60 pound sub-sample gives greater precision than schemes A or B with 10 and 30 pound sub-samples. This is understandable because the total amount of raw product drawn is greater as a result of the larger amount taken in a single sub-sample.

Table 5: Precision levels resulting from various combinations of totes and sub-samples for given levels of confidence and sub-sample sizes.

10 pound sub-samples — 95% confidence.

Level of precision (Dollars per ton)

Number of totes	1	2	3	4	5	6	7	8	9	10
5	±4.76	±3.56	±3.10	±2.85	±2.68	±2.56	±2.48	±2.41	±2.35	±2.31
4	±5.22	±4.50	±3.47	±3.18	±2.99	±2.86	±2.76	±2.69	±2.63	±2.58
3	±6.03	±4.58	±4.00	±3.67	±3.49	±3.24	±3.20	±3.10	±3.04	±2.98
2	±7.38	±5.62	±4.89	±4.62	±4.23	±4.14	±3.91	±3.80	±3.72	±3.65
1	±10.44	±7.96	±6.93	±6.36	±5.99	±5.71	±5.55	±5.38	±5.26	±5.12

Number of sub-samples per tote

10 pound sub-samples — 90% confidence.

Level of precision (Dollars per ton)

Number of totes	1	2	3	4	5	6	7	8	9	10
5	±3.92	±2.98	±2.61	±2.39	±2.25	±2.15	±2.08	±2.02	±1.97	±1.94
4	±4.38	±3.34	±2.91	±2.67	±2.51	±2.40	±2.31	±2.26	±2.21	±2.17
3	±5.06	±3.84	±3.36	±3.08	±2.93	±2.72	±2.68	±2.61	±2.55	±2.50
2	±6.20	±4.72	±4.10	±3.88	±3.55	±3.48	±3.28	±3.19	±3.12	±3.06
1	±8.77	±6.68	±5.82	±5.34	±5.02	±4.80	±4.66	±4.52	±4.42	±4.30

Number of sub-samples per tote

30 pound sub-samples — 95% confidence.

Level of precision (Dollars per ton)

Number of totes	1	2	3	4	5	6	7	8	9	10
5	±4.16	±3.34	±3.02	±2.85	±2.74	±2.67	±2.62	±2.57	±2.53	±2.51
4	±4.65	±3.74	±3.38	±3.18	±3.07	±2.98	±2.92	±2.87	±2.83	±2.80
3	±5.37	±4.33	±3.91	±3.69	±3.54	±3.45	±3.38	±3.32	±3.28	±3.24
2	±6.58	±5.75	±4.79	±4.51	±4.34	±4.22	±4.14	±4.06	±4.01	±3.97
1	±9.31	±7.49	±6.77	±6.38	±6.14	±5.97	±5.84	±5.75	±5.68	±5.62

Number of sub-samples per tote

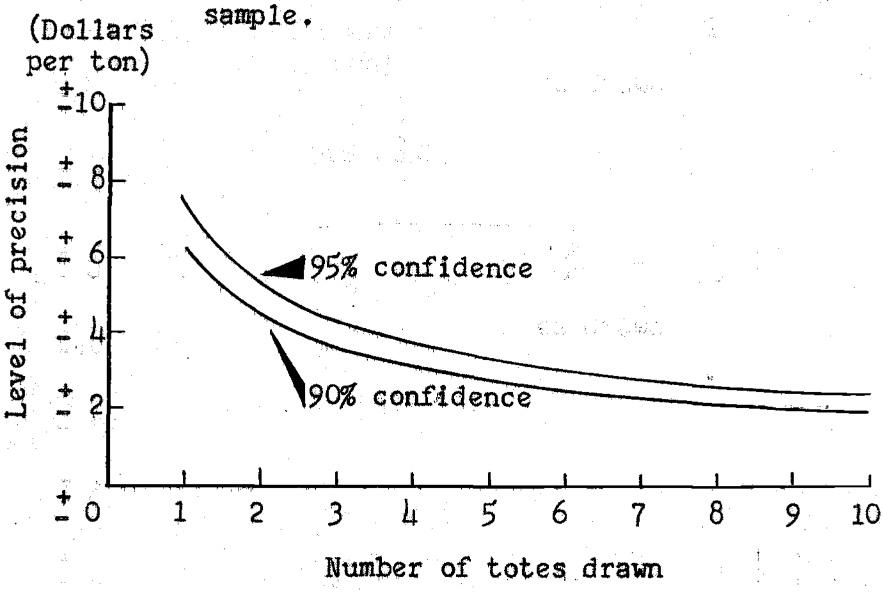
30 pound sub-samples — 90% confidence.

Level of precision (Dollars per ton)

Number of totes	1	2	3	4	5	6	7	8	9	10
5	±3.49	±2.81	±2.54	±2.40	±2.30	±2.24	±2.19	±2.16	±2.12	±2.11
4	±3.90	±3.14	±2.84	±2.67	±2.57	±2.50	±2.45	±2.41	±2.38	±2.35
3	±4.51	±3.63	±3.28	±3.09	±2.97	±2.89	±2.83	±2.79	±2.75	±2.72
2	±5.52	±4.83	±4.02	±3.79	±3.64	±3.54	±3.47	±3.41	±3.37	±3.33
1	±7.81	±6.28	±5.68	±5.35	±5.15	±5.01	±4.90	±4.83	±4.76	±4.71

Number of sub-samples per tote

Figure 4: The number of totes required to achieve selected levels of precision with 90% and 95% confidence, taking a single sixty pound continuously drawn subsample.



The 95 and 90 per cent confidence levels have been used in this study to state precision of various sampling schemes. However, other levels might be chosen by industry to provide satisfactory levels of confidence. The confidence level indicates, under repeated sampling of a normal population, the percentage of estimates falling within a given range of the mean. This percentage of all estimates is stated as a confidence level for a particular estimate derived from the sampling scheme.

Estimated Costs of Sampling

The sampling costs of the three sampling schemes--using 10, 30 and 60 pound sub-samples--have been estimated according to procedures presented in Chapter 3.

