MILLER, Thomas Alan, 1937-
AGGREGATION ERROR IN REPRESENTATIVE FARM
LINEAR PROGRAMMING SUPPLY ESTIMATES.

Iowa State University of Science and Technology, Ph.D., 1967
Economics, agricultural

University Microfilms, Inc., Ann Arbor, Michigan

AGGREGATION ERROR IN REPRESENTATIVE FARM
LINEAR PROGRAMMING SUPPLY ESTIMATES

by

Thomas Alan Miller

A Dissertation Submitted to the
Graduate Faculty in Partial Fulfillment of
The Requirements for the Degree of
DOCTOR OF PHILOSOPHY

Major Subject: Agricultural Economics

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

Head of Major Department

Signature was redacted for privacy.

Dean of Graduate College

Iowa State University
Of Science and Technology
Ames, Iowa

1967

TABLE OF CONTENTS

	Page
CHAPTER I. INTRODUCTION	1
General Problem Area	1
The Aggregation Error Problem	6
Objectives of the Study	9
CHAPTER II. A REVIEW OF REPRESENTATIVE FARM SUPPLY ESTIMATION AND AGGREGATION ERROR	11
Use of the Representative Farm	11
The Aggregation Problem	16
Solutions to the Aggregation Problem	20
CHAPTER III. SUFFICIENT CONDITIONS FOR ERROR-FREE AGGREGATION	27
The Aggregation Problem	27
Exact Aggregation Over All Products	31
Exact Aggregation of One Product	52
CHAPTER IV. ADDITIONAL THEORY ON AGGREGATION ERROR	57
Direction of Aggregation Error	57
Situations Causing Aggregation Error	78
CHAPTER V. PROGRAMMED SUPPLY FUNCTIONS FROM DIFFERENT STRATIFICATIONS	88
The Population Defined	89
The Sampling Procedure and Data	90
The Stratification Procedure	91
The Representative Farm Models	100
The Resulting Supply Functions	112

	Page
CHAPTER VI. EVALUATION OF PROGRAMMED SUPPLY FUNCTIONS	119
State Supply Estimates	119
Area 3 in Detail	134
CHAPTER VII. GUIDES TO NEW STRATIFICATIONS	145
Postprogramming Analysis of the 36 Farms	145
New Delineations of Representative Farms	153
CHAPTER VIII. SUMMARY	158
The Problem and Objectives	158
The Theory of Aggregation Error	159
Aggregation Error in Empirical Work	162
Representative Farm Linear Programming Models	165
ACKNOWLEDGEMENTS	168
BIBLIOGRAPHY	169
APPENDIX	174

CHAPTER I. INTRODUCTION

General Problem Area

The analysis of regional and national supply relationships is an important part of the study of adjustments in agriculture. Knowledge of supply relationships provides a basis for understanding agricultural adjustments to changing demands, changing input supplies and new technologies. An understanding of both supply relationships and the agricultural adjustment process is invaluable to decision makers at all levels of the agricultural economy. It enables farmers to plan their operations for higher profits. It allows farm input suppliers to accurately predict the demand for their products. It provides policy makers and consumers with better insights into the changing role of agriculture.

This thesis is concerned with one of the inherent problems of a commonly used tool of supply analysis. This is the problem of aggregation error in representative farm linear programming models. As a first step, it is desirable to review briefly the use of representative farm linear programming models as a technique of supply analysis.

Supply estimation models

The analysis of supply relationships generally begins with the estimation of supply functions. Supply functions are defined as the relationship between the price of a commodity and the quantity supplied by producers. Estimation of these supply functions in agricultural economics research is accomplished primarily by two methods -- linear regression analysis of time-series data and "synthetic" supply estimation using linear programming. Regression analysis is useful to determine supply relationships at a relatively high level of aggregation. It requires adequate time-series data for all important independent variables postulated to affect the quantity of a commodity supplied. The effect of each independent variable comes directly from the estimated regression equation. The supply function is estimated by the price-quantity relationship, and the other independent variables in the equation are interpreted as shifters of this supply function.

Linear programming supply estimation uses a variable pricing model of a producing unit. The producing unit may be either a farm or a region. The variable pricing process generates a synthetic supply curve for this unit, subject to the usual assumptions of linear programming. The programmed supply curves of individual producing units are then summed into area, regional or industry aggregate estimates, depending on the definition of the original producing unit and the estimate desired. This approach generally shows what farmers would produce if they maximized profits. This technique is particularly useful in determining the production potential of the agricultural industry.

In regression models, the effect of past changes in both the structure of agriculture and in technology can be determined by means of dummy variables and time trends. As long as these changes are gradual and continue at the historical rate, regression estimates provide good extrapolations into the future. It is difficult, however, for regression models to appraise the impact of anticipated changes in agriculture when no historical data are available. For example, people who are responsible for formulating agricultural policy or appraising alternative farm programs are often faced with the task of determining the likely outcome of specific program provisions. It is common for these new programs to provide a new set of institutional restraints for which there is little or no historical basis for estimating response. Regression models are inadequate for answering questions of this type.

Linear programming models possess a potential for appraising the effect of new variables on supply. These models simulate the decision-making process of the producing unit under study. The response of the producing unit to any new variable can then be estimated irrespective of the variable's historical relevance. This application of linear programming is relatively new in economic research. It does have the potential for improving forecasts of supply relationships within a rapidly changing agriculture.

Representative farm supply models

The unit of analysis (producing unit) in linear programming models may be a region such as the Corn Belt, a smaller area such as northeast Iowa, a group of farms such as Iowa cash-grain farms or the individual farm itself. The well-known Heady-Egbert models are examples of studies considering the producing region as the unit of analysis (11, 12 and 22). Alternatively, many recent adjustment studies have treated individual or representative farms as the unit of analysis (17, 25 and 36). The supply curves for the individual farms are then summed or aggregated into the desired regional supply estimates.

Use of the representative farm as the unit of analysis has several advantages over area linear programming supply models (8, p. 1440 and 34, p. 86). An advantage is that it permits analyzing the impact of aggregate changes in agriculture at the individual farm level -- thus providing a key to the relation between macro and micro variables in the economic relation under study. This analysis makes it possible for policy makers to appraise the effects of alternative programs at both the national level and the individual farm level. Representative farm linear programming models have a definite advantage over both area linear programming models and regression models when this individual farm interpretation is required.

Another advantage of including representative farms in linear programming models is that it simulates the response decisions as being made by the managerial units that actually make them. Such models have the potential of being more realistic than alternative models treating the area as the unit of analysis. Area models must employ the rather abstract concept of the linear programming model simulating the "decision-making process" of an entire producing area.

Representative farm models have an advantage over the area models in allowing restriction of resource mobility among farms in the area. Using the area as the unit of analysis is equivalent to the assumption that all of the area's resources may be fully used. However, subaggregates of the area, usually individual farms, often restrict resource usage to less than the area total. An example is a farm with excess labor available and restrictive land. Programming individual farms and restricting resource mobility among farms will force this labor to go unused. However, treating the area as the unit of analysis will allow full use of this farm's labor -- equivalent to the assumption of sufficient resource mobility to achieve full use of all the area's labor and land.

The representative farm model estimation process

The previous discussion provides background for understanding the technique of synthetic supply estimation by representative farm linear programming. Normally, the estimation process itself is accomplished through the following steps:

(1) The commodity and population are defined for which supply estimates are desired.

(2) Data are collected on a fairly large sample basis on the resources, costs and alternatives found on individual farms in the population.

(3) Generally, sampling rates high enough to achieve desirable reliability of these data will cause the number of sampled individual farms to far exceed linear programming capabilities, if all are to be programmed. Therefore, the next step is to stratify all of the sampled farms into a much smaller number of groups and to delineate a representative farm (sometimes referred to as a typical or benchmark farm) to represent each group.

(4) A linear programming model is developed for each representative farm and that farm's supply function is estimated by variable pricing techniques.

(5) The supply functions of the representative farms are expanded to estimates of the supply functions for each group of sample farms. This expansion provides supply estimates for each of the groups of farms delineated in step 3.

(6) The supply functions of these groups are summed horizontally to obtain supply functions for the individual farms sampled in step 2.

(7) Finally, the sample estimates are expanded into the desired population supply estimates defined in step 1.

There are, of course, some variations to this process. Occasionally a complete sample of individual farms is obtained in step 2, making step 7 unnecessary. It sometimes is possible to program all farms in the sample, thereby omitting steps 3 and 5. Steps 6 and 7 may often be combined into one operation. However, the basic process is the one outlined in steps 1 through 7.

The Aggregation Error Problem

Aggregation error is one of three possible sources of error in representative farm linear programming supply estimates. These three sources or types of error have recently been discussed quite aptly by John Stovall (45, p. 478):

> 1. Specification error arises because the programming model fails to reflect accurately the conditions actually facing the farm firm for a given length of run. Specification error may include errors in the technical coefficients,

the resource restrictions or product and input prices.[1]

2. Sampling error arises when the distribution of the model's parameters over all firms in the population is not known but is estimated by sampling techniques.

3. Aggregation error[2] as defined by Frick and Andrews is "the difference between the area supply function as developed from the summation of linear programming solutions for each individual farm in the area and summation from a smaller number of typical or benchmark farms" [15, p. 696].

Each of these types of errors can be related to the seven steps in the estimation procedure previously outlined. Specification error is associated primarily with step 4 where the individual farm linear programming models are developed and supply functions for each estimated. Sampling error arises in step 2 and in step 7 where the sample data are obtained and the population totals are estimated from sample results. Aggregation error would arise from step 3 and step 5 when the sample farms are stratified into groups, the representative farms are delineated for each group and the representative farm supply functions are aggregated into group supply estimates.

