May 5, 1969

To the Graduate Council:

I am submitting herewith a thesis written by Billy Wayne Nelms entitled "Analysis of Egg Handling Cost in Tennessee." I recommend that it be accepted for nine quarter hours of credit in partial fulfillment of the requirements for the degree of Master of Science, with a major in Agricultural Economics.

James S. Snell
Major Professor

We have read this thesis and recommend its acceptance:

Charles L. Cleland

B. D. Raskopf

Accepted for the Council:

Hilton A. Smith
Vice Chancellor for
Graduate Studies and Research

ANALYSIS OF EGG HANDLING COST IN TENNESSEE

A Thesis

Presented to

the Graduate Council of

The University of Tennessee

In Partial Fulfillment

of the Requirements for the Degree

Master of Science

by

Billy Wayne Nelms

June 1969

ACKNOWLEDGEMENTS

The author wishes to express his gratitude to all persons who have made this study possible. The author is especially grateful to Dr. James G. Snell, committee chairman, for the guidance, patience, and constructive criticism which he provided throughout this study. Special thanks are due also to Dr. Merton B. Badenhop, Dr. Charles L. Cleland, and Mr. B. D. Raskopf, members of the author's graduate committee, without whose help this thesis could not have been written.

To Dr. Thomas J. Whatley, Head of the Department of Agricultural Economics and Rural Sociology, gratitude is extended for making the writer's graduate studies possible.

The writer is grateful to Mrs. Anna Norwood for typing drafts of this thesis.

A very special indebtedness is expressed to my wife, Linda, and son, Tony, for their patience and devotion during the writing of this thesis.

ABSTRACT

This study was designed: (1) to develop economically and statistically sound cost functions for the assembly and distribution operations of the Tennessee egg industry, and (2) to determine if economies of scale were present in the assembly and distribution operations.

The data were obtained from a population of 75 egg handling firms in Tennessee. Basic data were obtained from expense invoices, payroll accounts, operating statements, income tax statement, and other accounting records related to the 1967 assembly and distribution operations of the 75 firms.

Least squares stepwise regression was used to develop cost functions and quasi economies of scale functions. A series of models was developed in both the cost function and the economies of scale analysis. After arriving at various statistical and economic conclusions, the superior equation for each analysis was selected.

The superior cost function selected from the analysis with regard to both the assembly and distribution of eggs was composed of three exogenous variables: (1) reciprocal of the volume of eggs handled, (2) wage rate, and (3) cases of eggs handled per man-hour. These three variables explained 94 percent of the variability in assembly cost per case of eggs and 55 percent of the variability in distribution cost.

The superior quasi economies of scale function for both assembly and distribution employed the reciprocal of the volume variable. By utilizing this functional form, 84 percent of the variability in assembly cost was explained by variation in volume, and 26 percent of the variability in distribution cost was explained by variations in volume. Thus, it was concluded that this functional form developed the best quasi economies of scale curves. Since the equations were developed by least squares analysis and did not establish a method to connect minimum cost points, the true economies of scale curves were not developed. However, it was concluded that this analysis produced evidence that economies of scale exist in the assembly and distribution operations.

The study does not reveal specifically how to reduce cost or how to change the size of operation to obtain lower cost. Instead, it provides a means by which a firm may gain a priori knowledge of factors that greatly effect assembly and distribution cost as well as a priori knowledge of the relative importance in the size of operation.

TABLE OF CONTENTS

CHAPTER I

INTRODUCTION

The egg enterprise is an important source of cash income on Tennessee farms. During the past 8 years the cash farm income from the sale of eggs has averaged $25,762,000 per annum. Eggs accounted for 5 percent of the gross sales of Tennessee farm commodities in 1960 and 4 percent in 1967.[1]

Over the past few years, the gross sales from eggs changed very little in comparison to the changes in the structure of the egg industry. In 1960, Tennessee produced 957 million eggs in comparison to 1,172 million in 1967. This increase in production only appears important when viewed with respect to the number of egg producing farms during these periods. During the period 1960-1967, the number of farms producing eggs declined 75 percent, while the average value of eggs sold per farm increased from $63 to $3,016.[2] However, the changing structure is most evident from the fact that of the 10,000 farms selling eggs in Tennessee in 1967, about 300 farms possessed 62 percent of the laying hens, produced 75 percent of the eggs sold, and had an average gross revenue of $70,000 per farm.[3]

---

[1] United States Department of Agriculture, Tennessee Agriculture Statistics (Washington: Government Printing Office, August 1968), p. 47.

[2] Ibid.

[3] Ibid.

Tennessee ranks 25th in egg production in the United States.[4] Five out of the top ten egg producing states border Tennessee and provide strong competition to Tennessee's egg industry.

Even though Tennessee produces fewer eggs than its population consumes, the availability of eggs in the surrounding states deprives Tennessee of this market potential. This competition provided by the surplus egg producing states that surround Tennessee has caused egg producers in Tennessee to change their methods of egg production drastically in recent years.

Egg producers in Tennessee have met this strong competition from large egg producing states by utilizing new technology and increased knowledge of egg production. These factors have enabled the egg producer to concentrate his production into large, integrated, and financially strong firms. Many of the producers today perform all of the operations from raising the replacement stock to delivering the eggs to retail outlets.[5]

I. OBJECTIVES

The general purpose of this study was to obtain information that would aid the development of the commercial egg industry in Tennessee

---

[4]United States Department of Agriculture, Poultry and Egg Situation (Washington: Government Printing Office, November 1968), p. 21.

[5]H. B. Jones and H. R. Smolly, Vertical Interpretation Methods of Producing and Marketing Eggs, University of Georgia Experiment Station Bulletin No. N.S. 120 (Athens: University of Georgia, September 1966), p. 8.

by providing cost information on the assembly and distribution of eggs. The specific objectives were:

1. to establish estimating equations for the cost of assembling and distributing eggs, and
2. to determine if economies of scale exist in the Tennessee egg industry.

II. JUSTIFICATION

The increased production and the increased market potential for Tennessee eggs have presented many new problems. The total industry has become very competitive. If the Tennessee egg industry is to survive, the allocation of resources must be in an efficient manner.

The cost of assembling and distributing eggs is an important part of the total cost of egg production, and there is a need for some explanation of the factors that affect assembly and distribution cost. The producers desire to know what can be changed in their operation that will make them competitive with other operators. An insight into the cost function relative to the transportation of eggs will provide means by which the egg producers may analyze their competitive position.

III. PROCEDURES

The analysis of the data in this study was performed by stepwise regression. In the first part of the study, various regression models were developed in order to synthesize a statistically and economically sound cost function. The second part of the study utilized regression analysis as a means by which to identify the existence of economies of scale. In this part, regression models that produced nonlinear curves were used and compared to linear functions established for given size classes. From this type regression analysis, it was possible to compare costs for various size firm.

IV. ORGANIZATION OF THESIS

Six chapters were utilized in order to systematically present the contents of this study. Chapter I contains a brief description of the Tennessee egg industry, the objectives of the study, justification of the study, a brief synopsis of the procedures, and this section which presents the organization of the thesis. Chapter II deals with the procedures employed in the study. In this chapter, the source and type of data is viewed first, and then the statistical procedures are developed in length. To conclude this chapter, the general limitations of the study are presented. Chapter III presents the synthesis of the cost equations for the assembly and distribution of eggs. Chapter IV contains the synthesized proofs for the existence of economies of scale in the assembly and distribution segment of the egg industry. Chapter V presents generated cost functions for the assembly and distribution operations. Chapter VI summarizes the results of this study.

CHAPTER II

RESEARCH PROCEDURE

I. SOURCE OF DATA

This study was based on a survey of 75 egg handling firms in the State of Tennessee. An egg handling firm was defined in this study as any firm that assembles and/or distributes eggs. Assembly costs were defined as the cost incurred in the transportation of eggs from the point of production to a point from which the eggs were distributed to various outlets for consumption. Distribution costs were defined as those costs incurred in the transportation of eggs from a point of preparation for market to the market outlet.

The 75 firms were selected from the complete population of 137 firms in the state that produced at least 2,600 cases of eggs per year. The limitation of the survey population to those firms that produced over 2,600 cases of eggs per year was established because they accounted for 99 percent of all the eggs produced in Tennessee. This limitation also allowed for the collection and analysis of similar type cost factors for each firm. The 75 firms were stratified into two groups. The first stratum was composed of those firms that handled 20,800 or more cases of eggs per year. The second stratum consisted of those firms that handled from 20,799 to a minimum of 2,600 cases of eggs per year.

The population of 75 firms was selected from the two strata. The entire first stratum of 28 firms was entered into the population. The remainder of the population was selected from the second stratum by simple random sampling. Each of the firms in the stratum was assigned a number, and then the selection was made by using a table of random numbers.

Fieldmen visited the firms during January and February, 1968. Detailed cost information pertaining to assembly and distribution operations was obtained. Basic data were obtained from expense invoices, payroll accounts, operating statements, income tax statements, and other accounting records relating to the 1967 assembly and distribution operations of the 75 firms.

After the assimilation of the data, the data were divided into: (1) firms that assembled eggs, and (2) firms that distributed eggs. The data were reviewed and those factors affecting assembly and distribution cost were determined. Seven variables were selected from the data.