Labor requirements and costs for each of the three schemes are shown for various jobs involved in sampling (Table 6). Labor standards and costs of drawing a tote are similar for all three schemes. Requirements and costs vary, however, as the amount taken as a sub-sample varies. As can be seen in Table 6, cost of labor required to draw and grade a 10 pound sub-sample is estimated to be about 14 cents, for the 30 pound sub-sample, 43 cents, and for the 60 pound sub-sample, 85 cents. It is evident that labor costs for the three schemes are closely associated with the weight of the sub-sample drawn, and increase rapidly as the weight of sub-sample increases.

Major items of equipment used in the sampling operation include the fork lift truck for transporting and dumping tote bins, a hopper and conveyor for receiving the beans and drawing the sub-sample, a mechanical sizer for sizing the samples drawn and an accurate set of platform scales. Season's fixed and variable costs for these and other minor items of

Table 6. Labor standards and costs for drawing and grading green bean samples using two-stage sampling techniques and selected sizes of sub-samples.

Job 1/	Units for standards and costs	Labor standards	Estimated costs 2/
		(Man-minutes per unit)	(Dollars per unit)
	10 lb. Sub-samples		
Select and obtain primary sample unit	1,000 lb. tote bin	.738	.029
Dump primary sample unit	1,000 lb. tote bin	1.115	.044
Obtain secondary sample unit	10 lb. sub-sample	.503	.018
Grade and size secondary sample unit	10 lb. sub-sample	4.035	.124
	30 lb. Sub-samples		
Select and obtain primary sample unit	1,000 lb. tote bin	.738	.029
Dump primary sample unit	1,000 lb. tote bin	1.115	.044
Obtain secondary sample unit	30 lb. sub-sample	1.509	.054
Grade and size secondary sample unit	30 lb. sub-sample	12.105	.373
	60 lb. Sub-samples		
Select and obtain primary sample unit	1,000 lb. tote bin	.738	.029
Dump primary sample unit	1,000 lb. tote bin	1.115	.044
Obtain secondary sample unit	60 lb. sub-sample	3.018	.108
Grade and size secondary sample unit	60 lb. sub-sample	24.210	.746

1/ Job descriptions are presented in Appendix Table 6.

2/ Wage rates used to convert physical labor requirements to costs are: Select and obtain tote $2.365; dump tote $2.365; obtain sub-sample $2.15; grade and size sub-sample $1.85.

equipment are shown in Appendix Table 5. Fixed costs were computed on the basis of installed replacement costs and include an allowance for depreciation, taxes, insurance, interest on investment, and fixed repairs and maintenance (see footnote Appendix Table 5 for the magnitude of these allowances). Variable costs include the estimated costs of variable repairs and maintenance, power and fuel. Season's costs of each item of equipment have been converted to costs per hour based on the estimated number of hours of use per season. Costs per primary or secondary unit were then calculated according to the rate of sustained hourly capacity of each item of equipment. These are shown in Table 7 for the three sampling schemes. Here again it is apparent that costs per sub-sample for the mechanical grader and platform scales increase rapidly as the weight of he sub-sample increases.

Table 8 shows estimated labor and equipment costs incurred in drawing a primary sample, and in drawing and grading a sub-sample for each of the three sampling schemes. It is evident that grading costs constitute a large percentage of total costs of sampling.

Table 7. Equipment standards of performance and costs for drawing and grading bean samples using two-stage sampling techniques and selected sizes of sub-samples.

Item of Equipment 1/	Units for standards and costs	Equipment standards 2/	Estimated costs
		(Units per machine hour)	(Dollars per unit)
		10 lb. Sub-sample	
Fork lift truck	1,000 lb. tote	9.7	$.07
Hopper and conveyor	1,000 lb. tote	9.7	.02
Tote bins	1,000 lb. tote	9.7	.01
Handtruck	10 lb. sub-sample	14.9	.00
Lug boxes	10 lb. sub-sample	14.9	.00
Mechanical grader	10 lb. sub-sample	14.9	.10
Platform scales	10 lb. sub-sample	14.9	.02
		30 lb. Sub-sample	
Fork lift truck	1,000 lb. tote	9.7	.07
Hopper and conveyor	1,000 lb. tote	9.7	.02
Tote bins	1,000 lb. tote	9.7	.01
Handtruck	30 lb. sub-sample	4.5	.00
Lug boxes	30 lb. sub-sample	4.5	.00
Mechanical grader	30 lb. sub-sample	4.5	.31
Platform scales	30 lb. sub-sample	4.5	.07
		60 lb. Sub-sample	
Fork lift truck	1,000 lb. tote	9.7	.07
Hopper and conveyor	1,000 lb. tote	9.7	.02
Tote bins	1,000 lb. tote	9.7	.01
Hand truck	60 lb. sub-sample	2.2	.00
Lug boxes	60 lb. sub-sample	2.2	.00
Mechanical grader	60 lb. sub-sample	2.2	.62
Platform scales	60 lb. sub-sample	2.2	.14

1/ Fixed and variable equipment costs are shown in Appendix Table 5.

2/ Hours in operating season are: Fork lift truck 1,104; all others 368.