---

[1]This definition does not agree completely with the traditional concept of specification error. Failure to incorporate appropriate activities and restraints and incorrect specification of the objective function are obviously specification errors. In contrast, errors in estimating technical coefficients and product and input prices appear to be more a problem of sampling than a problem of specification.

[2]Stovall feels (and this writer agrees) that the popular use of the term "aggregation bias" is unfortunate because of its connotation of a systematic rather than random error. It is not clear that errors in aggregation are biased, as is discussed in Chapter IV of this thesis.

It is obvious that these three types of error are not independent. Data from the specification and sampling phases are used for the aggregation phase; hence, decisions made for these two phases influence aggregation error. In a similar manner, error arising from the sampling phase may contribute to specification error. Stovall also mentions that these three types of errors are not intended to form an exhaustive set. They do, however, provide a meaningful beginning upon which to build a detailed study of possible errors.

There is a gap in knowledge of both theoretical and practical aspects of the aggregation error problem (33, p. 532). The specification error problem encompasses the whole area of using linear programming in farm management and production economics. This area has been the frequent subject of research projects in the last decade. The problem of sampling error is primarily a problem of statistics. Solutions to it are to be found through the theory of sampling. In contrast, the aggregation error problems involved in the technique of representative farm linear programming supply estimation have often been mentioned in the economics literature, but formal research procedures for reducing these problems have not been developed. As a result, the aggregation error problem was selected as the subject for this thesis.

The problem of dependence among the three types of error is handled by assuming all of the decisions involving specification and sampling have been made. Aggregation error is analyzed within this framework. It is assumed that the linear programming model is an adequate representation of the farm's decision-making process and that the estimated farm's supply curve is correct for the length of run and other economic conditions -- assumptions which hold the specification problem constant. Likewise, the assumption is made that the sampling rate is sufficient to hold sampling error to desirable levels. It is possible to view aggregation error within this framework and work towards procedures for understanding and controlling it.

Objectives of the Study

Aggregation error in representative farm linear programming supply estimates arises from two of the steps in the previously described seven-step estimation process (see p. 5). In step 3 the sample farms are stratified into a smaller number of groups and the representative farms are defined. Relevant questions at this point are (1) what factors should be used in the stratification or grouping of sample farms, (2) how many groups should be developed and (3) how should the representative farms be determined. Aggregation error also arises in step 5 where the supply functions for each of the representative farms are expanded into supply functions for each group of sample farms. Questions here pertain to determination of the correct expansion factors or aggregation coefficients. Current suggested solutions to these problems appear to be based mainly on intuitive concepts that are expected to minimize aggregation error.

The objectives of this study are therefore as follows:

(1) To review the current state of literature and theory relating to the problem of aggregation error and to evaluate these theoretical concepts for practical usefulness.

(2) To explore the theoretical aspects of developing error-free or minimum-error aggregates and if possible to develop additional theory in this area.

(3) To compare empirical supply estimates of Iowa pork and beef resulting from different stratifications of representative farms as a means of determining the relative magnitude of the aggregation error and possible factors that contribute to it.

(4) To utilize the results of the first three objectives to recommend practical procedures that may be followed by research workers in controlling aggregation error.

Generally, these four objectives outline the main subjects for the remainder of this thesis. Chapter II includes a review of the literature that applies to aggregation error in representative farm linear programming supply estimation. Chapters III and IV explore the theory of aggregation error and discuss theoretical concepts that appear useful in controlling aggregation error. Empirical supply estimates are developed in Chapter V for pork and beef in Iowa. These results are then discussed in Chapters VI and VII along with the recommendation of practical procedures that may be followed by researchers in holding aggregation error to tolerable levels. Chapter VIII summarizes the major points of the analysis.

CHAPTER II. A REVIEW OF REPRESENTATIVE FARM SUPPLY ESTIMATION AND AGGREGATION ERROR

Agricultural supply analysis is an important facet of economic research. Cochrane provides an excellent overview of the general supply relation in agriculture, emphasizing the different concepts of supply theory (7). Nerlove and Bachman present a comprehensive survey of the type of research done in this area (33). They categorize the two general empirical methods used in studying supply as (1) the "constructive method" involving derivation of supply functions from a combination of production function data and individual behavior and (2) statistical analysis of time-series data (33, p. 541). This division is consistent with the two general methods discussed in Chapter I. The constructive method employs classical production function analysis, farm budgeting and linear programming of both representative farms and producing areas. The time-series analysis most often involves using multiple regression techniques on data capturing the historical influence of certain independent variables affecting supply. As Nerlove and Bachman point out, these two methods complement each other in many respects.

An excellent bibliography of material discussing agricultural supply analysis appears at the end of the Nerlove and Bachman article (33, pp. 551-554). The method of supply analysis involving linear programming of representative farms is well represented in this bibliography.

Use of the Representative Farm

Representative farms provide a suitable starting place for reviewing literature pertaining to the aggregation problem. The concept of the representative farm (or the equivalent but more general representative firm) dates back some time in economic literature. Marshall noted that the representative firm was the particular kind of firm which could obtain the necessary internal and external economies of production such that the normal supply price of a commodity would just equal the firm's normal expenses of production. He stated that (29, p. 317):

> A representative firm must be one which has had a fairly long life, and fair success, which is managed by normal ability, and which has normal access to economies, external and internal, which belong to the aggregate volume of production; account being taken of the class of goods produced, the conditions of marketing them and the economic environment generally.

Taussig adapted Marshall's idea of a representative firm as " . . . a firm not far in the lead, not equipped with the very latest and best plant and machinery, but well equipped, well led and able to maintain itself permanently with substantial profits" (46, p. 180). He also stated that prices would tend to adjust to the costs of production of such a firm.

Representative farms in farm management

Both Marshall and Taussig considered the representative firm in the more abstract conceptual sense to explain economic principles of supply and equilibrium price rather than as an empirical tool. Elliott reviewed the idea as it applied to agricultural economics research. He defined a representative or typical farm as "a modal farm in a frequency distribution of farms of the same universe; or it is representative of what a group of farmers are doing who are doing essentially the same thing" (13, p. 486).. He viewed the concept as useful in making recommendations to farmers and as useful in budgeting and income studies. He outlined a specific and rather detailed quantitative procedure for identifying the typical farm(s) in a given farming area.

Elliott's later work was a part of the numerous type-of-farming studies that predominated in farm management research during the decade beginning in the 1920's. This work is surveyed and evaluated by Wilcox (52). Generally these studies were based on the premise that blanket recommendations to the so-called "average farmer" were too general to be of much value. Elliott stated (14, p. 2):

> What is needed is a segregation of farmers into specific groups of given sizes of farms and in homogeneous type-of-farming areas so that a correct appraisal can be made of the needs of typical groups and a true interpretation of the effect which changing conditions are likely to have upon them.

Results of such a type-of-farming study were reported for Iowa in 1929. Holmes designated five "farm type areas" in the state (23). These areas generally followed soil-type boundaries and conformed quite closely to combinations of areas used later in Iowa for agricultural adjustment studies.

<u>Representative farms in supply estimation</u>

The general process of estimating aggregate supply functions from representative farms is discussed by Barker and Stanton (2). It had its origin in 1932 when John D. Black created a strong and widespread interest in applying the representative farm budgeting procedure to problems of supply estimation in agricultural economics research. He proposed using what he called the "farm step-up method" of constructing synthetic supply curves, after observing that regression procedures were useful to estimate short-run supply but not much good for long-run estimates. He observed (3, p. 99):

> One would no doubt begin with the existing set of prices, then raise the price of the major product, leaving the price of others as it is now; and for each set of prices determine by observing the probable effect on net income the proportion of products that would be most advantageous.

This information could then be used to construct a synthetic supply curve. Black stated that the advantages of this method were (1) that it gets the information directly in the form needed for forecasting changes in specific types-of-farming and (2) that it is useful in analyzing the possible effects of new institutions and structures where regression procedures could not be extended.

Sherman Johnson and Black later directed a project studying supply responses of milk production to milk prices in five areas of New England and five in the Midwest (31). The project was a cooperative project between U.S. Department of Agriculture, several agriculture experiment stations and the Committee of Research in the Social Sciences of Harvard University. The supply studies were made in 1936-38. This work is interesting since the basic method used in stratifying individual farms and specifying representative farms is still in use at the present time.

The general method used was to budget representative farms to determine their supply response and then multiply these results by appropriate factors to estimate the response of the universe of interest. Six hundred and twelve farms were budgeted throughout the 10 areas studied (31, p. 120). Considerable thought and time were used in developing the representative farms used in the studies. In each of the areas studied, a preliminary survey of 100 to 200 farms was made to serve as a basis for individual farm selection. Farms were then stratified by factors likely to affect response. For example, factors chosen in Wisconsin were (1) land quality, (2) size, (3) existence of family labor, (4) operator age and (5) operator tenure. Farms in these groups were then selected for the final budgeting; a total of 24 farms were used in the Wisconsin example (6).

An analysis was made of the effect of using different numbers of representative farms in the Vermont contributing project to this regional study (1). The particular area of Vermont studied contained 213 individual farms and all were budgeted for the regional study. However, work was also done on a smaller sample of 26 farms " . . . selected on the basis of factors that were thought to be significantly related to production response . . ." (1, p. 50). The sample of 26 farms proved unsatisfactory for the purposes intended. The researchers concluded that a random sample of 50 farms appeared to be the minimum that would be reasonably satisfactory under the conditions found in that area of Vermont. However, it was reported in the regional project report that "with careful judgment in selection, half that number gave good results and involved much less work" (31, p. 120). This statement referred to budgeting a judgment or purposive sample of the individual farms in population of interest.

During the decade of the 1950's the need for up-to-date agricultural supply research was noted by Cochrane, Schultz and others (7, 37).