1. total volume of eggs handled
2. wage rate per hour
3. total miles traveled per year
4. number of sources or outlets
5. number of trucks
6. total labor in hours
7. size of truck in cases

Four variables not directly available in the data were computed;

1. percent truck capacity utilized
2. cases per mile traveled
3. cases per man-hour
4. cases per source

From inspection of the raw data, the first variable computed (percent truck capacity utilized) was deemed of economic importance due to the wide variation in capacity utilized and its apparent correlation with the cost per case of eggs for each firm. The high cost firms appeared to utilize less of their truck capacity than the lower cost firms. The last three computed variables were deemed of importance to the study due to their previous significance in a study by White in 1963.[1]

II. GENERAL COST FUNCTIONS

The first objective was to establish estimating equations for the cost of assembling and distributing eggs. A stepwise regression analysis was chosen as the statistical tool to determine the functional relationship between the cost factors obtained from the survey and the cost per case of eggs assembled or distributed. The stepwise regression provided a means by which multiple linear regression equations could be computed in a stepwise manner. The term stepwise defines the manner by which the variables were added to the regression equation. The variable

---

[1]Morris White, Relationship of Marketing Methods to Cost of Assembly, Grading, and Packaging Table Eggs, Auburn University Experiment Station Bulletin No. 348 (Auburn: Auburn University, October 1963), p. 14.

that enters the regression equation first was the variable that reduces the error sum of squares by the greatest amount when regressed upon the dependent variable in a simple regression. The remaining variables enter the regression equation in a stepwise manner according to their partial correlation with the dependent variable, after considering the variables which had already been added.[2] The variables that were entered were subject to a constraint termed the F value.[3] The F value constraint was determined ex ante and the entrance of a variable into the regression equation was permitted only if the calculated F value of the variable was greater than the established F value. In this study, the F value was set very low (.001) to allow all variables to enter. This was used so that economic consideration could be given to all variables even though statistically some were of little significance.

To perform a regression analysis that would give estimating equation for both cost of assembly per case of eggs and cost of distribution per case of eggs, it was necessary to divide the analysis into two processes. The group of variables related to assembly cost was analyzed separately from the data relative to distribution cost. The classification and descriptions of the variables for both assembly and distribution were the same.

---

[2]The University of Tennessee Computing Center, "BMDO2R Stepwise Regression" (Knoxville: The University of Tennessee, September 1, 1965), p. 1.

[3]Taro Yamane, Statistics: An Introductory Analysis (New York: Harper and Row, 1967), p. 676.

The regression analysis was set up in a manner which allows for the determination of a single estimating equation for each of the two operations. To insure that the estimating equations selected to represent assembly cost and distribution cost per case of eggs were the best available from the data, considering both the economical and the statistical side, a procedure was developed which allowed various equation forms to be compared.

The first step was to regress all of the variables associated with each of the costs upon each of the dependent variables. This enabled the identification of any multicollinearity that existed in these variables. From this analysis, the variables were divided into two groups by the use of statistical and economic interpretation with regard to the value of each variable relative to the regression analysis. The independent variables were divided into the following groups and designated as models.[4]

Model A-1.1 and Model D-1.1

X_1 = total volume of eggs assembled

X_2 = wage rate per hour

X_3 = total miles traveled per year

X_4 = number of sources

X_5 = number of trucks

X_6 = total labor in hours

---

[4]Model A-i.i refers to assembly cost functions, and Model D-i.i refers to distribution cost function.

X_7 = size of truck in cases

X_8 = percent truck capacity utilized

Model A-2.1 and Model D-2.1

X_1 = total volume of eggs distributed

X_2 = wage rate per hour

X_7 = size of truck in cases

X_8 = percent truck capacity utilized

X_9 = cases per mile traveled

X_{10} = cases per man-hour

X_{11} = cases per source

After the selection of the two basic models, it was necessary to develop additional statistical constraints that would insure that the estimating equation selected was the best available from these data. Two constraints were employed which developed two additional models for each set of variables. Model A-1.2 and Model A-2.2 were developed from the first models by eliminating those variables that added less than .10 to the regression sum of squares upon its entry into the regression analysis. Those models relative to distribution were designated as Model D-1.2 and Model D-2.2.

A third equation form or model was established by using an entrance constraint. Those variables in the second model that were not statistically significant at the 5 percent level were excluded from the regression equation. These models were designated as Model A-1.3, Model A-2.3, Model D-1.3 and Model D-2.3.

The stepwise regression analysis was performed by a computer. After the analysis, the dependent variables were estimated by six different estimating equations. To select the best estimating equation, it was necessary to test and compare each of the equations from both a statistical and an economic standpoint.

A statistical test explained by Fox was used to test the statistical significance of each set of models.[5] The test provides a means to test the significance of the coefficients of determination with relation to other coefficients of determination. The formula was:

$$F = \frac{\dfrac{RSS_1 - RSS_2}{n \pm j}}{m.sq._1}$$

where

RSS = residual sum of squares of large R^2,

RSS = residual sum of squares of small R^2,

$n \pm j$ = number of variables added or deleted, and

$m.sq._1$ = mean square error of the large R^2.

The testing process was performed in a systematic manner. All of the model in the first group (Model A-1.1, A-1.2, A-1.3) were tested against each other to determine if one was statistically superior to the others in its ability to estimate assembly cost per case of eggs. The same procedure was used on the distribution models. If the models

---

[5]Karl A. Fox, Intermediate Economic Statistics (New York: John Wiley and Sons, Inc., 1968), pp. 233-237.

were not statistically different, the selection of the best model was performed by employing economic criteria.

The second step was to test the second group of models (Model A-2.1, A-2.2, A-2.3). The same procedure was used as in the first group. The best estimating equation available in this group of models was selected. The third step consisted of the comparison of the two equations selected from the two sets of models. The equation that proved statistically and/or economically superior to the others was chosen as the estimating equation that best estimated assembly or distribution cost per case of eggs.

A third set of models was developed to analyze the cost of assembly and distribution operations, by incorporating the results from Chapter III and IV. The cost equations that were determined to be superior in Chapter III were incorporated with the superior quasi economies of scale equations in Chapter IV. The generated cost equations were compared with the cost equations of Chapter III by employing the F-test as in previous parts of the study. The models were designated as Model A-G and Model D-G.

III. ECONOMIES OF SCALE

Economies of scale are defined by Bressler to be, "The levels of cost that may be expected from the operations of plants of various size, when operations are organized as efficiently as possible under given conditions."[6] Economies of scale may also be considered as those points that represent the lowest average cost achievable by variation in the scale of operation. From the standpoint of this study, economies of scale were deemed of importance due to the very large range in both size and cost. No attempt was made to produce the causes of economies that might exist but rather to determine only their existence or nonexistence.

Economies of scale were analyzed from a cost standpoint based on economic theory. From economic analysis, the costs incurred by a firm were attributable to the production process. The production process was a technical process whereby the outputs were a function of inputs. The costs to the production process were directly related to the cost of the inputs. From this analysis of production, the basis for the study of economies of scale was derived.[7]

The factors available in this study which lent themselves to an analysis of economies of scale were: (1) total cost per case of eggs handled, and (2) total cases of eggs handled. By comparing the relationship between these two variables, it was possible to determine the existence of lower handling cost associated with large volume firms contrasted to higher handling cost for the smaller firms.

Engineering cost data are often used to determine the optimum level of economies to scale. Engineering data allows the determination

---

[6] R. G. Bressler, Jr., "Research Determination of Economies of Scale," Journal of Farm Economics, 27:526-539, August 1945.

[7] H. B. Jones, Economies of Scale in Commercial Egg Packing Plants, University of Georgia Experiment Station Bulletin No. N.S. 120 (Athens: University of Georgia, September 1966), p. 9.

of those points of lowest cost per given output, given the cost of each input, which is assumed to be the same for all firms. In this study, the objective was to determine the existence of economies in the existing industry. The cost information available for analysis was obtained from accounting data which did not allow for prediction or analysis of the optimum economies of scale for the industry.[8]

The economies of scale were analyzed by a regression analysis. By using the regression approach, it was possible to get an estimate of the economies of scale.[9] This was a quasi economies of scale analysis since the estimated regression equation was fitted through the scatter of cost observations rather than through the locus of least cost observations.

It was necessary to establish a regression equation form that would fit a line to the data in such a manner so as to explain a large amount of the variability in the dependent variable by the variation in the independent variables. Four different regression equations were used in the attempt to arrive at a high coefficient of determination. A simple linear model was chosen along with three other models that entered the regression analysis in linear fashion, but if plotted they were of nonlinear form.

Model 1. $Y = a + b_1X_1$

---

[8] Milton Friedman, Price Theory (Chicago: Aldine Publishing Company, 1967), p. 169.

[9] R. C. Bressler, Jr., op. cit., p. 528.

Model 2. $Y = a + b_1X_1 + b_1X_1^2$

Model 3. $Y = a + b_1 \frac{1}{X_1}$

Model 4. $\text{Log } Y = a + b_1 \log X_1$

Y = assembly or distribution cost per case of eggs.

X_1 = total volume of eggs, in cases, assembled or distributed per firms per year.