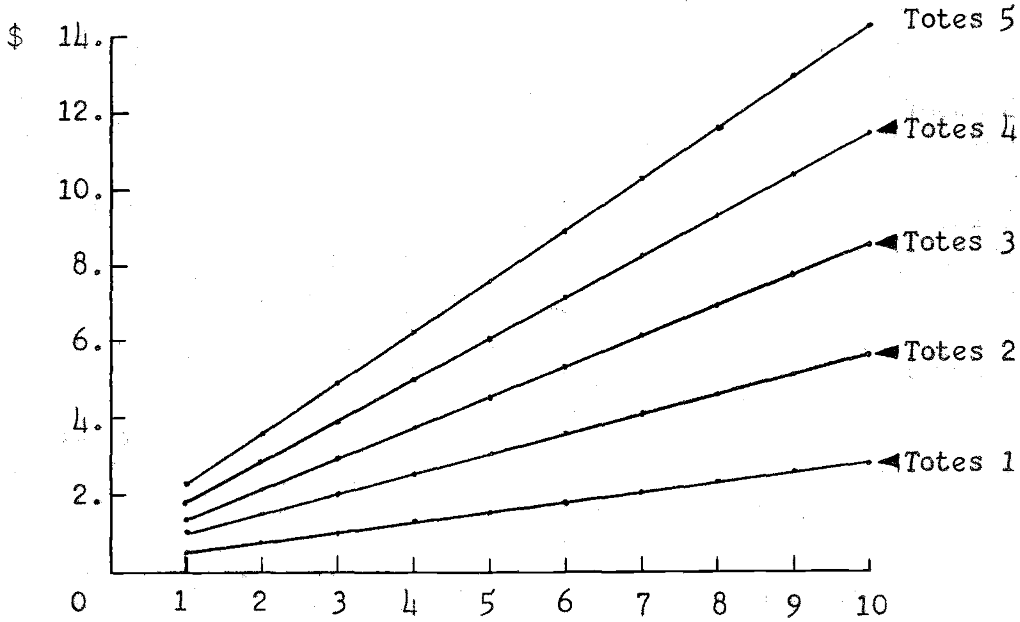
Table 8. Estimated costs of obtaining a tote and drawing and grading sub-samples of size 10, 30 and 60 pounds.

Element of cost	Estimated cost of drawing a primary unit	Estimated costs of drawing a sub-sample 1/			Estimated costs of grading a sub-sample		
		10 lb.	30 lb.	60 lb.	10 lb.	30 lb.	60 lb.
	(Dollars)	(Dollars)	(Dollars)	(Dollars)	(Dollars)	(Dollars)	(Dollars)
Labor	$ .07	$.02	$.05	$.11	$.12	$.37	$.75
Equipment	.10	.02	.07	.14	.10	.31	.62
Total	$ .17	$.04	$.12	$.25	$.22	$.68	$1.37

1/ The ten and thirty pound sub-samples were drawn randomly by plant personnel; the sixty pound sub-sample was drawn mechanically in a continuous manner.

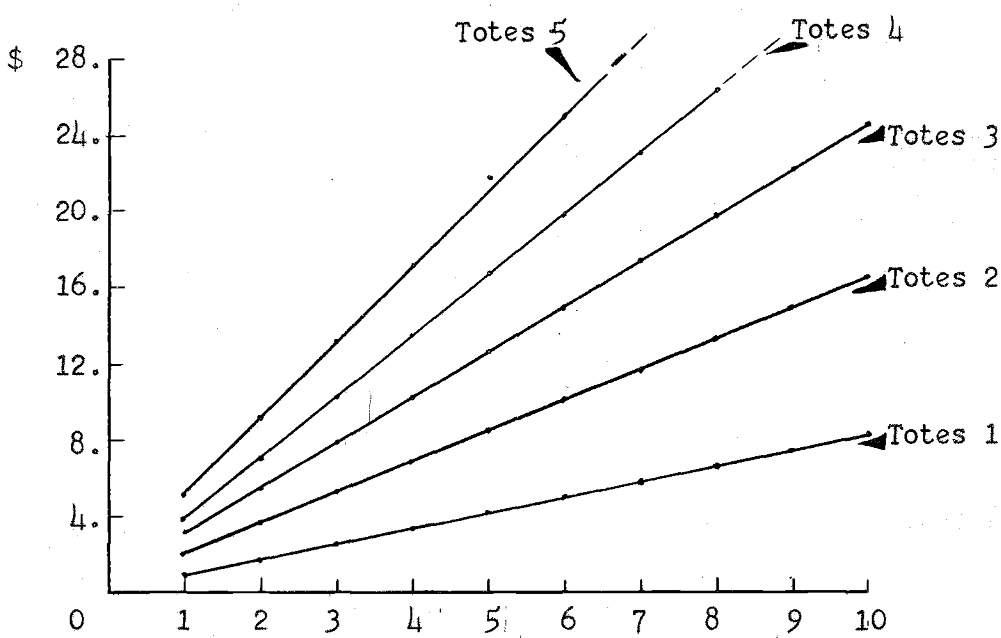
Utilizing the break-down of costs as presented in Table 8 and the cost function (equation 2.3), total costs of sampling, using various combinations of numbers of totes and sub-samples were determined for each of the three sampling schemes. These are shown in Figures 5, 6 and 7. Because the number of sub-samples for Scheme C was limited to one, only the number of totes were varied in Figure 7. For Schemes A and B both the number of totes and sub-samples have been varied. For reasons of simplicity total cost functions are presented here as continuous linear functions. In reality these functions would be discontinuous in nature because fractions of units, whether they be totes or sub-samples, are not drawn.

Figure 5: Total costs incurred in drawing samples consisting of various n and m using a ten pound sub-sample, 1962.



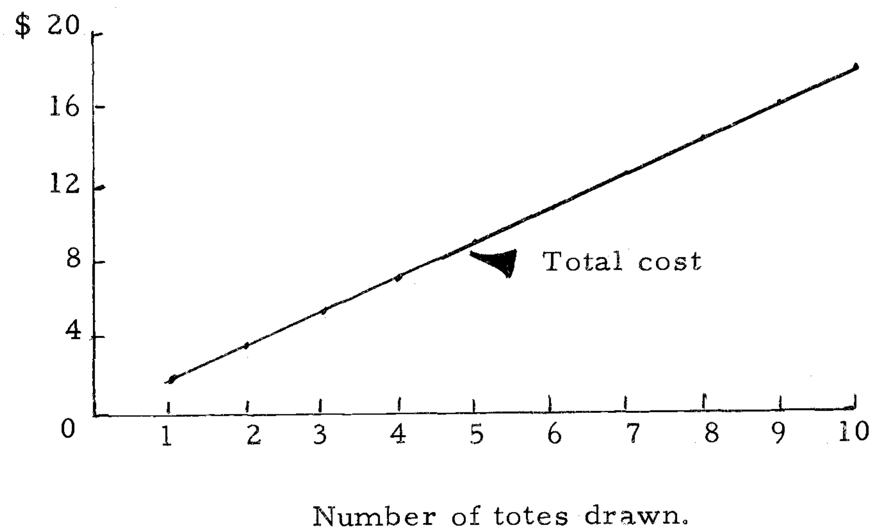
Number of ten pound sub-samples drawn per tote.