Empirical supply studies during this period were oriented strongly toward regression analysis of time-series and cross-sectional data. However, interest was renewed in the process of synthetic supply estimation using representative farms with the advent of the electronic computer and linear programming techniques (17, 25, 36). Basically the process used was the same as in the original study led by Johnson and Black with linear programming replacing the former budgeting procedure; McKee and Loftsgard summarize the usual procedure (28). Much of this more recent work was also coordinated by regional cooperative projects between the Economic Research Service, U.S. Department of Agriculture, and the various state experiment stations.[1] The Lake States Dairy Adjustment Study is a completed project of this type (47).

The Aggregation Problem

The aggregation problem has been discussed by economists in various contexts. Thiel and others discussed aggregation in reference to the problem of estimating macro-parameters from their micro-counterparts (48, 24, 30). An example of this problem is the estimation of the relation between total population income, population size and total consumption of a commodity given that consumption of each household is a known

[1]These studies include Appraisal of Opportunities for Adjusting Farming to Prospective Markets (W-54); Supply Response and Adjustments for Hog and Beef Cattle Production (NC-54); An Economic Appraisal of Farming Adjustment Opportunities in the Southern Region to Meet Changing Conditions (S-42); Great Plains Regional Technical Committee on Wheat Adjustment Research (GP-5); Lake States Dairy Adjustment Study; and Northeast Dairy Adjustment Study.

function of family income and family size. The aggregation problem in this context arises from the dependence of one individual's behavior on another. Thiel developed the conditions necessary for "perfect aggregation" and defined an "aggregation bias" that would result if the conditions are not met.

Another problem of aggregation is essentially one of choosing the right level of detail to employ in various parts of linear programming models. In discussing aggregation problems in interregional competition models, Heady divided the problems into those of (1) input or resource aggregation, (2) output or product aggregation, (3) producing unit or firm aggregation and (4) objective function aggregation (19). He felt the major aggregation problem was one of balancing enough detail to give meaningful results against clerical, computing and other limitations. On one hand, the quest for realism makes delineation of many specific products, inputs, producing regions and firms desirable. On the other hand, the number of divisions must be kept low enough to be consistent with the time and facilities available for research.

With the increased use of representative farms in synthetic supply estimation, research workers became more aware of the problems of aggregating farm supply functions into aggregate supply functions. This concept of the aggregation problem is equivalent to the third problem listed by Heady, producing unit or firm aggregation. Nerlove and Bachman listed it as one of the main problems in the representative farm linear programming method of agricultural supply analysis (33). This aggregation problem arises when farm supply functions are used to estimate total or aggregate supply functions and is the primary concern of this thesis.

Generally, researchers felt that solutions to the farm supply function aggregation problem should involve following correct procedures in delineating the typical or representative farms to be programmed. The earlier Johnson and Black work attempted this approach. They delineated homogeneous groups of individual farms with similar response patterns for the particular stimuli under study and budgeted a representative farm for each group. The methods used in the early study were not too successful. Christensen had in fact tested the significance of the five factors used to stratify farms in affecting final results in the Wisconsin portion of the study (6). He reported that no significant relationship could be found between the various stratification factors and changes in production.

Beginning in the late 1950's, several researchers suggested formal procedures for delineating homogeneous groups of farms. Plaxico suggested a method for determining the causal factors of differences in supply relationships (34). This method involved the systematic testing of hypotheses concerning output and various characteristics postulated to affect supply. These hypotheses were tested by programming a small number of farms in each case, fitting regression lines through the points programmed and testing the regression coefficients for significance. He said that ". . . characteristics shown to have no effect or a linear effect on supply need not be considered in developing weighting factors to be used in aggregating the various firm responses" (34, p. 89).

In a paper on problems of defining typical resource situations, Thompson recognized the importance of resource ratios in affecting supply response but doubted that farm supply functions could be aggregated without error. He said (49, p. 42):

> . . . a group of individual farms having different resource ratios will not necessarily allocate their resources in the same manner as would a single farm having resource ratios similar to those of the group as an entity. As a consequence of this, it does not appear possible to choose a typical farm so that the differences in response between the typical and individual farms systematically cancel each other, i.e., so that the response of the entire group is a known function of that of the typical farm.

During this period the term "aggregation bias" became popular. This bias or error was simply the difference between (1) aggregate results obtained from analyzing every farm separately and (2) aggregate results obtained from the short cut of analyzing representative farms. The main problem was then one of choosing the kind and number of representative farms that will minimize the amount of aggregation bias.

Lee Day suggested that the possibilities for completely eliminating aggregation bias were unlikely within a manageable number of representative farms. He cited the need for empirical research to determine the " . . . magnitude and direction of bias at various prices associated with different levels of variance in proportionality of resources before we can reach informed judgments about the number of representative farms and appropriate procedures for structuring these farms" (8, p. 1442).

Carter considered representative farms primarily for the purposes of making recommendations to farmers (5). He felt that many studies passed lightly over the problem of representative farm selection and that "refinements could be made in the actual selection of representative farms if it were possible to isolate the primary characteristics of farms and farmers that tend to dominate or strongly influence the particular decision under study" (5, p. 1454). Carter also observed that characteristics that influence one decision may not have the same effect on another; therefore, any empirical use of a representative farm must be closely tied to a specific purpose or problem.

Plaxico and Tweeten discussed the potential application of representative farms in public policy evaluation, in projections research and in studying adjustment proposals (35, p. 1463):

> Given that the criteria for selecting or identifying representative farm units is dictated by the use to be made of the units, some specific questions encountered are: (1) should variables other than physical resources be considered, (2) are the analyses to be normative or predictive in nature, (3) what is an acceptable degree of variability around a mean of a variable of interest in a defined population, (4) should the unit be representative of a present population or of some population projected to exist in a future date in time.

They observed that increasing the number of representative farms programmed would reduce within-group variance and they expressed a need for research on the problem of defining populations which can be depicted by a representative unit.

Solutions to the Aggregation Problem

Coincident with the greater awareness of the aggregation problems involved in representative farm linear programming supply models, the literature began to include the first tangible steps toward a solution to the problem.

General solutions

H. O. Hartley provided the first of these steps. He noted that there could be serious discrepancies between the area supply estimates achieved by linear programming a few representative farms and the one achieved by programming every farm in the sample. "In particular one would have misgivings about those supply functions which depend on farm economic items not accounted for by the classification . . . " (18, p. 8). For example, Hartley felt that the aggregation error may be quite serious if the original stratification of sample farms was on the basis of land and buildings while the computed supply functions were dependent on a third factor such as labor. He outlined a method by which the supply curves could be corrected for factors omitted in the classification.

Hartley's correction method avoided making a three-way classification to include the omitted factor and thereby increase the required number of representative farm linear programming solutions "beyond practical limits." He instead proposed variable factor programming on the omitted factor to establish a functional relationship between the factor and supply of the product in question. Then he developed a statistical adjustment of the programmed supply functions to account for the influence of the omitted factor. This statistical adjustment reduced the aggregation error in the original supply functions. However, it is not clear from the paper that the required variable factor programming and correction could be accomplished with less effort than the additional programming necessary if the farms were originally stratified by the third factor. This computational complexity of the correction method renders it a less than adequate solution to the aggregation problem.

Richard Day was first to state conditions which, if met, would assure exact aggregation (9). He showed by means of the duality theorem of dual linear programming that under suitable conditions a single linear programming model for the aggregate is equivalent to the summing of all the solutions of the set of individual firms. Sufficient conditions for this equivalence are proportional variation of resources and net returns among all firms in the set and identical technical coefficients for all firms. The definition of proportional variation requires that every individual firm resource (net return) vector must be a scaler multiple of every other resource (net return) vector.

Day called firms meeting these sufficient conditions "proportionally heterogeneous." He went on to evaluate the possibility that actual firms may meet these conditions, observing that "to obtain realistic aggregative results from 'adding up' the solutions of these typical firm models, it may be necessary to stratify farms into much smaller typical farm aggregates than is now being done" (9, p. 812). He recommended a detailed analysis of sample survey data to determine the extent of proportional heterogeneity among all relevant factors before deciding at what level aggregates should be formed. He indicated that scale by itself was not a useful criterion for stratifying firms. It should be used only if certain factors vary nonproportionally with respect to scale.

Finally, Day observed that the realistic objective is only to hold the aggregation error to a tolerable degree. He suggested parametric programming and ranging on the constraint and net return vectors to obtain an idea of the sensitivity of the solutions of a given model to nonproportional changes in the resources and net returns. This procedure would give an idea of the existence of aggregation bias or error in particular conditions.

Richard Day's work represents a valuable contribution to the theory of aggregation -- it provides the foundation for the additional theory developed in Chapter III of this thesis. His article was followed by an increasing awareness of the problem of aggregation error and the initiation of research projects oriented specifically toward finding a workable solution to the problem. Results from two of these projects have appeared in the literature, one by George Frick and Richard Andrews and another by Seamus Sheehy and R. H. McAlexander.

Empirically tested solutions

Frick and Andrews tested several methods of grouping farms to minimize aggregation error in linear programming supply functions (15). For comparison they computed five sets of linear programming milk supply functions for a defined universe of 51 farms:

(1) by programming every farm in the universe of 51 farms,

(2) by programming one representative farm computed as the mean of the resources of the 51 farms,

(3) by programming six different-sized representative farms classified on the basis of number of dairy stanchions

(4) by programming five representative farms derived by grouping farms according to the most limiting resource and

(5) by programming six different-sized representative farms classified on the basis of potential numbers of dairy stanchions.

The percentage errors between the last four supply functions and the first were found to be (15, p. 698):

Method	Error
(1)	(control)
(2)	+ 17.7 percent
(3)	+ 15.3 percent
(4)	+ 6.6 percent
(5)	+ 15.6 percent

In addition, they observed the supply elasticity estimates from methods (1) and (4) were somewhat similar, but that supply curves from methods (2), (3) and (5) were consistently more elastic.