After the performance of the regression analysis, the models were tested against each other in order to determine if the estimating equations were statistically different. The previously discussed F-test was utilized.[10] The regression equation that best fitted the data was selected by statistical analysis. If one equation proved to be statistically different from the other three models, it was chosen. If two or more equations tested out to be statistically adequate in their abilities to explain the variation in cost per case, then the statistically best equation was chosen with regard to the highest coefficient of determination.

The selection of the final equation produced the evidence of either economies of scale or no economies of scale. If the estimated cost equation produced decreasing cost per case of eggs as the volume of eggs handled increased, it was evident that economies existed in the industry. Increasing cost per case of eggs would produce diseconomies of scale.

---

[10] Refer to page 11.

Another statistical analysis was also employed as a means of testing the existence of economies of scale in the egg assembly and distribution operations. The development of the nonlinear cost curves in the previous part of this study gave an insight into the cost associated with various size firms. This part of the study was developed in order to estimate these variations in cost by linear curves for selected size groups of assemblers and distributors. This was done to determine if the linear cost curves could produce very close approximation of the continuous nonlinear cost curves.

In the construction of linear cost curves for selected size firms, zero-one variables were used. Due to the small number of observations in the study, it was appropriate to use the zero-one variable approach. Zero-one variables enable the development of linear equations for subsets of the sample population while keeping the degree of freedom high. They also provided a means by which to obtain the equations in a one step regression analysis.[11]

Zero-one variables were exactly what the name implied. By dividing the total sample population into distinct subclasses, it was possible to assign a zero or one value to each firm in a given subclass. A new variable was formed by classifying the firms in a subclass as zero or one. When the firm was assigned a value of one this meant that this firm was part of a given subclass of firms, and this class of firms was

---

[11]Shaul Ben-David and W. G. Tomek, "Allowing for Slope and Intercept Changes in Regression Analysis," unpublished Ph.D. thesis, Cornell University, 1968, pp. 1-14.

identifiable by the variable formed. When the firsm were designated as zero within the variable, this meant that the firms were not part of this subclass and were excluded from the regression analysis when the variable was entered. This procedure allowed the identification and control of predetermined subclasses of the sample population when synthesizing cost equations by a regression analysis.[12]

In order to develop zero-one variables for this part of the study, the 29 firms that assembled eggs were divided into the following classes:

Large class (338,000-45,000 cases/year)

Medium class (44,999-14,000 cases/year)

Small class (13,999-2,600 cases/year)

The 69 firms that distributed eggs were also divided into three classes.

Large class (338,000-26,000 cases/year)

Medium class (25,999-8,000 cases/year)

Small class (7,999-2,600 cases/year)

Zero-one variables were appropriate since they provided a method which allowed for the analysis of cost relationships between classes from the standpoint of intercept, slope, and both intercept and slope changes. In this part of the study of economies of scale, all three approaches were employed.

The existence of economies with relation to intercept changes was analyzed first. Average fixed cost was defined as that cost incurred

---

[12] J. Johnston, Econometric Methods (New York: McGraw-Hill Book Company, Inc., 1960), p. 221.

when production was zero. The estimated alpha coefficient from the regression analysis could be considered as an estimate of fixed cost. However, if the estimated alpha coefficient was to be a "true" estimate of fixed cost, there had to be an observation of zero production. Since production never equaled zero in the study, the alpha coefficient could not be considered a true estimate of the fixed cost.

In order to analyze the difference in intercept values for each of the classes, it was necessary to develop an estimated alpha coefficient for each class while assuming that the slope coefficient for all the classes was constant.[13]

To perform the regression analysis, it was necessary to develop a regression equation that would produce an estimated alpha coefficient for each class while holding the slope coefficient constant for all classes. The first step in developing the regression equation was to divide the basic linear regression equation for the entire sample population into three equations. The basic regression equation was:

$$Y = a_1 + b_1X_1$$

where

Y = estimated cost per case of eggs handled per firm,

a_1 = intercept coefficient,

b_1 = slope coefficient, and

---

[13] Robert Ferber and P. V. Verdoorn, Research Methods in Economics and Business (New York: The MacMillan Company, 1962), pp. 369-372.

X_1 = total cases of eggs handled per year per firm.

From this basic equation the following equation forms were developed for the three classes of egg handlers.

Small class $Y = a_{12} + b_1X_1$

Medium class $Y = a_{13} + b_1X_1$

Large class $Y = a_{14} + b_1X_1$

The alpha terms represent the estimated "fixed cost" for the respective class. The "X" variables are the volume of eggs handled by the firms. In order to determine the intercept values for each class and hold the slope coefficients constant, the following equation was developed from the three above.

$$Y = a_1 + b_1X_1 + a_{12}X_{12} + a_{13}X_{13} + a_{14}X_{14}.$$

The alpha coefficients denote the estimated fixed cost for the respective classes as in the previous three equations. The X_{12}, X_{13} and X_{14} variables represent the zero-one variables. Each of these variables was composed of firms that were coded as one for the appropriate class and zero when the firm did not belong in the particular class. For example, X_{12} was composed of firms in the small class which were coded as one to allow the regression of this subset of firms upon cost per case of eggs per firm. All other firms were coded as zero with respect to X_{12}, thus they were not entered into the regression analysis. The beta coefficient represents the constant slope. The variable X_1 represents the total cases of eggs handled per firm.[14]

The equation $Y = a_1 + b_1X_1 + a_{12}X_{12} + a_{13}X_{13} + a_{14}X_{14}$ was indeterminate by regression analysis due to the singular matrix that prohibited inversion of the necessary matrix. Perfect linear multiple correlation existed among the X variables.[15]

In order to correct the singular matrix situation, one of three restrictions had to be employed. By assuming that the summation of the alpha coefficients ($a_{12} + a_{13} + a_{14}$) was equal to zero, the singular matrix could be avoided. This procedure was not used due to the harshness of the assumption from an economic standpoint. From an economic standpoint, it seemed irrational that the summation of the fixed cost for each class would be zero. The second restriction was to assume that the basic alpha coefficient, determined when regressing the entire sample population upon cost per case of eggs, was equal to zero. From an economic standpoint, this also seemed irrational. By analyzing cross-sectioned data, there was no valid reason to assume that average fixed cost was zero. The third restriction was selected as the assumption to be made in order to eliminate the perfect linear multiple correlation. Here it was assumed that the alpha coefficient for the small class was equal to zero. By doing this, the average fixed cost for the small

---

[14] W. G. Tomek, "Using Zero-One Variables with Time Series Data in Regression Equations," Journal of Farm Economics, 45:814-822, November 1953.

[15] Shaul Ben David and W. G. Tomek, op. cit., p. 9.

class took on the value of the basic alpha value. This allowed moments about the mean to be used in the regression equation.[16]

With the employment of this restriction upon the unsolvable equation forms, the applicable regression equation was established.

$$Y = a_1 + b_1X_1 + b_{13}X_{13} + b_{14}X_{14}$$

where

Y = estimated handling cost per case of eggs per firm,

a_1 = intercept value of the small class,

b_1 = slope coefficient for all classes,

X_1 = total volume of eggs handled per firm per year,

$b_{13} + a_1$ = intercept value of the medium class,

X_{13} = 1 if the firm is medium class, 0 elsewhere,

$b_{14} + a_1$ = intercept value of the medium class, and

X_{14} = 1 if the firm is large class, 0 elsewhere.

A stepwise regression analysis was performed by an electronic computer. The intercept values for each class were computed, and the estimating equation for each class was established. In order to determine if the alpha coefficient associated with the different size firms were statistically different from each other, a t-test was employed.[17]

---

[16] W. G. Tomek, op. cit., p. 817.

[17] Paul G. Hoel, Elementary Statistics (New York: John Wiley and Sons, Inc., 1967), p. 172.

A null hypothesis was established saying that the intercept values equaled zero.[18] If the test failed to reject the hypothesis, then it was accepted that economies of scale did not exist in the industry. If the hypothesis was rejected when some difference did exist in the cost for different size firms, it was concluded that, from an economic standpoint, economies of scale were present.

The second step in this part of the study was to analyze average variable costs by employing dummy variables. The firms were divided into mutually exclusive groups, and the assumption was made that the intercept values for each group or class was the same. After dividing the basic equation into three equations (one for each class), the equations were recombined by allowing one intercept coefficient to represent all classes, and by the use of dummy variables which provide a means to analyze all three size groups in one regression equation.[19] The equation form was:

$$Y = a_1 + b_1X_1 + b_{12}(X_{12}X_1) + b_{13}(X_{13}X_1) + b_{14}(X_{14}X_1)$$

where

Y = estimated handling cost per case of eggs per firm,

a_1 = intercept value which is constant for all classes,

b_1 = slope coefficient for sample population,

---

[18] $$t_{\frac{a}{2}}, n\text{-}k = \frac{B_i = 0}{S_{bi}}.$$

[19] W. C. Tomek, op. cit., pp. 814-818.

X_1 = total volume of eggs handled per firm per year,

b_{12} = slope coefficient for small class,

X_{12} = 1 when small class, 0 elsewhere,

b_{13} = slope coefficient for medium class,

X_{13} = 1 when medium class, 0 elsewhere,

b_{14} = slope coefficient for large class, and

X_{14} = 1 when large class, 0 elsewhere.