Figure 6: Total cost incurred in drawing samples consisting of various n and m using a thirty pound sub-sample, 1962.



Number of 30 pound sub-samples per tote.

Figure 7: Total cost incurred in drawing samples consisting of various numbers of totes using a 60 pound sub-sample, 1962.



Optimum Sampling

Variance and cost estimates derived in the previous sections are now brought together for the purpose of determining optimum sampling plans. The approach taken will consist of the following steps: (1) Optimum sampling plans will be determined for each of the three sampling schemes for selected levels of precision,[1/] (2) comparisons will be made of the optimum samples determined in (1) above for each of the schemes in order to determine the scheme which provides selected levels of precision at the lowest

[1/] It has been assumed that the level of precision is the more critical criterion for determination of optimum sampling procedures, and that costs of sampling play a rather minor role in decisions involving sampling procedures in processing plants.

costs, (3) generalizations will be made with regard to other sampling schemes involving sub-samples of weights other than those included in the study (10, 30 and 60-pounds), and (4) a comparison will be made of the simple random method of selecting the sub-sample as opposed to the continuous and mechanical means of drawing the sub-sample.

In determining optimum sampling plans for each of the three sampling schemes and for selected levels of precision, the optimum number of secondary units is first determined. Equation 2.10 is used in determining the $m_{(opt)}$. Among tote and among sub-sample variation as well as costs associated with each sampling unit are used in equation 2.10 in computing $m_{(opt)}$. The $m_{(opt)}$ computed in this manner provides an estimate of the number of sub-samples per primary unit regardless of the number of primary units drawn or the level of precision to be achieved. For Scheme A utilizing the 10 pound sub-sample the $m_{(opt)}$ is

$$m_{(opt)} = \frac{(4.8826)\sqrt{2}}{\sqrt{33.02 - 23.89}} \qquad \sqrt{\frac{.1744}{.2695}} = 1.8$$

which is rounded to two. For Scheme B utilizing the 30-pound sub-sample, $m_{(opt)}$ is

$$m_{(opt)} = \frac{(3.992)\sqrt{2}}{\sqrt{29.18 - 15.94}} \sqrt{\frac{.1744}{.8085}} = .72$$

which is rounded to one. Because only one sub-sample was drawn in Scheme C an optimum number of sub-samples could not be determined.

Having determined the optimum number of sub-samples to be drawn from a tote, the number of totes to be drawn to achieve various levels of precision was determined. This was done by inserting the values of $m_{(opt)}$ and the variance estimates into equation 2.6 and solving for n.

Table 9 shows optimum sampling solutions necessary to achieve selected levels of precision for each of the three sampling schemes. For example, in order to estimate the value per ton of green beans within a range of plus or minus five dollars with an assurance that 95 per cent of all possible samples that could be drawn would fall within that range, the optimum sample size of the three schemes would be: (1) Scheme A - two 10-pound sub-samples drawn from each of three totes, (2) Scheme B - one 30-pound sub-sample drawn from each of four totes and (3) one 60-pound continuously drawn

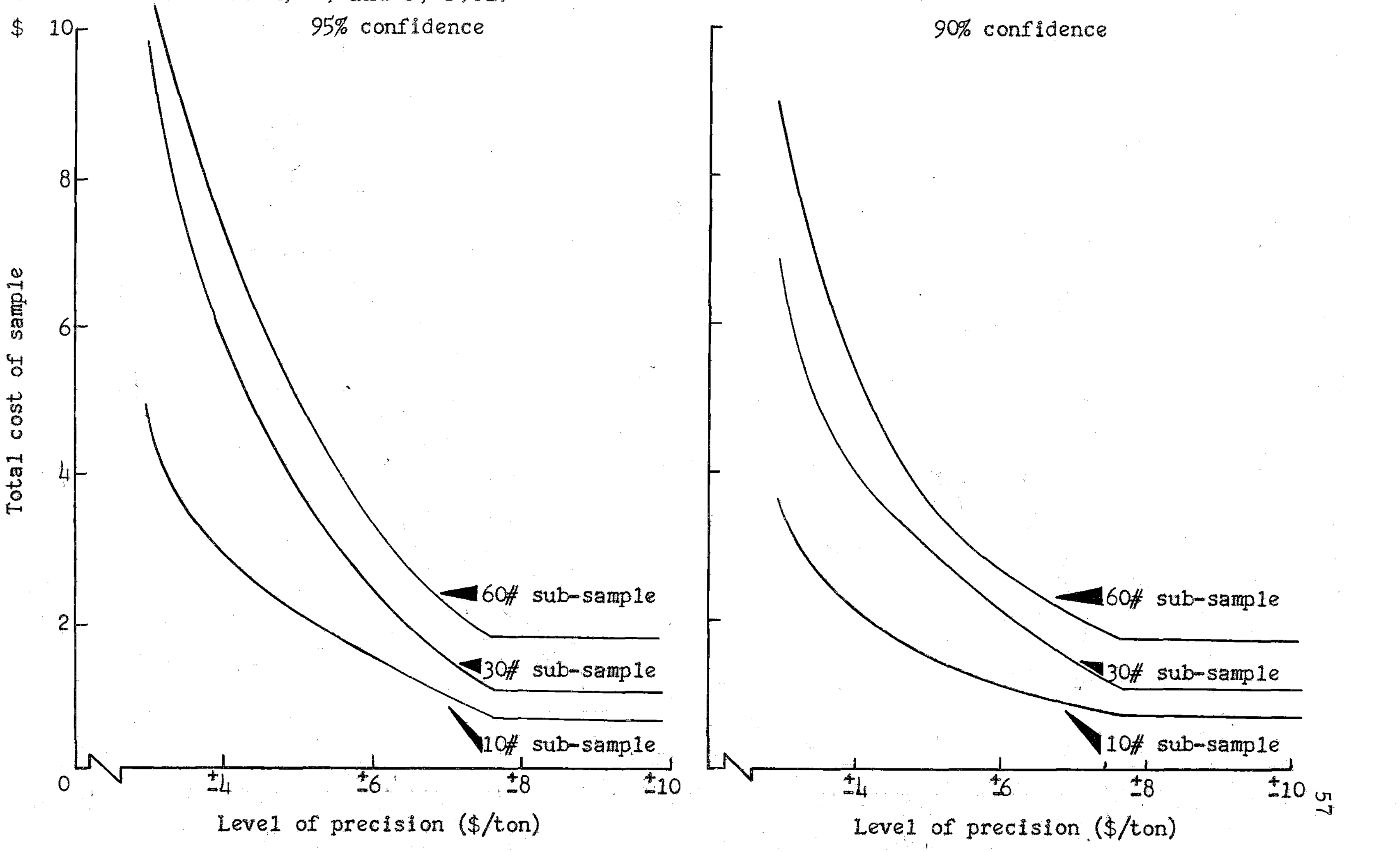
Table 9. Optimum sample sizes and costs for sampling schemes utilizing 10, 30 and 60 pound sub-samples for various levels of precision and confidence.