Although grouping farms according to the most limiting resource gave the estimate with the least aggregation error, Frick and Andrews observed it had the disadvantages of (1) ignoring size of farm, a factor that may be important for some estimates, (2) forming classes that were hard to project to future dates and (3) becoming quite involved when handling more than one product. It did, however, provide a practical and workable solution to the problem of holding aggregation error within tolerable bounds.

Another study of aggregation error was the subject of a Ph.D. thesis by Seamus Sheehy (41) and was reported in an article by Sheehy and R. H. McAlexander (42). They compared the aggregate output estimates developed under two methods of selecting representative or "benchmark" farms.

These methods were (1) the "conventional method" where farms were classified on the basis of absolute levels of certain resources and (2) the "homogeneous restriction method" which made use of the level and productivity of the resources on sample farms. As applied, the "homogeneous restriction method" was analogous to the method Frick and Andrews used when they classified farms by the most limiting resource.

Four sets of representative farms were delineated and programmed under the "conventional method" (42, p. 689). These were:

(S_1) two farms, dairy and nondairy,

(S_2) seven farms, stratified as dairy or nondairy and by amounts of cropland,

(S_3) ten farms, stratified by dairy or nondairy, amount of cropland and herd size and

(S_4) sixteen farms, stratified by dairy or nondairy, amount of cropland, herd size and amount of labor available.

Thirteen representative farms were constructed and programmed under the "homogeneous restriction method" -- (S_5).

Estimated supply functions under the "conventional method" moved to the left as more detail was used in the stratification -- that is, the supply estimates decreased when more representative farms were used. Sheehy and McAlexander didn't estimate the percentage shifts but inspection of the results suggests that supply function S_4 represented 10 to 20 percent less production than S_1 over most of the prices programmed. Supply functions S_2 and S_3 fell between functions S_1 and S_4.

The estimated supply function using the "homogeneous restriction method" (S_5) was even further to the left than S_4; the difference between

S_1 and S_4 appeared comparable to the difference between S_4 and S_5. Sheehy and McAlexander felt that supply function S_5 could be " . . . considered virtually unbiased" after they had surveyed each of the 13 groups of individual farms represented and found them each to be relatively homogeneous within groups (42, p. 692). They concluded that the selection of representative farms on estimated restrictions of the commodity in question rather than on the absolute level of resources reduced aggregation error (42, p. 693).

In addition to this empirical work, Sheehy's thesis contains excellent sections on theory relating to aggregation error (41). His arguments are necessarily detailed and are reviewed and discussed at appropriate places in the next two chapters. Other works that could have been covered more fully in this review are handled in a similar manner, notably those of Richard Day and John Lee. Their contributions, along with Sheehy's, form an integral part of the theory developed in the next two chapters.

CHAPTER III. SUFFICIENT CONDITIONS FOR ERROR-FREE AGGREGATION

The logical sequence for the initial phases of the study of aggregation error is the formal statement of the problem followed by the development of solutions to the problem in a general theoretical framework. This chapter presents the aggregation problem algebraically and then discusses the specific conditions that are sufficient for error-free aggregation in a completely general model. Chapter IV discusses the direction of aggregation error and possible causes of this error.

The Aggregation Problem

A rigorous definition of the aggregation problem provides the point of departure for the analysis. Consider the linear programming model representing the g-th farm of a set of n farms, which is the problem of selecting a vector of production levels, X_g, such that profit is a maximum, resource limits are respected and no production levels are negative. In vector notation we solve for X_g such that

(3.1) $$\pi_g = Z_g X_g$$

is a maximum subject to

(3.2) $$B_g X_g = C_g$$

and

$$X_g \geq 0$$

where π_g = total net returns to the g-th farm,

Z_g = the 1 by m vector of activity net returns for the g-th farm,

X_g = the m by 1 vector of activity levels to be chosen by the g-th farm,

B_g = the k by m matrix of input-output coefficients for the g-th farm and

C_g = the k by 1 vector of available resources of the g-th farm.

This is standard linear programming form with the necessary slack vectors included to reach equality in the relations of Equation 3.2. Heady and Candler provide an excellent discussion of both the assumptions involved in linear programming and how it may be applied to farm management problems (21). The reader is urged to refer to their book for these details. For the purposes of this analysis, we will merely assume that a suitable solution has been found to the specification error problem (i.e., that the desired supply response for the farm is adequately determined by the variable pricing linear programming model).

The desired end result is the total production of all of the n farms in the set. If the optimum solutions are obtained for all n farms and totaled, the desired solution for the aggregate set of n farms becomes

$$\sum_{g=1}^{n} X_g .$$

This sum is <u>free</u> of aggregation error since it represents the exact sum of the solution vectors of all of the n individual farms in the set. Hence, it becomes the logical standard against which all other procedures may be judged.

As mentioned earlier, the number of farms in the set for which the estimates are desired usually makes it impractical to obtain optimum solutions for all of them, simply because of the computational costs involved. An alternative procedure involving greater abstraction becomes necessary to make the problem computationally feasible. The customary procedure is to define a representative farm within the set and determine its optimum production by linear programming. The production of the set as a whole is then estimated by the appropriate weighting of this representative farm result. If the representative farm is defined as the average farm in the set (i.e., its resources are the total resources of the set divided by n), then the appropriate weighting factor becomes n, the total number of farms. In this case, an alternative procedure yields the same estimate -- merely define the representative farm as the sum of the entire set and omit the weighting step. Including the total resources in the representative farm avoids the step of dividing the set's resources by n and the associated step of multiplying the optimal solution by n.

Thus, the alternatives often used are (1) to sum the total resources over all farms and to determine the optimum solution for the aggregate as a whole or (2) weight the results obtained for the representative farm. Since these procedures as they were described yield equivalent results, either one may be used in discussing the aggregation problem with no loss in generality. Choosing method (1) because of its fewer steps, the more abstract alternative then may be expressed in one problem of selecting a vector of aggregate area production levels, X, such that

(3.3) $$\pi = Z\ X$$

is a maximum subject to

$$(3.4) \qquad B\,X = C$$

and

$$X \geq 0.$$

The nonsubscripted symbols now represent the entire set of farms where the symbols of Equations 3.1 and 3.2 with g subscripts represented the individual farms. The dimensions of the matrices are the same in both cases. Since the individual farm resources are summed to obtain the resources of the aggregate set,

$$C = \sum_{g=1}^{n} C_g.$$

Exact aggregation may now be defined as the situation in which the levels of the various activities in the second formulation are exactly the same as those obtained by programming each farm separately and summing, that is

$$X = \sum_{g=1}^{n} X_g.$$

Conversely, aggregation error is defined as the situation in which

$$X \neq \sum_{g=1}^{n} X_g.$$

This definition of aggregation error is completely general in that it includes all m outputs or components of the vector X. The next section will discuss conditions that will eliminate aggregation error from all products. It may, of course, be possible to achieve exact aggregation for one or more of the m outputs while at the same time having aggregation error in the estimated levels of the remaining outputs. This somewhat narrower concept of exact aggregation is discussed at the end of the chapter, along with conditions that may be met to achieve it.

Exact Aggregation Over All Products

The central question now becomes, given the set of n farms and the aggregation problem specified in the previous section, what conditions are sufficient among the set of farms to achieve exact aggregation?

Proportional heterogeneity

Reference was made in Chapter II to Richard Day's work concerning sufficient conditions for exact aggregation of all products. He develops sufficient conditions for exact aggregation which he defines as the requirement of "proportional heterogeneity" (9). The conditions are that:

(3.5) $$B_1 = B_2 = . . . = B_n = B$$

(3.6) $$Z_g = \gamma_g Z$$

where γ_g is a scalar greater than zero for all g and

(3.7) $$C_g = \lambda_g C$$

where λ_g, a scalar greater than zero and less than one for all g, represents the proportion of the set's resources that the g-th farm possesses. Equation 3.5 states that all farms in the set possess identical matrices of input-output coefficients; Equation 3.6 states that the farms have only proportional variation in net return expectations; Equation 3.7 states that the farms have only proportional variation in resource or constraint vectors.

If these conditions hold, exact aggregation may be attained -- that is, one representative farm may be programmed instead of the n individual farms. The weighted solution for this representative farm will exactly equal the sum of the n individual farm solutions.

Day presents proof of the sufficiency of these conditions for eliminating aggregation error through the duality theorem of linear programming. In addition to fulfilling the previously defined requirements of exact aggregation, he notes that the condition

$$R = \frac{1}{n} \sum_{g=1}^{n} R_g$$

would also be achieved in a set of firms conforming to the conditions of the equations where R is the "average marginal net revenue productivities" of the resources in the set and the R_g are the vectors of marginal net revenue productivities of resources of the individual farms. These values represent the solution of the dual linear programming problems.

Day has an excellent discussion of the implications of the conditions of proportional heterogeneity from an operational standpoint. It would appear that the requirement of proportional heterogeneity is quite restrictive and that, at best, only very small sets of actual farms could be found that would meet it. For example, if two individual farms differ in one resource by a certain ratio, they must differ in all other resources by that identical ratio if they are to be represented by a representative farm without aggregation error. The odds against this happening appear quite high, considering the large number of resources that are relevant to the production decisions of a farm and the relatively high degree of divisibility of a large portion of these resources. Although Day does point out that such proportional variation may be found more often in the real world than we may first expect, he states that ". . . it may be necessary to stratify farms into much smaller typical farm aggregates than is now being done" (9, p. 812).

The significance of this requirement may be viewed in light of the problem of dividing a large number of individual farms into groups such that each group can be represented without error by a representative farm. In view of the extremely small coverage of each representative farm, a very large number will be required if exact aggregation is to be achieved; we may, in fact, approach programming every individual farm. Thus, the achievement of exact aggregation appears extremely difficult just because of the computational effort required.