This equation was unsolvable by regression analysis due to perfect linear multiple correlation. To correct this situation, the same restriction was employed as in the intercept changes. The small class coefficient was assumed to be zero. The new equation form was:

$$Y = a_1 + b_1X_1 + b_{13}(X_{13}X_1) + b_{14}(X_{14}X_1)$$

where

Y = estimated handling cost per case of eggs per firm,

a_1 = intercept coefficient for all classes,

b_1 = slope coefficient for small class,

X_1 = total volume of eggs handled per firm per year,

$b_{13} + b_1$ = slope coefficient for medium class,

X_{13} = 1 when medium class, 0 elsewhere,

$b_{14} + b_1$ = slope coefficient for large class, and

X_{14} = 1 when large class, 0 elsewhere.

A stepwise regression analysis was performed and the resulting slope coefficients were tested with regard to their statistical significance by the t-test. The null hypothesis was that the slope coefficients for all classes were equal.[20] If the test accepted the hypothesis, it was proven statistically that economies of scale did not exist in the egg assembly and distribution industry with regard to average variable cost. If the hypothesis was rejected, it was proven statistically that economies of scale did exist in the industry. The rejection of the hypothesis only excluded the existence of economies from a statistical standpoint and not from an economic standpoint.

The third step performed in this part of the study was the use of zero-one variables to establish separate cost functions for each class of egg assemblers and distributors. By developing separate cost functions, it was possible to determine if economies of scale existed among the three classes with regards to total cost. The same procedures were used in the development of the regression equation as in the first two parts. Zero-one variables were employed in such a way as to allow both intercept and slope changes.[21] To eliminate perfect linear multiple correlation, the restriction that the small class intercept and slope coefficient equaled zero was also used. The equation form developed was as follows:

---

[20] Since the initial assumption was that b_1 (small class) equaled zero, the null hypothesis is the same as testing the hypothesis that bi = 0.

[21] Shaul Ben-David and W. G. Tomek, *op. cit.*, p. 6.

$$Y = a_1 + a_{13}X_{13} + a_{14}X_{14} + b_1X_1 + b_{13}(X_{13}X_1) + b_{14}(X_{14}X_1)$$

where

Y = estimated handling cost per case of eggs per firm,

a_1 = intercept coefficient for small class,

$a_{13} + a_1$ = intercept coefficient for medium class,

$a_{14} + a_1$ = intercept coefficient for large class,

b_1 = slope coefficient for small class,

$b_{13} + b_1$ = slope coefficient for medium class,

$b_{14} + b_1$ = slope coefficient for large class,

X_1 = total volume of eggs handled per firm per year,

X_{13} = 1 when medium class, 0 elsewhere, and

X_{14} = 1 when large class, 0 elsewhere.

The stepwise regression analysis was performed by electronic computer. The estimating equation for each class was developed. The t-test was used to determine the statistical significance of each equation. The null hypothesis was that both the slope coefficient and the intercept coefficient equaled zero. If the null hypothesis was accepted by the test, it was proven statistically that economies of scale were not present in the assembly and distribution segment of the egg industry. If it was rejected, the equations were statistically important in their explanation of economies of scale.

The conclusion of this study of economies of scale in the assembly and distribution segment of the egg industry consisted of the comparison of nonlinear curves versus short linear curves in their explanation of economies of scale.

CHAPTER III

COST FUNCTIONS

In this chapter, those factors affecting the cost of assembling and distributing eggs were analyzed by a stepwise regression in order to establish equations that would adequately estimate the cost per case of eggs handled in each of these operations. Those cost associated with the assembling of eggs were analyzed first, and then the costs associated with the distribution operation were analyzed. A total of six estimating equations were developed for each of the operations.

I. ASSEMBLY

The analysis of the assembly cost was associated only with those variables that affected the cost incurred in the transportation of eggs from the point of production to a collecting point from which the eggs were distributed to various outlets for consumption. The variables and criteria used in the analysis were presented in the previous chapter.

Model A-1.1. Those variables that composed Model A-1.1 entered the regression analysis in a stepwise fashion.[1] The resulting estimating equation was:[2]

---

[1]The variables are shown in the order which they entered the regression equation.

[2]Standard errors of the regression coefficients are shown in the parentheses below the coefficients.

$$Y = .72 - .76X_8 \quad -.004X_7 \quad +.13X_2$$
$$(.20) \quad (.005) \quad (.07)$$
$$-.008X_5 \quad -.00005X_6 \quad -.000003X_1 \quad +.000001X_3$$
$$(.03) \quad (.00004) \quad (.000001) \quad (.000002)$$
$$-.0007X_4$$
$$(.003)$$

(8)

$R^2 = .69$ Se = 14 cents[3]

<u>Model A-1.2</u>. This model was derived by allowing only those variables in Model A-1.1 that reduced the residual sum of squares by .10, upon their entry into the regression equation, to enter the model. The following estimating equation was produced:

$$Y = .71 - .77X_8 \quad -.0004X_7 \quad +.13X_2$$
$$(18) \quad (.004) \quad (.05)$$
$$-.000003X_1 \quad +.00006X_6$$
$$(.000001) \quad (.00002)$$

$R^2 = .68$ Se = 13 cents

<u>Model A-1.3</u>. This model was developed by refining Model A-1.2 and omitting those variables in Model A-1.2 which were not statistically significant at the 5 percent level. The result was:

---

[3]Se is the standard error of estimate in cents.

$$
\begin{array}{llll}
Y = .70 & -.81X_8 & +.13X_2 \\
 & (.17) & (.05) \\
 & -.000003X_1 & +.00006X_6 \\
 & (.000001) & (.00002) \\
R^2 = .67 & & \text{Se} = 13 \text{ cents}
\end{array}
$$

Model A-2.1. This model was composed of seven variables that contributed to the variability in assembly cost per case of eggs. The regression analysis produced the following regression equation:

$$
\begin{array}{lllll}
Y = .82 & -.57X_8 & -.006X_{10} & +.11X_2 & \\
 & (.16) & (.003) & (.05) & \\
 & -.0005X_7 & -.04X_9 & -.0000005X_{11} & -.00000005X_1 \\
 & (.0004) & (.03) & (.000002) & (.0000003) \\
R^2 = .79 & & & \text{Se} = 11 \text{ cents} &
\end{array}
$$

Model A-2.2. This model was composed of five variables from Model A-2.1 that reduced the residual sum of squares by .10:

$$
\begin{array}{llll}
Y = .81 & -.57X_8 & -.006X_{10} & -.11X_2 \\
 & (.15) & (.003) & (.04) \\
 & -.0006X_7 & -.04X_9 & \\
 & (.003) & (.02) & \\
R^2 = .79 & & & \text{Se} = 11 \text{ cents}
\end{array}
$$

<u>Model A-2.3</u>. Three variables that were statistically significant at the 5 percent level in Model A-2.2 were defined as Model A-2.3. The resulting estimating equation was as follows:

$$Y = .78 - .62X_8 \quad - .010X_{10} \quad + .10X_2$$
$$\quad (.15) \qquad (.002) \qquad (.04)$$

$R^2 = .76$ Se = 11 cents

After the development of the estimating equations, it was necessary to determine if any one of these equations could be proven to be both statistically and economically superior to all the others in its ability to estimate assembly cost per case of eggs. An F-test, as explained in Chapter II was employed as a means by which to test each equation against the others.

A systematic procedure was followed. The Model A-1 equations were tested against each other first. It was hypothesized that the equations were not staitstically different at the 5 percent level of significance.

Model A-1.1 vs Model A-1.2

$$F_{(3,21)} = \frac{\frac{.8959 - .8837}{3}}{.0198} = .202$$

(Critical F at 5% level = 3.07)

Model A-1.1 vs Model A-1.3

$$F_{(5,21)} = \frac{\frac{.8959 - .8652}{5}}{.0198} = .303$$

(Critical F at 5% level = 2.68)

Model A-1.2 vs Model A-1.3

$$F_{(2,24)} = \frac{\frac{.8837 - .8652}{2}}{.0178} = .505$$

(Critical F at 5% level = 3.40)

The null hypothesis that the estimating equations were not significantly different from a statistical standpoint was accepted in all three cases. With these results, it was both statistically and economically correct to select Model A-1.3 as the most appropriate estimating equation of the three. Statistically, Model A-1.3 was proven to be equal to the other equations even though it contained fewer variables. Economically, Model A-1.3 was just as adequate as the other cost functions with regard to the relationship of the exogenous and endogenous variables.

The Model A-2 equations were tested next. The same procedure was used as in the testing of the Model A-1 equations, and the null hypothesis was the same.

Model A-2.1 vs Model A-2.2

$$F_{(2,22)} = \frac{\frac{1.016 - 1.015}{2}}{.0136} = .037$$

(Critical F at 5% level = 3.44)

Model A-2.1 vs Model A-2.3

$$F_{(4,22)} = \frac{\frac{1.016 - .9636}{4}}{.0132} = 1.07$$

(Critical F at 5% level = 2.82)

Model A-2.2

vs $$F_{(2,24)} = \frac{\frac{1.015 - .9636}{2}}{.0121} = 2.17$$

Model A-2.3

(Critical F at 5% level 3.40)

The null hypothesis that the estimating equations were not significantly different at the 5 percent level was accepted in this series of tests. Model A-2.3 proved to be as adequate as the other two equations in its ability to estimate the assembly cost per case of eggs.