	Optimum sample solutions and costs								
	10-pound sub-sample			30-pound sub-sample			60- pound sub-sample		
Level of precision (plus or minus dollars per ton)	n	m	Cost (Dollars)	n	m	Cost (Dollars)	n	m	Cost (Dollars)
				95% Confidence					
± $10.00	1	2	.71	1	1	1.08	1	1	1.79
± 7.50	1	2	.71	1	1	1.08	1	1	1.79
± 5.00	3	2	2.14	4	1	3.93	3	1	5.38
± 4.00	4	2	2.85	6	1	5.90	4	1	7.17
± 3.00	7	2	4.99	10	1	9.83	6	1	10.75
				90% Confidence					
± $10.00	1	2	.71	1	1	1.08	1	1	1.79
± 7.50	1	2	.71	1	1	1.08	1	1	1.79
± 5.00	2	2	1.43	2	1	1.97	2	1	3.58
± 4.00	3	2	2.14	4	1	3.93	3	1	5.38
± 3.00	5	2	3.57	7	1	6.88	5	1	8.96

sample from each of three totes. Estimated costs incurred in drawing and grading each of these optimum plans would be $2.14, $3.93, and $5.38, respectively. Based on the range of precision and confidence considered in Figure 8, sampling Scheme A provides the least cost method of achieving given levels of precision and confidence. It should be pointed out here that each of the three schemes will yield a given level of precision with various combinations of n and m, as demonstrated in Figure 3, but the cost of drawing and grading these samples would have been greater than combinations presented in Table 9.

With regard to the question of sampling schemes using sub-sample weights different from those considered in this study, it appears, based on the magnitude of the variability obtained by the use of 10, 30 and 60-pound sub-samples, that the larger sub-sample tends to reduce among sub-sample variation. However, as has previously been pointed out sampling costs increase very rapidly as sub-sample size increases. Because of this, a more efficient method of reducing among sub-sample variation would be to increase the number of small sub-samples rather than to increase the size of the sub-sample. It would follow then that a sample scheme

Figure 8: Cost of achieving various levels of precision at 95 and 90 per cent confidence by use of sampling schemes A, B, and C, 1962.



using the 30-pound sub-sample would provide a fixed level of precision at a lower cost than one utilizing a 40-pound sub-sample. The same would be true for schemes using 10 and 20 pound sub-samples. The question is immediately raised about the possibility of the use of sub-samples smaller than 10 pounds. Data provided by this study do not permit a conclusion to be made regarding this matter. Further experimentation along these lines would be of interest.

In estimating $m_{(opt)}$ for sample scheme A and B the values determined were 1.8 and .72. Without rounding this would indicate about 18 and 21 pounds of beans taken per tote for each of the two methods. Based upon these results it might be concluded that 20-25 pounds of produce drawn from the tote would give optimum results. The question is raised as to whether this should be drawn randomly or in a continuous manner.

With regard to the comparison of the simple random method of selecting the sub-sample as opposed to the continuous method, logically it would seem that the continuous method is the more precise of the two. This is the case because a sample drawn continuously would contain produce from every section of the tote. A means by which the comparison can be made is to compare the precision levels achieved by drawing six ten pound sub-samples and two thirty pound sub-samples with that of a continuously drawn sixty pound sub-sample. If this is done, assuming a single tote is drawn in each case, the variance of the mean for sampling scheme A is 8.54, for sampling scheme B, 14.59 and for sampling scheme C, 14.42. The lower variation for Scheme A indicates greater precision. There is therefore some divergence in anticipated and actual results. The costs incurred in the three schemes are approximately equal as the same total sample weight is drawn in each case.

A comparison of the among tote variation for the continuous samples between 1961 and 1962 data might indicate why there is a divergence from the anticipated result in the comparison of the two methods. Appendix Table 4 shows that estimated variation is smaller in 1961 than 1962. Further research is necessary to determine the cause of the difference in the estimates. However, it can be concluded logically that the continuous method of sampling provides equal or greater precision than does the random method of sampling.

There is another consideration pertinent to a comparison of the two methods of obtaining sub-samples. It appears that mechanical selection of the sub-sample may provide some advantage in that the simplicity of selection of the sub-sample eliminates the human element of error. This might ensure that probability sampling procedures are followed.

CHAPTER 5

SUMMARY AND CONCLUSIONS

Raw product moving off the farm varies with respect to quality attributes. To determine the proper payment to the grower, it is necessary to ascertain the value of the product delivered. This may be done by either grading the whole load or sampling portions of the load. The latter method is least expensive and is commonly used in the Oregon fruit and vegetable processing industry.