There is hope, however. Day's condition of proportional heterogeneity is a sufficient condition for exact aggregation -- not a necessary condition. The door is left open to both the development of less binding sufficient conditions and to the definition of the minimum conditions necessary for exact aggregations.

Qualitatively homogeneous output vectors

Less binding sufficient conditions for exact aggregation may be defined using the concept of qualitatively homogeneous output vectors (abbreviated as QHOV). The first step is to define this concept both intuitively and then somewhat more rigorously. A theorem and proof of the sufficiency of these requirements follows.

An intuitive idea of the relaxed requirements is gained by considering the optimum solutions of a set of individual farms as determined by linear programming. Assume the set of farms under consideration is similar to the extent necessary for all individual optimum solutions to include identical sets of activities. Such a set of individual farms may vary in both resource and net return vectors if this variation is not great enough to cause a change in the set of optimum activities common to all farms in the group. The variation in resource vectors among farms will, of course, cause differences among farms in optimum activity levels. The important point is that the identity of the activities in the optimum solutions must be the same for all farms. Farms meeting this requirement will be defined as having qualitatively homogeneous output vectors.

To make a more rigorous specification of the new conditions, consider the optimum solution of each individual farm. The optimum solution for the g-th farm may be expressed as a column vector

$$X_g = \begin{bmatrix} x_{1g} \\ x_{2g} \\ \cdot \\ \cdot \\ \cdot \\ x_{mg} \end{bmatrix} .$$

Previously, m was defined as the number of production processes considered by the farm plus the number of slack vectors necessary to permit nonuse of all resources and k was defined as the number of resources or constraints. Observe that $m > k$ for this formulation since k is also the number of required slack vectors that are included in m to achieve equality in the restraints. For each optimum solution, X_g is made up of at most k activities that are greater than zero and at least m minus k activities that are equal to zero.[1]

For each of these farms we could express a streamlined output vector as

$$X_g' = \begin{bmatrix} x_{1g}' \\ x_{2g}' \\ \cdot \\ \cdot \\ \cdot \\ x_{kg}' \end{bmatrix}$$

by omitting the m minus k activities which are common to each farm and equal to zero. We note now that the X_g' (streamlined output vectors) for farms having qualitatively homogeneous output vectors will all consist of the same k basic activities. All such farms will have the same resources limiting, the same resources in disposal and the same real processes in their final solution vectors.

<u>Theorem I</u> Sufficient conditions for exact aggregation are (1) that all farms have identical coefficient matrices, that is, that $B = B_g$ for all g and (2) that all farms have qualitatively homogeneous output vectors (QHOV).

---

[1]These k activities are often called the basic variables in the literature, while the remaining activities are called nonbasic variables. The theorem generally developed is that an optimum solution involves at most k unknowns at nonzero values (where k equals the number of equations). For example, see (10, Theorem 2, p. 75).

<u>Proof</u> For farms meeting conditions of the theorem, the original linear programming problem may be reduced to the more trivial problem of solving a set of k equations in k unknowns

(3.8) $$B' X_g' = C_g$$

where $B' = B_g'$ is the k by k part of the coefficient matrix corresponding to the k activities in X_g'. This reduced equation is then equivalent to the original constraint set, Equation 3.2, with the unused activities (columns) of the coefficient matrix omitted and the zero elements of X_g omitted. This is no more than saying that if the identity of the final basic activities is known in advance, the linear programming problem may be solved simply as a set of simultaneous equations.

Similarly, the solution to the aggregate farm may be determined by the relation

(3.9) $$B' X' = C$$

which is developed in a similar fashion to the reduced equations for the n individual farms.

Summing Equation 3.8 over all n farms gives

(3.10) $$B' \sum_{g=1}^{n} X_g' = \sum_{g=1}^{n} C_g.$$

Since $$\sum_{g=1}^{n} C_g = C$$

by definition, it is obvious from Equations 3.9 and 3.10 that

$$X' = \sum_{g=1}^{n} X_g'.$$

All that remains is to include the m minus k zero level elements to both vectors to complete the proof that

$$X = \sum_{g=1}^{n} X_g.$$

The conditions of the theorem are hence sufficient conditions for exact aggregation.[1]

The conditions of the theorem are general in respect to the price or revenue vectors used; hence, the theorem covers variable-price programming. This is because the consideration of different prices merely has the effect of further restricting the groups of farms that meet the requirement of QHOV.[2] To have exact aggregation under varying sets of prices, all farms in the group must meet the conditions of the theorem for every set of prices considered. In other words, the farms must all have solutions meeting the requirement of QHOV for the first set of prices, have a set of possibly different solutions but again meeting the requirement of QHOV for the second set of prices and so on for all price combinations considered.

The conditions of Theorem I are substantially less binding from an operational standpoint than are the original ones developed by Day. Now some range of different resource situations and net return expectations can be covered by a representative farm without aggregation error.

---

[1]Theorem I, its proof and excerpts of this discussion appear in Agricultural Economics Research (32).

[2]Assuming of course that the group of farms all have identical input-output coefficient matrices.

Moreover, there is no restriction on the type of variation that may occur in these vectors among farms, as long as the amount of variation is within the limits that allow all individual farms in the group to have solutions made up of the same activities. The only restrictions are that to be represented by a representative farm without aggregation error, all of the individual farms in the group must have (1) identical input-output matrices and (2) QHOV.

On the negative side, the conditions of Theorem I are defined as a requirement of the solutions to the individual farms rather than as a requirement of the farms themselves. They thus provide a less than ideal solution to the problem of delineating representative farms to eliminate aggregation error. A translation of these conditions into requirements on the data of the individual farms is needed to make the process truly operational. This translation may be approached as follows.

An alternative statement

Interpretation of Theorem I into restraints on the coefficients rather than on the solutions of the individual farms requires a close look at the dual linear programming solutions of the farms.[1] Define R and R_g (for g = 1 . . . n) as m by 1 vectors representing the dual solutions for the linear programming problems of the aggregate farm and the n individual farms, respectively.

[1]A formulation and discussion of the dual linear programming problem may be found in Heady and Candler (21, pp. 90-107).

Each of the m by 1 dual solution vectors has a one to one correspondence with the m activities (including slack or disposal activities) in the respective primal linear programming problems. The values in the dual solution vectors represent the marginal decrease in the functional value, π, that would result from a marginal increase of one unit of the activity in question. As a result, the values are all nonnegative for an optimum solution and have the following economic interpretation. For disposal activities, they represent the loss in net returns that would be brought about from one more unit of the activity (i.e., one less unit of the restrictive resource being used) -- hence, they represent the marginal productivity of these resources in the optimal primal solutions. These values are often called the shadow prices. For real (producing) activities, the values in the dual solution represent the loss in net returns that would be caused by including one unit of a nonprofitable activity in the optimal primal solution -- which could also be interpreted as the change in that activity's Z value which would be required to make it profitable enough to be included in the optimal primal solution.

At the time Theorem I was developed, it was observed that a parallel argument could be developed for aggregation of the dual solutions over the same set of n farms (32, p. 56). Under the conditions of Theorem I,

$$C = \sum_{g=1}^{n} C_g \qquad \text{lead to} \qquad X = \sum_{g=1}^{n} X_g.$$

For the dual solutions, a similar argument could be developed to show that if

$$Z = \frac{1}{n} \sum_{g=1}^{n} Z_g, \text{ then } R = \frac{1}{n} \sum_{g=1}^{n} R_g.$$

Restated, if the net return vector of the aggregate farm is defined as the average of the net return vectors of the individual farms and if the conditions of Theorem I are met, then the dual solution of the aggregate farm will be equal to the average of the dual solutions to all of the individual farms.

In discussing Theorem I, John Lee makes an interesting extension of this argument (26, p. 59). He observes that the statement for aggregation of the dual solutions could be modified slightly to show that if $Z_g = Z$ for all g, then $R_g = R$ for all g. If all individual farms in a set meeting the conditions of Theorem I have identical net return vectors, then they will all have identical dual solution vectors. The dual solutions will be the same for all farms and constant over the range of resource ratios represented by the individual farms. Lee makes the observation that, utilizing this relationship, the observed ranges of the resource ratios can be used as a criteria for grouping individual farms on the basis of observable characteristics. He then states a new aggregation theorem, hereafter referred to as Theorem II, and explains it as follows (26, p. 59):

[Theorem II]

> "Sufficient conditions for exact aggregation are (1) that all farms have identical coefficient matrices, (2) that all farms have the same net returns expectations and (3) that the range of resource ratios be such that the dual solution vector is the same for all farms."

> This is simply the dual counterpart of Miller's theorem. It would delineate sets of farms identical to those delineated by the original theorem. However, it may be more useful since it lends itself to interpretation in terms of observable characteristics. The link between theorem and application is the empirical task of determining the exact ranges of resource ratios over which the marginal revenue product is constant.

Lee considered the proof of Theorem II to be quite obvious and omitted it from his discussion. The proof may be easily constructed in the following manner. First consider the dual solution vectors $R = R_g$. They correspond to the same m activities that are in the primal solution vectors X and X_g; when an activity is in the optimal primal basis, its value in R is zero; when an activity is not in the primal basis, then its value is not equal to zero. From this correspondence we see that $\bar{R} = \bar{R}_g$ for all g, in fact, implies QHOV. Then by Theorem I, QHOV and identical coefficient matrices imply exact aggregation.

Inspection of this proof reveals that it in no way requires condition 2 of Theorem II that all farms have identical net return expectations. The remaining two conditions are actually sufficient conditions for exact aggregation and the theorem is true without condition 2. As such, it should be clarified that the groups of farms delineated by Theorem II are identical to those delineated by Theorem I only when all individual farms have identical net return vectors. A given group of individual farms may meet the conditions of Theorem I and still have different net return vectors. Such a group would have to be further subdivided in order to meet the requirements of Theorem II.