After the selection of the economically and statistically superior estimating equation from each set of models (Model A-1 and A-2), the F-test was utilized to determine if a statistical significant difference existed between the two equations.

Model A-1.3

vs $$F_{(1,26)} = \frac{\frac{.9636 - .8652}{1}}{.0132} = 7.46$$

Model A-2.3

(Critical F at 5% level = 4.22)

The test rejected the null hypothesis that the two equations were equal in their abilities to estimate assembly cost per case of eggs. From these tests, the one equation that most adequately estimated the cost per case of eggs assembled from the date available in this study was:

$$Y = .78 - .62X_8 - .010X_{10} + .10X_2$$
$$(.15) \quad (.002) \quad (.004)$$

where

Y = assembly cost per case of eggs in cents,

X_8 = percent truck capacity utilized in assembly,

X_{10} = cases of eggs assembled per man-hour, and

X_2 = wage rate per hour.

This equation states that assembly cost per case of eggs is a function of the percent of truck capacity utilized, cases of eggs assembled per man-hour, and wage rate per hour. The variation of these three variables explains 76 percent of the variability in assembly cost. If the percent truck capacity utilized is increased by one unit, it is estimated that on the average the cost per case of eggs assembled will decrease by .62 units. A one unit increase in cases of eggs assembled per man-hour is estimated to reduce assembly cost, on the average, by .01 units. The equation also estimates that assembly cost will, on the average, increase 10 cents, if wage rates increase by one dollar. Thus this equation tells not only of what variables assembly cost is a function, but it also tells the relationship between these variables and assembly cost.

II. DISTRIBUTION

The stepwise regression analysis was performed upon the cost factors that were associated with the distribution operation in the egg industry. The variables were divided into two sets as explained in Chapter II. The criteria used in the regression procedures were identical to those used in the development of the assembly equations.

Model D-1.1. This model consisted of eight variables that were assumed to affect the cost of egg distribution. The variables entered the regression equation with the statistically most important entering first. The estimating equation developed was:

$$
\begin{array}{llllll}
Y = .36 & - .001X_7 & + .18X_2 & - .16X_8 & + .000005X_6 \\
 & (.004) & (.05) & (.21) & (.00001) \\
 & - .000005X_1 & + .001X_3 & + .0000010X_4 & - .01X_5 \\
 & (.000001) & (.005) & (.0000009) & (.01)
\end{array}
$$

$R^2 = .54$ Se = 21 cents

Model D-1.2. Six variables were entered into regression Model D-1.2. The six variables were chosen because they increased the regression sum of squares by .10 upon their entry into Model D-1.1. The variables entered the stepwise regression in linear fashion to form the following estimating equation for distribution cost per case of eggs.

$$
\begin{array}{lllll}
Y = .36 & - .001X_7 & + .19X_2 & - .19X_8 & + .00005X_6 \\
 & (.0004) & (.05) & (.21) & (.00001) \\
 & - .000005X_1 & + .001X_4 & & \\
 & (.0000009) & (.0005) & &
\end{array}
$$

$R^2 = .53$ Se = 21 cents

Model D-1.3. Model D-1.3 was established by using an entrance constraint. Those variables in Model D-1.2 that were not statistically significant at the 5 percent level were excluded from the regression equation. The synthesized estimating equation was:

$$Y = .18 - .001X_7 + .19X_2 + .00005X_6$$
$$(.0004) \quad (.05) \quad (.00001)$$
$$- .000005X_1 + .001X_4$$
$$(.0000009) \quad (.0005)$$

$R^2 = .53$ Se = 21 cents

Model D-2.1. Model D-2.1 was composed of seven variables. Three of the variables (X_1, X_2, and X_7) were taken directly from the questionnaire and four (X_8, X_9, X_{10} and X_{11}) were computed from the data as explained in Chapter II. The estimating equation resulting from the stepwise regression of distribution cost per case of eggs upon these variables was;

$$Y = .57 - .02X_{10} + .19X_2 - .0000008X_1$$
$$(.004) \quad (.05) \quad (.0000005)$$
$$- .05X_9 - .22X_8 + .000006X_{11} - .0001X_7$$
$$(.03) \quad (.20) \quad (.000007) \quad (.0004)$$

$R^2 = .56$ Se = 21 cents

Model D-2.2. This model was developed from those variables in D-2.1 that reduced the residual sum of squares by .10 when they entered the regression analysis.

$$Y = .36 - .02X_{10} + .20X_2 - 0000008X_1 - 104X_9$$
$$(.004) \quad (.04) \quad (.0000004) \quad (.03)$$

$R^2 = .55$ Se = 21 cents

Model D-2.3. Three variables that were statistically significant at the 5 percent level in Model D-2.2 were defined as Model D-2.3. The equation developed was:

$$Y = .35 - \underset{(.003)}{.02X_{10}} + \underset{(.04)}{.20X_{2}} - \underset{(.0000003)}{.0000007X_{1}}$$

$R^2 = .53$ Se = 21 cents

After the development of the six estimating equations, they were tested in the same manner as the assembly equations. By using the F-test, it was possible to determine if any one of the equations was superior to the others in its ability to estimate distribution cost per case of eggs. The same systematic procedure was followed as in the assembly section.

Model D-1.1 vs Model D-1.2

$$F_{(2,61)} = \frac{\frac{3.209 - 3.129}{2}}{.0456} = .877$$

(Critical F at 5% level = 3.15)

Model D-1.1 vs Model D-1.3

$$F_{(5,61)} = \frac{\frac{3.209 - 3.092}{5}}{.0456} = .023$$

(Critical F at 5% level = 2.37)

Model D-1.2 vs Model D-1.3

$$F_{(3,63)} = \frac{\frac{3.129 - 3.092}{3}}{.0421} = .028$$

(Critical F at 5% level = 2.75)

The null hypothesis (Ho: D-1.1 = D-1.2 = D-1.3) was accepted in all of the test. This says that the cost per case of eggs distributed may be estimated by any of the three equations with statistically equal results. It also says that Model D-1.3, which contains only five variables, is as statistically adequate as any of the others in its estimation of distribution cost. Thus, Model D-1.3 was selected as the best estimation equation developed from the variables in this set of models.

The Model D-2 equations were tested next.

Model D-2.1 vs Model D-2.2

$$F_{(3,62)} = \frac{\frac{3.340 - 3.255}{3}}{.0427} = .665$$

(Critical F at 5% level = 2.75)

Model D-2.1 vs Model D-2.3

$$F_{(4,62)} = \frac{\frac{3.340 - 3.176}{4}}{.0427} = 1.24$$

(Critical F at 5% level = 2.52)

Model D-2.2 vs Model D-2.3

$$F_{(1,65)} = \frac{\frac{3.255 - 3.126}{1}}{.0421} = 3.06$$

(Critical F at 5% level = 3.99)

The equations in Model D-2 proved not to be significantly different statistically. Since Model D-2.3 contained only three variables and estimated distribution cost per case of eggs as well as the other two equations, it was selected as the most significant of the Model D-2 equations developed.

The equations selected from each set of models, as the superior estimating equation developed from that set of variables, were compared next:

Model D-1.3

vs

Model D-2.3

$$F_{(2,64)} = \frac{\frac{3{,}126 - 3{,}092}{2}}{.0434} = .391$$

(Critical F at 5% level - 3.14)

The test showed that no difference existed in the statistical ability of Model D-1.3 or Model D-2.3 to estimate distribution cost per case of eggs from the data available. Economically, there was no way to conclude that one variable was superior to another in its relationship to distribution cost. Therefore, Model D-2.3 was selected as the equation that best estimated distribution cost per case of eggs.

$$Y = .35 - \underset{(.003)}{.02X_{10}} + \underset{(.04)}{.20X_2} - \underset{(.0000003)}{.0000007X_1}$$

where

Y = distribution cost per case of eggs in cents,

X_{10} = cases distributed per man-hour,

X_2 = wage rate per hour, and

X_1 = total volume of eggs distributed in cases.

This model indicates the functional relationship of the distribution cost per case of eggs and the three cost factors. The variation in cases distributed per man-hour, wage rate per hour, and total volume of eggs distributed in cases explained 53 percent of the variability in distribution cost per case of eggs. If the case of eggs distributed per man-hour and total cases of eggs distributed are increased by one unit each, then on the average the distribution cost will decrease by .02 and .0000007 cents per case respectively. The reverse is true with regard to wage rate. A one dollar per hour increase in wage rate will on the average increase distribution cost by 20 cents.

III. SUMMARY OF COST FUNCTIONS

A stepwise regression analysis was employed to develop cost equations from the data. The data were divided into two sets due to the high degree of multicollinearity. Estimating equations for both assembly cost and distribution cost were developed from each set of variables. By the use of statistical tests, it was proven that one set of variables (Model A-2.1 and D-2.1) was statistically superior to the others (Model A-1.1 and D-1.1) in their abilities to estimate assembly cost.

Model A-2.3 explained 76 percent of the variability in assembly cost per case with a standard error of estimate of 11 cents. All of the variables in this equation were statistically significant at the 5 percent level, and all of the beta coefficients were significant at the 5 percent level also. The signs associated with each variable were consistent with economic theory.