The sampling scheme utilized must satisfy both growers and processors that it will accurately and precisely reflect the value of product delivered. Probability sampling theory permits the estimation of the precision of sampling procedures. It is preferred to judgment sampling or purposive selection because the precision level achieved by the latter methods of sampling cannot be estimated. These techniques, although commonly used, contain unknown error which cannot be measured.

There exists a lack of knowledge and application of probability sampling techniques in the processing industry. The result is a variety of sampling techniques for which corresponding levels of precision are not known. This study was designed to point out the need for probability sampling and to provide methodology requisite to solving some basic but common problems in raw product sampling.

As a means of achieving these purposes, probability sampling procedures were applied to green beans for processing. Green beans are delivered to the processing plant in tote bins loaded on a truck. Two-stage sampling is used with the first stage being the selection of totes as primary units and the second stage the selection of sub-samples of product within the tote as secondary units.

One of the requirements of probability sampling is that every tote or sub-sample has a known chance of being drawn. When the sample meets the requirements of probability sampling the frequency distribution of all possible sample estimates can be developed. The precision of the sampling scheme used can then be determined by examining the spread or variance of the distribution of estimates about the mean.

When the costs of sampling, utilizing a given sampling scheme, are integrated with the variability existing within raw product, the optimum sample is determined as the least cost method of achieving a stated level of precision or conversly the greatest precision for a fixed total cost. Given the criterion to be achieved--precision or cost--the optimum size of sample for a given sampling scheme can be computed.

Three sampling schemes utilizing probability sampling procedures were utilized in sampling green beans. The tote bins selected in each case were drawn by simple random sampling. In scheme A 10-pound sub-samples were selected randomly from the tote. In sampling scheme B 30-pound sub-samples were drawn randomly from the tote. In scheme C a single 60-pound sub-sample was drawn continuously by a mechanical device from throughout the tote. In sample schemes A and B a number of sub-samples can be selected from a tote bin; however, in sampling scheme C only one sub-sample per tote could be drawn as the beans passed into the processing line.

The sampling study, utilizing the three sampling schemes, was conducted during two processing seasons at cooperating processing plants in the Willamette Valley. The facilities of one of the cooperating plants were used to make grade determination on the samples taken in the study. The samples were graded by hand to remove beans with visual defects and the remaining produce was sized mechanically so that the sample value could be determined. Sample estimates of value were then used to determine raw product variation.

The amount of variation existing in the raw product was estimated by each of the three schemes. In sampling schemes A and B it was found that the among sub-sample component of variation contributed more to total variation than did among tote variation. The among sub-sample component of variation was smaller for sampling scheme B, with a 30-pound sub-sample than for sampling scheme A with a 10-pound sub-sample. The decrease in the variance component in scheme B was small, however, in relation to the increase in the size of the sub-sample over that taken in scheme A.

In sample scheme C only one sub-sample per tote was taken and therefore the within tote variation could not be computed. However, the among tote mean square estimate is used in determining precision. The precision achieved by the continuous sampling method with a 60-pound sub-sample is greater than when a similar number of totes but only one sub-sample is drawn using the other sampling schemes.

Costs incurred in sampling were estimated using the synthetic method of cost determination. Labor standards and current wage rates were used to determine the labor cost of each sampling operation. Both annual fixed and variable equipment costs were determined. Costs were then divided into those components associated with drawing the primary unit or tote and those associated with drawing and grading the secondary unit or sub-sample. The cost of drawing the tote is constant for all sampling schemes; however, the costs of drawing and grading the sub-sample depend upon the amount of raw product taken as a sub-sample. Sampling scheme C with a 60-pound sub-sample had the highest cost while sampling scheme A with a 10-pound sub-sample had the lowest cost.

Utilizing the information regarding variance and cost, the optimum sample sizes to achieve various levels of precision and cost were computed for each sampling scheme. For example to estimate the value of a ton of beans within $\pm$\$5.00 at 95 per cent confidence the following sample sizes are optimum for each sampling scheme: (1) in sampling scheme A--three primary sampling units and two secondary sampling units; (2) in sampling scheme B--four primary sampling units and one secondary sampling unit and (3) in sampling scheme C one continuous sub-sample from three primary units. The costs incurred in drawing each of these samples were \$2.14, \$3.93 and \$5.38 respectively. The difference in total season's sampling costs incurred by the use of these optimum samples for schemes A, B, and C are shown below. Assume, for example, that a processing plant received 15,000,000 pounds of green beans in an operating season and the average load contained 9,000 pounds or nine totes per load. There would be a total of 1,667 loads to be sampled during the season. Total estimated costs of sampling this number of loads by each of the three sampling schemes using the ±$5.00 level of precision and 95 per cent confidence would be: scheme A--$3,567, scheme B--$6,551 and scheme C--$8,968. In each scheme other combinations of primary and secondary sampling units could have been used to achieve a similar level of precision but only at higher costs. A comparison of three sampling schemes over a range of precision levels and their resulting costs revealed that sampling scheme A with the 10-pound sub-sample was the least costly method of achieving given levels of precision. Sampling scheme C was the highest cost method of those considered.

Based on the estimates of among tote and among sub-sample variation for the 10, 30 and 60-pound sub-samples it would appear that the more product taken as a sub-sample the smaller the among tote variation. However, costs increase at a rate that more than offsets the gain in precision. The 10-pound sub-sample therefore provides the lowest cost method of achieving a given precision within the range of sub-samples considered.

A comparison of the simple random and continuous sampling methods is difficult to make because sub-samples taken by each method were of different weights. A comparison of the three sampling schemes using the same total sample weight revealed, contrary to expectations, a greater precision for the random sampling technique drawing a 10-pound sub-sample. Although the data do not completely support the conclusion, logically it would appear that the continuous sample should supply equal or greater precision than a simple random sample of the same total sample weight. There is an advantage in the simplicity of mechanically drawing the continuous sub-sample because it tends to eliminate human error in the selection of the sub-sample. Further research with regard to the continuous sampling method of obtaining sub-samples would be useful.