Although this requirement of identical net return vectors may in some cases force a larger number of groups, there is much to be gained from it. Lee included it in Theorem II because it makes the conditions of the theorem observable in the individual farm data. This eliminates the need for determination of the optimum solution before individual farms can be accurately grouped. Using Theorem II, farms can be classified and grouped solely on the basis of their resource vectors and coefficient matrices.

Lee showed how this may be accomplished by using Figure 1 and the following discussion (26, pp. 59-60).

> The potential of this approach may be demonstrated graphically. Suppose there exists a group of farms each possessing some combination of two resources C and L, each viewing the same three production processes, A_1, A_2, and A_3, with technical coefficients common to all farms, and each having identical net returns expectations. The situation is depicted in figure 1. All possible resource ratios are depicted by points on the horizontal bar C_1 C_1' derived by adding varying amounts of resource L to a fixed amount (C_1) of resource C. The net returns expectations for an initial set of product prices (or input prices) create a field of iso-revenue curves exemplified by the solid iso-net revenue curves shown.
>
> Given the situation described above and portrayed in figure 1, the output and net revenue for farms with C_1 of resource C and none of resource L will be zero. As the level of resource L increases in small increments from L_0 to L_1, L remains the limiting resource and net revenue increases in proportion to increases in L. In other words, as resource L is increased from L_0 to L_1, net revenue is maximized by moving up the iso-revenue field along the A_1 activity vector (since the A_1 vector represents the most efficient utilization of resources as long as resource L is limiting).
>
> Since net revenue changes in proportion to changes in the limiting resource, L, between L_0 and L_1, the marginal value product (shadow price) of L is constant over the same range. . . .

At this point Lee observed that all farms having resources C and L in ratios between C_1/L_0 and C_1/L_1 could be aggregated without error,

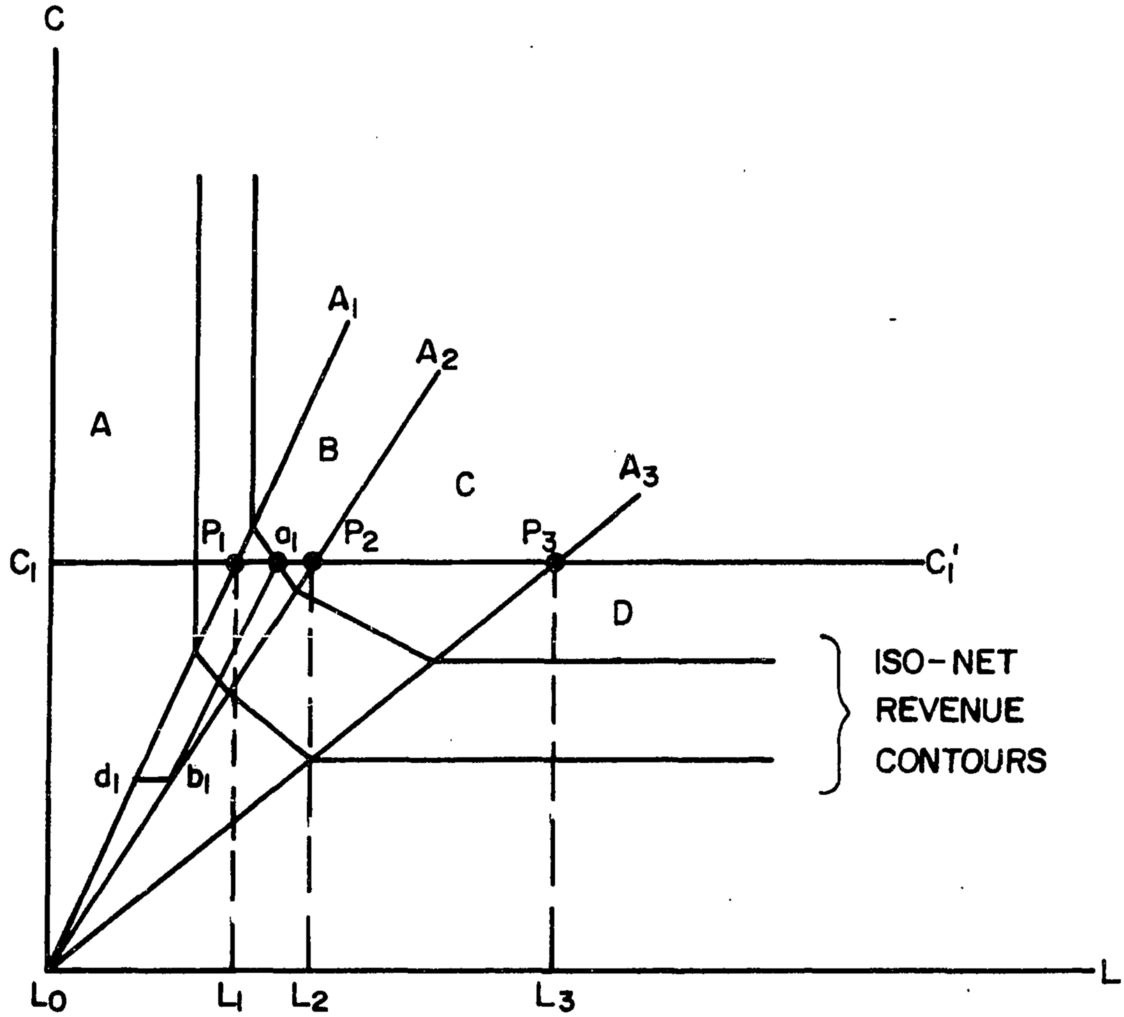


Figure 1 (26, p. 60). Diagram of a linear programming model of two resources and three activities for farms with varying labor-capital ratios since all of the requirements of Theorem II were met. This group of farms has identical input-output matrices and net return vectors by assumption. In addition, they all have identical dual solutions as reasoned using Figure 1.

Lee then went on to discuss the situation for farms with other resource ratios (26, p. 60).

> With resource combination C_1L_1 (denoted by point P_1), both resources are exactly used by activity vector A_1. With further increases in resource L (beyond L_1) both resources C and L are limiting. However, the full amount of both resources can be utilized and net revenue maximized by combinations of activities A_1 and A_2 (for example, L_0 b_1 of A_2 and b_1 a_1 = d_1 P_1 of A_1 in figure 1). The locus of resource combinations, P_1 P_2, is also the path of net revenue expansion as resource L is increased. This expansion path intersects the iso-net revenue field at constant angles (i.e., as L increases, the net revenue from A_2 substitutes for net revenue from A_1 at constant rates). Thus, between L_1 and L_2, the marginal revenue product of L is constant and the conditions of the dual to Miller's theorem [Theorem I] are again met. Note that the MVP of L between L_1 and L_2, while constant, is less than the constant MVP of L between L_0 and L_1. The reason is that as L is increased it becomes less scarce relative to resource C. This is reflected in the flatter slope of the iso-revenue curve. Obviously, farms with resource ratios between C_1/L_1 and C_1/L_2 can be aggregated without bias.
>
> As resource L is increased from L_2 to L_3, its MVP is again constant though lower than previously. Farms with resource ratios between C_1/L_2 and C_1/L_3 meet the conditions for exact aggregation. Beyond L_3 amounts of resource L, C becomes the only limiting resource; A_3 is the only activity in the solution and the MVP of L is constant at zero. Thus, all farms with resource ratios of C_1/L_3 or less can be aggregated without bias.
>
> With resource C fixed at C_1, the line L_0 P_1 C_1' represents the maximum efficiency net revenue expansion path as L is increased from L_0 to infinity. The angle at which this path cuts the field of iso-net revenue curves determines the marginal revenue product (shadow price) of L. . . .

This completes the discussion of all possible resource ratios that may occur in the set of farms represented in Figure 1. It is now clear that the maximum number of groups of farms required to eliminate aggregation error in this case is four. For the net revenue vectors reflected in the iso-revenue contours, only four dual solutions are possible. The activity vectors A_1, A_2 and A_3 themselves, along with the axis, form the dividing lines which separate the different dual solutions. All combinations of C and L falling between two such vectors will have the same dual solutions and may be aggregated without error.

An extension may be made to all possible net revenue vectors (as would be encountered in the generation of supply functions by variable pricing) as long as all farms in the set have identical net revenue vectors at each point aggregated. If these three activities are the only ones available to the farms, the four groups of farms A, B, C and D are the maximum number of groups required for exact aggregation. There is, of course, the possibility that with a single set of net revenue vectors, the number of groups may be reduced, since two or more groups may have the same dual solutions. However, it is worthy of repeating that no more than four groups would ever be required. No matter how many individual farms were in the original group, their behavior could be estimated without error by using only four representative farms.

It is apparent from Figure 1 that the sufficient conditions for exact aggregation expressed by Theorem I are much less restrictive than the conditions developed by Richard Day. In order to meet Day's requirement of proportional heterogeneity, a separate representative farm would be needed for every different point on the line C_1 C_1'. Each different point on this line contains a new resource ratio and a group of any two would not meet Day's requirements. It is possible, in fact, to visualize on Figure 1 the groups of farms which would meet Day's requirement. In order to achieve proportional heterogeneity, a group of farms in Figure 1 would all have to be situated on a straight line extending out from L_0. A group of farms with the C/L ratio expressed by a line like A_3 would be an example. In view of the extremely large number of such lines that may be required, these conditions are considerably more restrictive than the conditions of Theorem I.