The statistical test concluded that no statistical difference existed between the estimating equations developed from the distribution data. From the statistical standpoint, it was concluded that Model 2.3, which contained three cost variables, was the most appropriate equation. The signs in the equation were economically sound, and the economic relationship between the endogenous and exogenous variables was homogenous. The model explained 53 percent of the variability in distribution cost per case with a standard error of estimate of 21 cents.

The estimating equations for assembly and distribution cost were similar with regard to the variables included in the equation. Both assembly and distribution are functions of wage rates and the cases of eggs handled per man-hour. These two cost factors were statistically significant in both equations. The equations differed in that a large portion of the variability of assembly cost per case was associated with the percent of the truck capacity utilized, whereas in the distribution cost this factor was not statistically significant. In the distribution equation, the total volume of eggs distributed, in cases, was statistically significant, whereas in assembly it was omitted.

In the A-1.1 models, it was possible to explain a greater percent of the variation of the cost. With identical cost factors for both operations, this result needs explaining. By looking at some of the variations in factors relative to each operation, it is possible to understand why this happened.

Variation in truck capacity utilized in assembly ranged from 18 percent to 100 percent while in distribution the range was 66 percent to 100 percent. This gives an insight into the significance of this variable in assembly cost. The variation in the percent of the truck capacity utilized alone had a correlation coefficient of -.50, while in distribution the correlation coefficient was only -.04.

The reverse was true when the total volume of eggs was considered. The simple correlation coefficient of this variable was .11 in distribution and .07 in assembly. Thus, it is possible to say that in the analysis of assembly cost the more important variables were analyzed, whereas in the distribution analysis some important variables were not present in the data.

CHAPTER IV

ECONOMIES OF SCALE

A stepwise regression analysis was employed to determine if economies of scale were present in the assembly and distribution operations of the Tennessee egg industry. As explained in Chapter II, two approaches were developed to test the hypothesis that economies of scale were present in these operations. The first approach utilized simple linear regression equations, while the second approach employed dummy variables. The linear regression equation for assembly and distribution cost was developed first. Then the dummy variables were employed to determine if selected size groups, in both assembly and distribution, produced different intercept changes, slope changes, and both intercept and slope changes. From this type analysis, it was possible to compare the results of each method used.

I. LINEAR MODELS

As explained in Chapter II, four linear regression models were used in this part of the study. The equations for assembly cost per case were developed first.

Assembly. Assembly cost was analyzed by the four regression equations with the following results:

Model A-1: $$Y = .28 - .0000007X_1$$
$$(.0000005)$$
$$R^2 = .08 \qquad Se = 21 \text{ cents}$$

Y = assembly cost per case of eggs

X_1 = total volume of eggs assembled per year in cases

Model A-2: $$Y = .34 - .000003X_1 - .000000000008X_1^2$$
$$(.000002) \quad (.000000000005)$$
$$R^2 = .16 \qquad Se = 20 \text{ cents}$$

Model A-3: $$Y = .17 + 389.4 \frac{1}{X_1}$$
$$(32.9)$$
$$R^2 = .84 \qquad Se = 9 \text{ cents}$$

Model A-4: $$\log Y = 1.20 - .22 \log X_1$$
$$R^2 = .46 \qquad Se = 16 \text{ cents}$$

After the synthesis of each of the models, it was necessary to determine if one model was statistically superior to the others in its ability to estimate assembly cost per case. The F-test discussed in Chapter II was utilized. It was hypothesized that the equations were not statistically different at the 5 percent level of significance. Tentatively choosing Model A-3 as the best model due to its large R^2 and small Se, the following tests were made.

Model A-1 vs Model A-3

$$F_{(1,28)} = \frac{\frac{1.190 - .2095}{1}}{.0441} = 22.24$$

(Critical F at 5% level = 4.20)

Model A-2 vs Model A-3

$$F_{(1,28)} = \frac{\frac{1.080 - .2095}{1}}{.0415} = 20.99$$

(Critical F at 5% level = 4.20)

Model A-3 vs Model A-4

$$F_{(1,28)} = \frac{\frac{.7007 - .2095}{1}}{.0260} = 18.88$$

(Critical F at 5% level = 4.20)

The null hypothesis was rejected in all the tests. Model A-3 was proven to be statistically better than the other models. The model also was the best with respect to the estimation of economies of scale in the assembly operation. The fit of the curve through the data was much better in this model than any of the others. The quasi economies of scale equation had the following properties:

$$Y = .17 + \underset{(32.9)}{389.4}\,\frac{1}{X_1}$$

$R^2 = .84$ Se = 9 cents

The equation was economically sound with respect to the relationship of the exogenous and endogenous variables. The beta coefficient was statistically significant and the standard error of estimate was the lowest obtained from the group of models.

Distribution. Distribution cost was analyzed by four regression equations with the following results:

Model D-1: $$Y = .15 - .000002X_1$$
$$(.0000005)$$
$$R^2 = .11 \qquad Se = 28 \text{ cents}$$

Model D-2: $$Y = .59 - .000006X_1 - .00000000002X_1^2$$
$$(.000002) \quad (.000000000006)$$
$$R^2 = .21 \qquad Se = 27 \text{ cents}$$

Model D-3: $$Y = .30 + .1568.4 \frac{1}{X_1}$$
$$(344.48)$$
$$R^2 = .24 \qquad Se = 26 \text{ cents}$$

Model D-4: $$\log Y = 1.65 - .28 \log X_1$$
$$(.06)$$
$$R^2 = .26 \qquad Se = 26 \text{ cents}$$

The same test procedures employed in the test of the assembly models were used to test the distribution models. The null hypothesis that the models were not statistically different was tested with the following results:

Model D-1 vs Model D-3

$$F_{(1,68)} = \frac{\frac{5.275 - 4.541}{1}}{.0678} = 10.82$$

(Critical F at 5% level = 3.98)

Model D-2 vs Model D-3

$$F_{(1,68)} = \frac{\frac{4.713 - 4,541}{1}}{.0678} = 4.27$$

(Critical F at 5% level = 3.98)

Model D-3 vs Model D-4

$$F_{(1,68)} = \frac{\frac{4.541 - 4.419}{1}}{.0660} = 1.81$$

(Critical F at 5% level = 3.98)

The null hypothesis was accepted in the first two tests and rejected with regard to Model D-3 and Model D-4. Model D-3 was selected as the equation to use in developing the quasi economies of scale curve even though there was no significant statistical difference in it and Model D-4. The standard error of estimate was the same in both models, and the coefficient of determination in both models were about the same. Due to this equation form being significant in the assembly analysis, it was chosen as the best equation to estimate economies of scale in the distribution operation. The calculated equation was:

$$Y = .30 + \underset{(344.48)}{1568.4} \frac{1}{X_1}$$

$R^2 = .24$ Se = 26 cents

The explained variation in the assembly cost per case of eggs was 24 percent. The standard error of estimate was 26 cents which was the lowest obtained in the analysis. The beta coefficient was statistically significant at the 5 percent level, and the sign of the coefficient was economically sound.

II. DUMMY VARIABLE MODELS

To further explore the hypothesis that economies of scale exist in the assembly and distribution operation of the egg industry, dummy variables were employed. As explained in Chapter II, dummy variables were most appropriate in developing separate cost equations for selected groups of egg handlers. The criteria and procedures necessary to develop the use of dummy variables in this part of the study were discussed in Chapter II. The generated variables were entered into a stepwise regression analysis as a means by which to develop the quasi economies of scale curve. The intercept value was developed first, the slope values second, and the third part of the analysis encompassed both intercept and slope values for each size group. The assembly data were analyzed first.

Assembly. Holding the slope coefficient for each of the three classes (small, medium and large) constant, the intercept coefficient was developed for each class of assemblers as explained in Chapter II. The following equation was developed:

$$Y = .39 - .0000002X_1 - .23X_{13} - .19X_{14}$$
$$(.0000006) \quad (.09) \quad (.12)$$
$$R^2 = .25 \qquad Se = 20 \text{ cents}$$

X_1 = total volume of eggs assembled, in cases per year

X_{13} = 1 when medium class, 0 elsewhere

X_{14} = 1 when large class, 0 elsewhere

The t-test was used to test the null hypothesis that the intercept values differed for the various classes.

$$t_{(23)} = \frac{.23-0}{.09} = 2.56 \qquad t_{(23)} = \frac{.19-0}{.12} = 1.58$$

(Critical t at 5% level of significance = 2.07)

The test showed that the intercept value for the small class and the medium class assemblers was statistically different, whereas no significant difference was present between the small class and the large class. From economic theory, it was concluded that the average fixed cost for the small and medium class assemblers was different at zero level of production. It also produced evidence that the average fixed cost for the small class assemblers was not statistically different from the average fixed cost of the large class of assemblers at zero level of production.

The slope coefficients were determined next.

$$Y = .36 - .000008X_1 + .000004X_{13}X_1 + .000007X_{14}X_1$$
$$(.00001) \quad (.00001) \quad (.00001)$$
$$R^2 = .15 \qquad Se = 21 \text{ cents}$$

X_1 = total volume of eggs assembled, in cases per year

X_{13} = 1 when medium class, 0 elsewhere

X_{14} = 1 when large class, 0 elsewhere

The hypothesis that the slope coefficients were not statistically different at the 5 percent level of significance was tested by the t-test.

$$t_{(23)} = \frac{.000004 - 0}{.00001} = .40 \quad t_{(23)} = \frac{.000007 - 0}{.00001}$$

(Critical t at 5% level of significance = 2.07)

The test concluded that the slope coefficients for each class of assemblers was not statistically different from zero. These results indicated that the slope of the cost functions for the various sizes of firms did not change as the volume of eggs handled increases.