BIBLIOGRAPHY

1. Barnes, Ralph M. Motion and time study. 4th ed. New York, Wiley, 1949. 665 p.

2. Black, Gary. Synthetic method of cost analysis in agricultural marketing firms. Journal of Farm Economics 37:270-279. 1955.

3. Cochran, William G. Sampling techniques. New York, Wiley, 1953. 330 p.

4. Dickmann, F. H. J. and G. B. Davis. Costs and efficiencies in handling, storing and packing onions. Corvallis, Oregon State College, August 1960. 27 p. (Miscellaneous Paper No. 96)

5. Deming, William E. Some theory of sampling. New York, Wiley, 1950. 602 p.

6. Hansen, M. H., W. N. Hurwitz and W. G. Madow. Sample survey, methods and theory. New York, Wiley, 1953. 624 p.

7. Hendricks, Walter A. Introduction to experimental statistics. Washington D. C., The Graduate School of the U. S. Dept. of Agriculture, 1953. 251 p.

8. French, B. C., L. L. Sammet and R. G. Bressler. Economic efficiency in plant operations with specific reference to the marketing of California pears. Hilgardia 24:543-721. 1956.

9. Johnson, Palmer O. and Rao S. Munamarty. Modern sampling methods, theory, experimentation, application. Mineapolis, University of Minnesota, 1959. 87 p.

10. Kramer, Anuhud, W. L. Ogle and B. W. Clark. Sampling procedure for grading tomatoes. Food Technology 7:353-355. 1953.

11. Kramer, Anuhud _et al._ A procedure for sampling and grading raw green asparagus. Food Technology 5:212-214. 1956.

12. Li, Jerome C. R. Introduction to statistical inference. Ann Arbor, Edwards Brothers, 1957. 458 p.

13. Morrow, Robert Lee. Motion economy and work measurement. New York, Ronald, 1957. 458 p.

14. Pursel, John H. An analysis of costs and efficiencies of certain methods used in vining green peas. Master's thesis. Corvallis, Oregon State University, 1962. 57 numb. leaves.

15. Rowe, Gorden A. and Masao Matsumoto. Sampling tomato loads for variables and attributes. Berkeley, U. of California, 1960. Unpublished.

16. Sammet, L. L. and B. C. French. Economic-engineering methods in marketing research. Journal of Farm Economics 35:924-930. 1953.

17. Sitton, Gordon R. and L. L. Sammet. Economic factors in picking, assembling and grading sevillano olives. Berkeley, University of California, 1953. 83 p. (Mimeographed Report No. 155)

18. U. S. Dept. of Agriculture. Agriculture Marketing Service. U. S. standards for snap beans for processing. July, 1959. 7 p.

19. U. S. Dept. of Agriculture. Statistical Reporting Service. Vegetables for processing, acreage, production, value by states, 1954-59. Dec. 1961. 43 p. (Statistical Bulletin No. 299)

APPENDIX

Appendix Table 1. Raw product value estimates in dollars per ton as derived from repeated sampling of ten pound smaples by plant and year.

Load Designation	Tote no.	Plant no. 1 1961 (Load number)			Plant no. 2 1961 (Load number)	
		1-10	10-20	20-21	1-10	10-20
1	A	$129.86	$ 94.35	$ 98.49	$ 99.90	$ 92.12
		127.10	87.72	93.58	92.98	90.08
	B	125.28	91.44	105.79	94.18	94.86
		125.59	83.78	103.43	93.18	91.10
2	A	127.14	93.94		126.20	72.84
		117.50	93.28		128.54	104.14
	B	126.48	87.58		126.03	108.77
		128.83	100.74		131.11	112.62
3	A	127.77	101.75		88.46	115.57
		130.20	100.36		85.75	112.77
	B	130.63	102.51		84.04	116.18
		134.12	80.96		77.50	115.80
4	A	138.16	98.48		89.92	79.40
		133.30	97.11		96.10	77.56
	B	130.88	90.88		96.84	75.66
		125.70	98.10		102.42	80.06
5	A	120.60	96.56		112.55	81.20
		119.86	102.27		104.09	83.15
	B	119.74	102.60		118.76	87.64
		120.38	112.97		118.14	87.83
6	A	125.36	97.86		122.54	88.69
		125.74	102.51		116.94	90.90
	B	131.98	99.06		120.40	88.88
		131.18	98.60		120.59	91.94
7	A	98.92	105.00		114.05	117.40
		97.61	97.38		112.38	114.22
	B	96.05	102.71		115.61	112.74
		96.05	106.52		111.95	117.06
8	A	99.30	129.48		106.09	
		81.58	131.97		104.88	
	B	92.06	134.52		105.78	
		94.83	141.18		104.16	
9	A	96.06	90.32		106.52	
		98.16	90.32		112.10	
	B	98.10	87.97		113.50	
		102.38	88.44		106.28	
10	A	98.42	88.08		106.24	
		98.32	95.18		105.50	
	B	96.64	92.18		113.86	
		90.67	84.09		95.23	

Appendix Table 2. Raw produce value estimates in dollars per ton derived from repeated sampling of thirty pound samples by plant and year.