The value of Theorem II and the Figure 1 analysis is now apparent. The boundaries of the groups of farms having different shadow prices are formed by the technical coefficients themselves. These input-output coefficients contain the information needed to determine what range of resource ratios may be included in a group of farms and still allow them to be aggregated without error. It is not necessary to solve the individual linear programs in order to determine groups of farms which meet the requirements of Theorem I. The practical problem left is to determine the exact range of resource ratios over which the shadow prices are constant -- Figure 1 suggests that this may be accomplished by looking at the input-output coefficients themselves. The approach involves comparison of the ratios in which resources are required by the different activities with the ratios in which the resources are available to the individual farms.

Number of representative farms required

In his later work, John Lee explores a method for extending the previous analysis to the general case of any number of restrictions and activities (27). The reader is urged to refer to his analysis directly for details; it may, however, be summarized as follows.

The problem is one of grouping a set of farms having identical coefficient matrices and identical net return vectors into the number of groups that will assure elimination of aggregation error. The boundaries of the various groups are ascertained from information contained in the coefficient (B) matrix common to all farms.

Lee's first step is to divide the first row of the B matrix by the second. This gives all of the different ratios in which all the activities use resources c_1 and c_2. When arrayed from smallest to largest, these ratios become the critical boundaries of farms classified by the ratio of resource 1 to resource 2. For an example with three activities, the array of B ratios may be:

$$\frac{b_{11}}{b_{21}} < \frac{b_{13}}{b_{23}} < \frac{b_{12}}{b_{22}}$$

This array would delineate four groups of farms based on the ratio of these first two resources, c_1 and c_2:

Group	Resource ratios included
1	$\frac{c_1}{c_2} \leq \frac{b_{11}}{b_{21}}$
2	$\frac{b_{11}}{b_{21}} < \frac{c_1}{c_2} \leq \frac{b_{13}}{b_{23}}$

3 $$\frac{b_{13}}{b_{23}} < \frac{c_1}{c_2} \leq \frac{b_{12}}{b_{22}}$$

4 $$\frac{b_{12}}{b_{22}} < \frac{c_1}{c_2}$$

This classification would be similar to the Figure 1 example, except for the fact that in Figure 1 there were only two resources and the process stopped at this point.

With more than two resources, the above process is repeated for every possible resource ratio, subdividing the previous groups for every additional ratio considered. Thus, with three resources and three activities, the farms are divided into four groups on the c_1/c_2 ratios; each of these groups is subdivided into four more groups on the c_1/c_3 ratios; finally each of these 16 groups is subdivided into four more on the basis of the c_2/c_3 ratios. Thus, if the B matrix is 3 by 3, 64 groups of farms would be the maximum ever required to achieve exact aggregation.

In general, the maximum number of groups of farms required for exact aggregation with a B matrix of k rows and p real activities is

$$N = (p + 1)^{{}_kC_2} \tag{3.11}$$

where

$${}_kC_2 = \frac{k!}{2!(k-2)!} = \frac{k(k-1)}{2} .$$

This is the maximum number of groups that would be required for the case in which all of the elements of B were nonzero and all of the critical B ratios were different. Zero B elements mean that the activity does not require that particular resource and the ratios involving zero may be overlooked in the process. Defining d as the probability of a nonzero item in a particular location of the B matrix, then d^2 becomes the probability of a coefficient ratio composed of two nonzero elements. This allows correction of Equation 3.11 for d, which is also the density of the B matrix, resulting in the equation

$$(3.12) \qquad N' = (pd^2 + 1)^{\frac{k(k-1)}{2}}$$

where N' is the expected number of groups required.[1]

It is surprising to see how fast N' grows for increasing values of k, p and d. (See Table 1.) The number of groups required to eliminate

Table 1. Expected number of representative farms required to eliminate aggregation error for different sizes and densities of matrices

Rows k	Columns p	Density d	Approximate number required N'
3	3	1	64
10	20	.25	3,325
30	60	.25	2.68×10^{321}

aggregation error would, of course, never exceed the number of individual farms in the set to be analyzed. The high N' values for larger coefficient matrices would indicate that to assure exact aggregation, nearly as many representative farms may be needed as there are individual farms.

[1]The correction for density assumes that the distribution of zero values in each row of the B matrix is independent of the distribution of zero values in the other rows. If this assumption is not met, serious errors could arise from using this formula for low density matrices.

This observation backs up John Lee's contention that the number of representative farms required by Theorem I would reduce to Day's groups only when the number of activities reaches infinity (26, p. 61). Actually, the number of resources (rows) is an even more important factor, increasing N' as an exponential. As a result, both criteria become unmanageable for eliminating aggregation error when large B matrices are involved. For small B matrices the Theorem I requirements offer considerable improvement over Day's conditions. This point has already been made in discussing Figure 1.

The extremely large numbers of groups that may be required to eliminate aggregation error based on the conditions of Theorem I make some other course of action necessary. Three possibilities come into mind: (1) Identify still less binding conditions sufficient for exact aggregation, (2) develop a general measurement of aggregation error and a formal mathematical procedure for minimizing rather than eliminating it, given the number of groups of farms that can be worked with or (3) make an intuitive interpretation of the implication of Theorem I to the stratification procedure, with the goal again one of holding aggregation error to tolerable bounds rather than elimination of it.

John Lee discusses the second of these three possibilities (27). The method he proposes involves (1) identification of a measure of aggregation error in the multiproduct case, (2) grouping individual farms into the N' groups which are sufficient to eliminate aggregation error and (3) successive combinations of these groups so that aggregation error is minimized at each combination with the process stopping with one large group. With the information derived, the researcher can choose the number of different groups he wants to program based on the trade-off between the magnitude of the aggregation error and computing costs.

This formal procedure for minimizing aggregation error is the most ambitious of the three possibilities. There are, however, several obstacles to progress in this direction. The first is encountered in developing a choice criteria based on measuring aggregation error in a multiproduct situation. It is not clear that a general measure of aggregation error can be developed which is also specific enough to be relevant to a given situation. Lee's idea of "minimizing maximum aggregation error" is a debatable solution to this problem. The second obstacle is the size of the computational problem involved in starting with N' groups and working down through successive combinations. Considering the possible magnitude of N', there is even a question of whether a computer algorithm can be developed to economically handle the grouping problem from this direction. Such obstacles must be overcome before this formal procedure for minimizing aggregation error can be made operational.

An intuitive interpretation of the implication of Theorem I to the stratification procedure seems a better alternative. This is especially so since, as is brought out in Chapter IV, identification of significantly less binding sufficient conditions for exact aggregation is an unlikely possibility. On an intuitive level, Theorem I provides some guidelines to the problem of grouping farms to minimize aggregation error. It suggests that individual farms should be grouped into homogeneous groups on the basis of their coefficient matrices and then subdivided so that each cell would have the same optimum set of production activities. This idea will be developed more completely in Chapters VII and VIII after additional theoretical concepts of aggregation error are developed and the empirical results are discussed.

Exact Aggregation of One Product

Up to this point, the problem of exact aggregation has been approached with the idea of achieving it simultaneously for all activity levels or product estimates in the model. The problem may also be approached from the less general standpoint of considering only one product. This approach simplifies greatly the logic of achieving exact aggregation and tends to reduce the number of representative farms required. Such one-product supply estimates are often quite relevant in agriculture since many segments specialize in the production of one commodity.

Such one-product aggregation was the primary concern of the studies of Frick and Andrews and Sheehy and McAlexander reviewed in Chapter II of this theses (15, 42). These studies were both concerned with estimation of milk supply functions. In the areas in which these studies were concerned, dairying was the principal type of farming and milk supply functions were of major concern. This specialization made consideration of aggregation from the one-product standpoint logical and meaningful.

Although it is not explicitly stated, Sheehy's thesis deals primarily with the problem of minimizing aggregation error in the one-product sense. He builds most of his argument and procedure on the following statement (41, p. 22):

> If, in a group of farms, differences in the output of a commodity from farm to farm at a given price for that commodity are proportional to at least one resource and the commodity-resource relationship is not affected by constructing a benchmark farm from the individual farms, then the commodity output of the benchmark farm constructed by averaging the group farm resources can be expanded to an unbiased estimate of the group using the number of farms in the group as the expansion factor.

Knowledge of this principle provides a method which may be used in stratifying farms where estimates are desired for only one product. The idea is to identify a resource or set of resources on each individual farm with a level proportional to the product in question and to stratify farms based on this resource. For example, if there is a group of farms all having cropland proportional to corn production, then the total corn production may be estimated without aggregation error by use of a representative farm for the group.