The third step was to develop a regression equation allowing both intercept and slope changes for each class of distributors. The equation developed was:

$$Y = .70 - .00005X_1 - .56X_{13} - .52X_{14} + .00005X_{13} + X_1 + .00004X_{14}X_1$$
$$(.00001) \quad (.17) \quad (.14) \quad (.00001) \qquad (.00001)$$
$$R^2 = .51 \qquad Se = 17 \text{ cents}$$

The parameters were tested by the "rule of thumb" t-test to shorten the test procedures. This test can be tested for accuracy by employing the same t-test procedures that were used in the intercept and slope change analysis. The "rule of thumb" t-test is:

if $Sa > 1/2\ a$, reject Ho: $a_i = A$, 5% critical level

if $Sb > 1/2\ b$, reject Ho: $b_i = B$, 5% critical level

Sa represents the standard error of estimate of the intercept value.

Sb represents the standard error of estimate of the slope value.

The test rejected the null hypothesis in all cases. Therefore, it was concluded that the cost associated with each of the classes was statistically different. The reason that the slope coefficients were significant in this analysis, whereas they were not in the previous analysis, was due to the integration of both intercept and slope variation into the regression analysis. The estimating equations for each class were:

Small class: $Y = .70 - .00005X_1$

Medium class: $Y = .14 + .0000006X_{13}X_1$

Large class: $Y = .18 - .0000001X_{14}X_1$

When these equations were analyzed from an economic standpoint, it was apparent that economies of scale exist in the small and large class assemblers of eggs. The medium class estimating equation does not conform to the economic theory relevant to economies of scale. The small and large class assemblers showed a decrease in assembly cost as the volume of eggs assembled was increased, whereas a one unit increase in eggs assembled for a medium class firm results in a .000006 unit increase in cost.

Distribution. The same format was utilized in the analysis of economies of scale in the distribution operation as was used in the assembly analysis. The intercept coefficients for each class of distributors were developed first.

$$Y = .63 - .0000002X_1 - .17X_{13} - .34X_{14}$$
$$(.0000006) \quad (.08) \quad (.10)$$

By employing the "rule of thumb" t-test, it was apparent that the alpha coefficients for the medium and large class distributors were statistically different from zero at the 5 percent level of significance.

The synthesis of the beta coefficients for each class produced the following equation:

$$Y = .54 - .000002X_1 - .000005X_{13}X_1 - .00000001X_{14}X_1$$
$$(.0000005) \quad (.000005) \quad (.000001)$$
$$R^2 = .13 \qquad Se = 28 \text{ cents}$$

None of the beta coefficients are statistically different at the 5 percent level.

The alpha and beta coefficients were allowed to vary in the third step of this analysis.

$$Y = .63 - .0000002X_1 - .02X_{13} - .34X_{14} - .00001X_{13}X_1 - .0000001X_{14}X_1$$
$$(.0000006) \quad (.17) \quad (.10) \quad (.00001) \quad (.00001)$$
$$R^2 = .27 \qquad Se = 26 \text{ cents}$$

The t-test shows that only one parameter was statistically significant at the 5 percent level. The intercept value for the large class was statistically different from the small and medium classes. The equations for each class of distributors were:

Small class: $Y = .63 - .0000002X_1$

Medium class: $Y = .61 - .0000102X_3X_1$

Large class: $Y = .29 - .0000003X_4X_1$

It was apparent that the estimating equations for the small and medium class were very similar. The large class equation was different from the other two, and it implied that the cost of distribution was lower for the large class distributors.

III. SUMMARY OF ECONOMIES OF SCALE

Both the techniques utilized to study the relationship between size of firm and the cost of operation, indicated that as the volume of eggs handled increased over the complete volume range the cost decreased. All of the analysis of economies of scale were consistent with economic theory except for the increasing cost that occurred as the volume assembled increased within the medium class. This may be explained by having fitted the curve to the data by least squares. This technique only gives an estimate of economies of scale since it is fitted through the data rather than at minimum points of the data.

The dummy variable approach was in no way superior to the linear model approach. In the assembly analysis, the linear model approach was far superior to the dummy variable approach. The linear model method was able to explain 84 percent of the variability in assembly cost, whereas the dummy variable approach explained 51 percent. The amount of explained variation in distribution cost was very similar in both procedures, but the theory of economies of scale was best developed through the use of the linear model.

From this analysis, it was concluded that economies of scale were present in both the assembly and distribution operations of the Tennessee egg industry. It was also conclusive that the following regression equation best estimated the economies of scale curve.

$$Y = a + b_1 \frac{1}{X_1}$$

CHAPTER V

GENERATED COST FUNCTIONS

In Chapter III, Model A-2.3 and D-2.3 were proven to be the superior equations with regard to estimating assembly and distribution cost per case of eggs. Both of the equations were of linear form, and they produced a linear line when fitted through the data. In Chapter IV, Model A-3 and Model D-3 were proven to be the superior quasi economies of scale equation. These equations were of linear form, but they produced a curved line when fitted through the data.

By employing the results from each of these analysis, it was possible to generate cost curves that would develop the characteristics of the quasi economies of scale curve. From Chapter III, the variables that explained the greatest amount of cost variability in linear form were obtained. From Chapter IV, the function form of the volume variable that best fitted the data was obtained. The assembly model was developed first.

I. ASSEMBLY

Model A-G. Model A-2.3 ($Y = a + b_2X_2 + b_8X_8 + b_{10}X_{10}$) was the superior assembly cost equation in Chapter III, while Model A-3 ($Y = a + b_1 \frac{1}{X_1}$) was the superior assembly quasi economies of scale model in Chapter IV. All of the variables in each equation were entered into one equation and analyzed by stepwise regression. The following equation was developed:

$$Y = .24 + 316.07 \frac{1}{X_1} + .05X_2 - .007X_{10}$$
$$(25.84) \qquad (.02) \quad (.001)$$
$$R^2 = .94 \qquad \qquad Se = 6 \text{ cents}$$

The new equation excluded variable 8 (percent truck capacity utilized) since its slope coefficient was not statistically different from zero at the 5 percent level. To compare the significance of this model with that of Model A-2.3, the F-test as used in Chapters III and IV was employed.

Model A-2.3

vs $\qquad F_{(1,26)} = \dfrac{\frac{.9636 - .0846}{1}}{.0036} = 274.69$

Model A-G

(Critical F level at 5% level = 4.22)

From this test, it was evident that the generated model was greatly superior to Model A-2.3 in its ability to explain the variation in assembly cost per case of eggs.

II. DISTRIBUTION

Model D-G Model D-2.3 ($Y = a + b_1X_1 + b_2X_2 + b_{10}X_{10}$) was the superior distribution model in Chapter III, whereas Model D-3 ($Y = a + b \frac{1}{X_1}$) was the superior quasi economies of scale model. Since Model D-2.3 contained the volume variable (X_1), it was necessary to exclude it before combining the two equations. The variables $\frac{1}{X_1}$, X_2 and X_{10} were entered into the generated model and produced the following equation:

$$Y = .25 + 812.10 \frac{1}{X_1} + .18X_6 - .02X_{10}$$

$$\qquad\qquad (291.90) \qquad (.04) \qquad (.003)$$

$R^2 = .55$ Se = 20 cents

All of the variables are statistically significant at the 5 percent level. This model was compared to Model D-2.3 by the F-test.

Model D-2.3

vs

Model D-G

$$F_{(1,66)} = \frac{\frac{3.126 - 2.655}{1}}{.0408} = 11.54$$

(Critical F at 5% level = 3.99)

As in the assembly analysis, the generated model was proven to be statistically superior to the linear model developed in Chapter III.

III. SUMMARY OF GENERATED COST FUNCTIONS

The combination of the two models in both assembly and distribution proved valuable. Statistically, the new equations were superior to those developed in Chapter III. The amount of explained variation in assembly cost was increased 26 percent, while the explained variation in distribution cost was increased 4 percent. The standard error of estimate was reduced by 5 cents for assembly and 1 cent for distributions. All of the beta coefficients were statistically significant at the 5 percent level in both sets of equations, but in the generated models the coefficients were more significant.

Economically, the generated models proved most valuable. The variables included in the generated model were consistent with economic theory. Each of the variables was easy to obtain from records kept by the firms. A firm normally keeps accounting records of the volume of eggs handled, hours the employees work, and the wage rate paid to the workers. By utilizing these variables, a firm may make decisions on how to reduce cost of assembly or distribution. Basically each of the equations employes efficiency of the employees, total output of the firm, and wage rate to estimate the cost of assembly and distribution of a case of eggs for a given firm.

Even though these equations were acceptable ways to estimate assembly and distribution cost, they did not specify ways to reduce these costs. The "how" with regard to reducing cost was not available in the historical type data used. From the equation, it was apparent that a reduction in wage rate would reduce cost, but the ability of a firm to reduce wage rates was very limited due to government regulation of these rates. It was also apparent that an increase in employer efficiency would reduce cost, but the information obtained from accounting data was not relevant to such an analysis. To analyze efficiency, the firms needed an analysis of economic engineering data that contains the components of the operations. An increase in volume of eggs handled would also reduce cost, but the economic information for such an attempt was not available in the type data used in this analysis. For example, the demand for eggs would have to be considered before the volume was increased.