Load designation	Tote no.	Plant no. 1 1961 (Load number) 1-10	Plant no. 1 1961 (Load number) 10-20	Plant no. 1 1961 (Load number) 20-21	Plant no. 2 1961 (Load number) 1-10	Plant no. 2 1961 (Load number) 10-20	Plant no. 3 1962 (Load number) 1-10	Plant no. 3 1962 (Load number) 10-20
1		$122.84	$106.34	$ 90.82	$ 88.93	$ 89.03	$121.64	$114.49
	A	121.80	102.91	90.74	92.34	93.01	114.90	118.38
		116.10	91.29	99.48	91.17	87.64	124.34	118.18
	B	122.78	93.92	92.78	96.75	89.78	115.62	119.10
2		133.24	99.11		126.18	110.58	120.10	122.95
	A	134.48	97.24		128.56	109.47	121.08	119.33
		135.72	82.45		134.86	113.28	110.99	121.35
	B	130.64	91.96		134.90	109.84	116.47	118.72
3		136.17	99.58		81.54	114.05	121.92	117.09
	A	133.48	79.88		85.40	104.88	122.01	117.15
		131.56	101.38		76.56	112.06	119.35	121.73
	B	129.83	102.66		94.91	117.28	118.61	121.67
4		132.52	81.32		107.70	127.53	129.43	102.36
	A	132.92	90.72		104.16	127.53	124.04	100.71
		130.60	98.62		98.39	127.08	114.57	104.47
	B	129.50	95.51		103.02	127.98	115.05	102.37
5		121.84	88.88		109.68	85.92	125.66	130.10
	A	120.72	90.99		100.07	91.05	126.09	128.60
		119.44	90.84		114.15	80.73	126.93	120.93
	B	121.72	90.84		110.16	88.98	126.67	131.23
6		126.58	97.35		120.43	91.31	131.61	
	A	125.76	96.54		116.34	80.28	132.60	
		127.86	95.18		115.69	87.88	136.37	
	B	130.06	100.39		119.51	88.78	128.26	
7		99.22	98.88		111.80	109.28	103.32	
	A	95.92	98.73		114.22	111.55	103.89	
		92.93	96.28		115.06	109.46	108.05	
	B	106.92	98.60		111.92	107.86	99.49	
8		94.14	129.03		103.43		117.27	
	A	96.02	129.28		102.40		114.58	
		92.96	136.82		101.72		117.39	
	B	90.26	110.58		103.87		119.13	
9		90.87	88.20		104.42		111.98	
	A	95.46	87.88		109.60		108.96	
		99.30	82.93		92.28		114.22	
	B	99.30	89.61		88.06		116.36	
10		91.04	88.44		113.08		119.09	
	A	89.86	86.98		107.98		115.82	
		91.26	80.98		112.04		113.57	
	B	93.24	79.67		103.48		113.61	

Appendix Table 3. Raw product value estimates in dollars per ton derived from repeated sampling of sixty pound samples by plant and year.

Load designation	Tote no.	Plant no. 2 1961 (Load number) 1-10	Plant no. 2 1961 (Load number) 10-17	Plant no. 2 1962 (Load number) 1-10	Plant no. 2 1962 (Load number) 10-15
1	A	$ 92.18	$ 91.17	$111.01	$117.08
	B	92.17	89.90	108.56	116.48
2	A	128.53	113.90	102.13	81.12
	B	131.42	114.55	104.77	82.89
3	A	83.16	105.87	113.46	67.36
	B	81.38	115.98	113.20	67.52
4	A	101.14	122.54	80.57	115.09
	B	97.20	122.54	63.52	126.58
5	A	105.58	88.42	94.92	115.42
	B	108.94	83.87	80.36	119.07
6	A	118.00	88.20	86.11	
	B	121.12	86.01	83.37	
7	A	109.48	109.90	79.98	
	B	110.26	114.44	81.29	
8	A	98.95		122.32	
	B	97.72		118.98	
9	A	105.66		113.87	
	B	104.48		110.92	
10	A	104.82		106.32	
	B	106.06		100.82	

Appendix Table 4. Analysis of variance by plant and year for ten, thirty, and sixty pound sub-samples, 1961, 2.

Ten pound sub-samples

Estimates of	Plant no. 1 1961	Plant no. 2 1962
Among tote mean square	24.92	43.02
Among sub-samples mean square	21.26	27.12
Among tote variation	1.83	7.95
Among sub-samples variation	21.26	27.12

Thirty pound sub-samples

Estimates of	Plant no. 1 1961	Plant no. 1 1962	Plant no. 2 1961
Among tote mean square	36.05	19.13	29.56
Among sub-samples mean square	20.96	8.48	16.31
Among tote variation	7.55	5.32	6.63
Among sub-samples variation	20.96	8.48	16.31

Sixty pound sub-samples

Estimates of	Plant no. 2 1961	Plant no.2 1962
Among tote mean square	5.88	24.13
Among sub-sample mean square		
Among tote variation	5.88	24.13
Among sub-sample variation		

Appendix Table 5. Estimated costs for equipment used in sampling green beans, 1962

Item of equipment	Estimated replacement cost 1/	Useful years of life	Annual fixed cost 2/	Annual variable cost 3/	Seasons total cost 4/
	(Dollars)		(Dollars)	(Dollars)	(Dollars)
Fork lift truck (2,000 lb. capacity)	5050	10	833.25	345.44	1,178.79
Platform scales (125 lb. capacity)	706	10	116.49	12.99	129.48
Mechanical grader	4,150	20	477.25	85.71	562.96
Handtruck	56	15	6.59	.92	7.51
Hopper with conveyor	500	20	57.50	13.87	78.88
Lug boxes (100)	85	5	22.53	1.58	24.11
Tote bins (10)	140	5	37.10	2.57	39.67

1/ Based on current cost new plus estimated installation costs.

2/ Calculated as a percentage of replacement costs as follows: (a) Depreciation based on years useful life; (b) Insurance - 1%; (c) Taxes - 1%; (d) Interest - 3% and (e) Fixed repairs - 1.5%.

3/ Includes; (a) Repairs and maintenance as a percentage of replacement cost per 100 hours use; (b) Gasoline, oil, grease, and (c) electricity.

4/ Hours in operating season are; Fork lift truck, 1,104; all others 368.

Appendix Table 6. Job and job descriptions for labor standards for green bean sampling, 1962.

Job and job description
Select and obtain primary sample units - selection of primary sampling units by the fork lift operator and moving the primary unit selected to the area for drawing the secondary unit.
Dump primary sampling unit - begins when fork lift operator begins to dump the primary unit and includes the dumping of the primary unit into the hopper and the placing of the empty tote box to receive the beans. Also includes the handling and removal of the unsampled beans.
Obtain secondary sampling unit - moving lug boxes into position for sampling, drawing the secondary units from the primary units and taking sampled beans to weighing station.
Grade and size secondary sampling unit - weigh sample, move to grading platform and visually grade. Pass through sizer and weigh and record each grade and size. Remove empty lug boxes.