Therefore, it is important to realize that this analysis developed estimating equations from accounting data. The equations show the firm what factors affect the cost the greatest amount and the relationship of these variables to the total cost. They do not explain how to change the variables and reduce total cost.

CHAPTER VI

SUMMARY

The basic objectives of this study were: (1) to develop economically and statistically sound cost equations for the assembly and distribution operations of the Tennessee egg industry, and (2) to determine if economies of scale were present in the assembly and distribution operations.

The data were obtained from a sample population of 75 egg handling firms in Tennessee. The data were collected from accounting records and categorized as assembly and distribution cost factors. A total of seven cost factors were obtained directly from the survey data, while four were calculated from these factors.

The data were analyzed by a stepwise regression analysis. The analysis was divided into the assembly analysis and the distribution analysis, with each analysis employing the same statistical procedures. In the cost function analysis, two sets of variables were utilized in various functional forms to synthesize the ultimately superior equation for the assembly and distribution operation. Various regression equation forms were developed by applying both statistical and economical constraints. In the development of quasi economies of scale curves, two statistical procedures were developed. Four regression equations of linear form were employed first, and then the sample was divided into three classes. It was then possible to employ dummy variables as a means by which to develop a linear cost equation for each of the three classes.

In the analysis of the cost function, seven separate sets of variables were used. The seven equations were compared both statistically and economically, whereby it was possible to select the superior equations with regard to estimating assembly and distribution cost.

The superior equation with relation to estimating assembly cost was:

$$Y = .24 + \underset{(25.84)}{316.07} \frac{1}{X_1} + \underset{(.02)}{.05X_2} - \underset{(.001)}{.007X_{10}}$$

$$R^2 = .94 \qquad\qquad Se = 6 \text{ cents}$$

The equation was consistent with economic theory. Volume of eggs assembled (X_1), wage rate (X_2), and cases assembled per man-hour (X_{10}) were utilized in this equation to estimate assembly cost per case of eggs. These three variables have the capacity to explain 94 percent of the variability in assembly cost per case of eggs with a standard error of estimate of six cents. The equation also explains that a one unit increase in cases of eggs assembled (regressed in reciprocal form) will decrease assembly cost, on the average, by 316.07 units, whereas a one unit increase in wages will increase cost by .05 units. It is also conclusive that a one unit increase in labor efficiency, measured by cases assembled per man-hour, will decrease cost by .007 units.

The superior equation with relation to estimating distribution cost was:

$$Y = .25 + 812.10 \frac{1}{X_1} + .18X_2 - .02X_{10}$$
$$(291.90) \quad (.04) \quad (.003)$$
$$R^2 = .55 \qquad Se = 20 \text{ cents}$$

This equation also utilized the economically sound variables (volume, wage rate, and cases distributed per man-hour) to explain the variation in distribution cost per case of eggs. The equation explains that a one unit increase in cases of eggs distributed (regressed in reciprocal form) will, on the average, decrease cost per case of eggs by 812.0 units, and a one unit increase in cases of eggs handled per man-hour will decrease cost per case an average of .02 units. An increase in wage rates by one unit will increase cost by .18 units per case of eggs distributed. The equation has the capacity to explain 55 percent of the variability of distribution cost with a standard error of 20 cents.

Economically, these two equations proved to be most valuable to the assemblers and distributors of eggs. By utilizing these equations, a firm may make decisions on how to value their cost of operation. The equations do not tell "how" to reduce cost, but they do give an insight into the areas that will most adequately reduce the cost.

In the analysis of economies of scale, it was conclusive that economies of scale were present. A linear regression equation that produced curved lines when plotted was selected as the best equation for developing quasi economies of scale curves. The estimated economies of scale equation for the assembly operation was:

$$Y = .17 + 389.4 \frac{1}{X_1}$$
$$(32.9)$$
$$R^2 = .84 \qquad Se = 9 \text{ cents}$$

This equation has the capacity to develop a quasi economies of scale curve with 84 percent of the variability in assembly cost explained with a standard error of estimate of 9 cents.

The quasi economies of scale curve may best be produced in this distribution operation by the following equation:

$$Y = .30 + .1568.4 \frac{1}{X_1}$$
$$(344.48)$$
$$R^2 = .24 \qquad Se = 26 \text{ cents}$$

This equation does not explain as much of the variation in cost as the assembly equation, but it was the best obtainable by the method of analysis of the data in this study. It is still economically important since it concludes that economies of scale are present in the distribution operation.

This segment of the study presents information to the firms which increases a priori knowledge relevant to size of plant decisions. This is not a true economies of scale analysis since the curves are fitted by least squares and not through the minimum cost points.

The objectives of this study were accomplished, and the results were favorable. By employing the cost functions developed in this study, an assembly or distribution firm has a guide which may help in making important decisions with regard to minimization of cost. The quasi economies of scale curves developed add another tool to the arsenal of information needed by egg handlers in order to stay competitive in the industry.

I. LIMITATIONS OF THE STUDY

The data employed in this study were obtained from accounting records. This data exhibits the historical operations of the firms, thus when analyzed, it only produces a picture of what has happened. It tends to ignore the question of what could have happened since the element of "opportunity cost" is not obtainable. The analysis of accounting data, due to the absence of many factor cost associated with a given operation or the combining of many factor costs into one major cost, also ignores the question of "why" and "how" such things as economies of scale are relevant to a given operation.

To assess "cost" in a manner that would answer such questions as what could have happened, it is necessary to obtain information from the actual operation of the firm. This engineering type data allows the breakdown of each individual phase of the operation. The total cost picture may then be obtained by summing the individual stages. By breaking each of the operations into elementary stages, it is possible to derive an answer to what could have happened under given conditions. The analysis of the elementary cost factors also enable the study of how and why a cost is relevant to a given firm or a given operation.

Thus, it is necessary to realize that this study produces cost functions and quasi economies of scale functions from information of past operations. This study does not develop methods by which a firm may pattern changes that will allow it to arrive at a minimum cost in an attempt to reach the theoretical minimum of an economies of scale curve or even to change a given factor cost. This study only has the ability to point out areas where potential savings exist.

SELECTED BIBLIOGRAPHY

SELECTED BIBLIOGRAPHY

A. BOOKS

Ferber, Robert, and P. V. Verdoorn. Research Methods in Economies and Business. New York: The MacMillan Company, 1962.

Fox, Karl A. Intermediate Economic Statistics. New York: John Wiley and Sons, Inc., 1968.

Friedman, Milton. Price Theory. Chicago: Aldine Publishing Company, 1967.

Hoel, Paul G. Elementary Statistics. New York: John Wiley and Sons, Inc., 1967.

Johnston, J. Econometric Methods. New York: McGraw-Hill Book Company, Inc., 1960.

Yamane, Taro. Statistics: An Introductory Analysis. New York: Harper and Rowe, 1967.

B. AGRICULTURAL EXPERIMENT STATION PUBLICATIONS

Jones, H. B. Economies of Scale in Commercial Egg Packing Plants. University of Georgia Experiment Station Bulletin No. N.S. 120. Athens: University of Georgia, September, 1966.

Jones, H. B., and H. R. Smolly. Vertical Interpretation Methods of Producing and Marketing Eggs: An Economic Evaluation. University of Georgia Experiment Station Bulletin No. N.S. 120. Athens: University of Georgia, May, 1966.

White, Morris. Relationship of Marketing Methods to Cost of Assembly, Grading, and Packaging Table Eggs. Auburn University Experiment Station Bulletin No. 348. Auburn: Auburn University, October, 1963.

C. JOURNAL ARTICLES

Bressler, R. G., Jr. "Research Determination of Economies of Scale," Journal of Farm Economics, 27:526-539, August, 1945.

Tomek, W. G. "Using Zero-One Variables with Time Series Data in Regression Equations," Journal of Farm Economics, 45:814-822, November, 1963.

D. GOVERNMENT PUBLICATIONS

United States Department of Agriculture. Poultry and Egg Situation. Washington: Government Printing Office, November, 1968.

United States Department of Agriculture. Tennessee Agriculture Statistics. Washington: Government Printing Office, August, 1968.

E. OTHER SOURCES

Ben-David, Shaul, and W. G. Tomek. "Allowing for Slope and Intercept Changes in Regression Analysis." Unpublished Ph.D. thesis, Cornell University, 1968.

The University of Tennessee Computing Center. "BMDO2R Stepwise Regression." Knoxville: The University of Tennessee, September 1, 1965.

VITA

Billy W. Nelms was born in Rives, Tennessee, on September 18, 1940. He was educated in that city's schools and was graduated from Rives High School in 1958. In October, 1959, he entered the United States Air Force, receiving an honorable discharge in September, 1963. That same month he entered The University of Tennessee at Martin, receiving a Bachelor of Science degree in Agriculture in June, 1967. The following September, he entered the Graduate School of The University of Tennessee and is presently engaged in fulfilling the necessary requirements for a Master of Science degree.

He is married to the former Linda Faye Sellers of Elbridge, Tennessee.