Sheehy defines the resource which is proportional to the product of interest as the absolute restriction for that product. He shows that grouping farms by absolute restrictions eliminates aggregation error. It is interesting to note that such a grouping is completely consistent with the Theorem II requirements of identical marginal productivities on all farms for the resource in question. Sheehy revealed his awareness of this in a graph similar to Figure 2 (41, pp. 28-29). It shows the line OAB as the total product curve as labor is increased on farms with a fixed amount of some other resource, for example, capital. In the area to the left of A (type A farms) labor is an absolute restriction for the product. In this area the level of

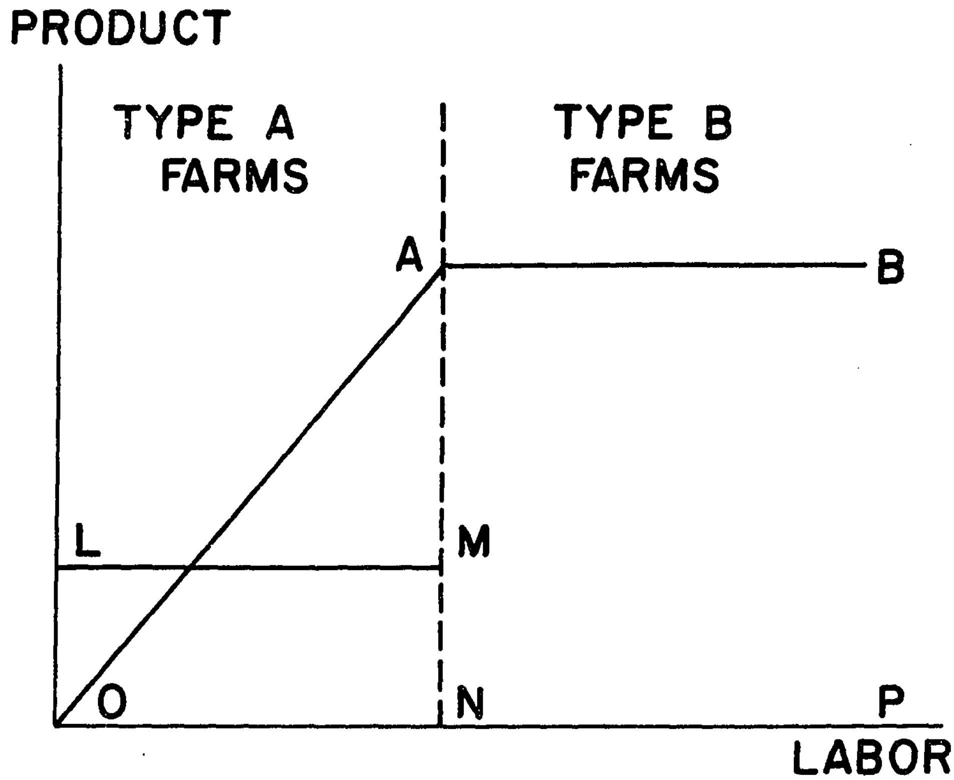


Figure 2. Total and marginal product curves for labor on farms with different labor-capital ratios.

product is proportional to the amount of labor, the total product curve is a straight line and the marginal productivity of labor is constant and shown by line LM. As labor expands, capital becomes the absolute restriction (type B farms) and the marginal productivity of labor is zero for these farms as shown by line NP. Type A farms would have a marginal productivity of OL for labor and a marginal productivity of zero for capital; type B farms would have a marginal productivity of zero for labor and some positive marginal productivity for capital. Thus, all farms within each of these groups would have the same marginal productivities or dual solutions. This would indicate that Sheehy's principle is analogous to a one-product form of Theorem II.

Exact aggregation under Sheehy's principle would appear to be promising indeed. However, limitations arise. First, the extension to two or more products enormously complicates the procedure. The problem of identifying the proportional resources to go with each of several products appears extremely complex. A second even more important problem is posed by the fact that for many optimum solutions no proportional relationship exists between a given product and any of the resources. Generally a proportional relationship would not exist unless the product in question is the sole user of one resource. This difficulty forces some compromise in the procedure and the solutions to such cases are not very satisfactory. Finally, the problem of identification of the proportional resources when they do exist becomes quite complicated in a practical situation even for one-product estimates. Sheehy spends a good deal of time in his thesis discussing a procedure designed to accomplish this (41, Chapter IV). His thesis should be referred to for the details of the procedure. The results of his empirical work have already been reviewed in the Sheehy and McAlexander article (see Chapter II, this thesis).

CHAPTER IV. ADDITIONAL THEORY ON AGGREGATION ERROR

This chapter explores some additional aspects of the aggregation error problem. The first step is an evaluation of the implication of the popular term "aggregation bias" to the direction of aggregation error. Second, theory is developed concerning the direction of aggregation error. Third, situations leading to aggregation error are specified. These situations include the analysis of the effect of variation in the coefficient matrices of the individual farms, a point neglected previously.

Direction of Aggregation Error

Aggregation error has been defined as the difference between the area supply estimate developed as the sum of the linear programming solutions for each individual farm in the population and the area supply estimated by a small number of representative farms. This error has been more popularly referred to in economic literature as aggregation bias. The term aggregation bias was used initially and only recently has a trend developed toward use of the term aggregation error (45, p. 478). Generally, the definition of the terms has been the same.

It would appear that the use of the term "bias" implicitly refers to systematic direction in the errors arising from aggregation. In statistical terminology, an estimator is called biased if it tends to under- or overestimate the desired parameter -- equivalent to saying that the estimator's expected value is not equal to the parameter to be estimated. Thus, the term aggregation bias carries the implicit connotation that the aggregate supply function estimated by a representative farm is expected to under- or overestimate the actual supply function.

There is evidence to suggest that the representative farm would overestimate the actual supply estimate. Such overestimation is implied by simple hypothetical examples and, to the extent that the implication has existed, no effort has been made toward expelling it. In retrospect, use of the term bias in describing errors in representative farm linear programming supply estimates may have contributed greatly to questions concerning the validity of the procedure.

A one-product example

An example of the reasoning relating to aggregation bias is found in the article by Sheehy and McAlexander (42, pp. 684-685). They demonstrated that for a situation in which two individual farms are producing one product with two resources, a representative farm overestimated the aggregate supply.

Sheehy and McAlexander began with the definition of an absolute resource restriction. "An absolute restriction is defined as one that limits absolutely the output of a particular commodity as the price of that commodity is raised indefinitely. That is, when the restriction is operative, all of the resource supply will be allocated to that commodity" (42, p. 684). Then they illustrated aggregation bias by assuming A was a farm with labor as an absolute restriction and B a farm with capital as an absolute restriction on the output of one specific product. They denoted the output of the product of farm A when the labor was absolutely restrictive as L_A (the output of the product on farm A when all of the labor is devoted to the product). They let C_A represent the output of the same product if capital rather than labor were absolutely restrictive. On farm B, they let C_B represent the restrictive output (capital assumed absolutely restrictive) and L_B, the potential output if labor were restrictive rather than capital. It then follows from the definition of an absolute restriction that:

$$(4.1) \qquad L_A < C_A \qquad \text{and} \qquad C_B < L_B$$

The combined output of the two farms separately programmed would be $(L_A + C_B)$. If, however, the resources of the two farms were first combined and the representative farm programmed, the output would be either $(L_A + L_B)$ or $(C_A + C_B)$, depending on whether labor or capital was restrictive on the combined farm. It follows from 4.1 that $(L_A + C_B) < (L_A + L_B)$ and that $(L_A + C_B) < (C_A + C_B)$. Therefore, Sheehy and McAlexander concluded that the output of the farms programmed separately would be less than the output when the resources of the two farms are combined (42, p. 685).

The reader is left to draw his conclusions about generalizing this result. It is quite possible that the writers didn't intend to imply that similar overestimation would occur in the multiproduct, multiresource case. However, this conclusion would appear reasonable from the example. In his thesis, Sheehy states, "An average benchmark farm representing a group of farms with different absolute restrictions gives an upward biased estimate of aggregate supply" (41, p. 25). At another point, he recognizes the possibility of negative aggregation error but does nothing toward placing it in the proper perspective (41, p. 33).

As a result, his work would at least contribute to the idea of a bias in terms of overestimation.

Relevant theoretical concepts

It is possible to treat some aspects of the direction of aggregation error in a more rigorous manner. Consider the notation developed in Chapter III. The first relationship that can be established is the direction of error in the estimate of the total maximum net returns (π) compared with the actual maximum net returns

$$\sum_{g=1}^{n} \pi_g.$$

If the assumptions are made that $Z_g = Z$ and $B_g = B$ for all $g = 1 \ldots n$, then the following theorem is true.

Theorem III The representative (aggregate) farm estimate of total maximum net returns for the set of farms is at least as great as the value found by summation of the individual farms; that is

$$\pi \geq \sum_{g=1}^{n} \pi_g.$$

Proof Observe that for all individual farms, the optimal feasible solutions are such that

$$B\, X_g = C_g.$$

Summing over n farms

$$B \sum_{g=1}^{n} X_g = \sum_{g=1}^{n} C_g = C.$$

Therefore, $\sum_{g=1}^{n} X_g$ is feasible for the aggregate problem but it is not necessarily optimum. Restated, the vector of activity levels representing the summation of the optimum solutions of the individual farms would identify a point within the constraint set of the aggregate farm, but there is nothing to assure that it would be optimal. Therefore,

$$\pi \geq Z \sum_{g=1}^{n} X_g = \sum_{g=1}^{n} Z X_g = \sum_{g=1}^{n} \pi_g .$$

This proves that under the stated conditions, the total maximum net returns estimated by the representative (aggregate) farm is at least as great as the summation of the values for the individual farms. Such an estimate is biased in the statistical sense -- in this case the expected value of the estimate is greater than the parameter to be estimated.

Positive aggregation error may be defined as the case in which the representative farm estimate is greater than the sum of the individual farms. If all activity levels exhibit positive aggregation error, then

$$\sum_{g=1}^{n} X_g < X.$$

With the aid of this definition, a consequence of Theorem III may be stated in reference to the relationship between positive and negative error in activities. First make the additional assumption that $Z \geq 0$ (the vector of activity net returns serving all farms is nonnegative).

<u>Corollary I</u> If one activity has a negative aggregation error, then some other activity must have a positive error.

<u>Proof</u> Denial of this is equivalent to asserting that for the optimum solutions,

$$\sum_{g=1}^{n} X_g \text{ dominates } X.$$

This is equivalent to saying that at least one activity level in the optimum $\sum_{g=1}^{n} X_g$ is higher than it is in X while no activity is strictly less. Therefore,

$$Z \sum_{g=1}^{n} X_g > Z\,X \qquad \text{and} \qquad \sum_{g=1}^{n} \pi_g > \pi$$

contradicting the primary theorem.

Thus, under the condition of nonnegative net return vectors, no activity can have negative aggregation error without at least one activity having positive error. The number of activities with negative errors can exceed the number of activities with positive errors. The only requirement is that there must be at least one with a positive error if there is one with a negative error.

The effect of the assumption that $Z \geq 0$ can now be seen. Existence of activities with negative net returns would allow activity estimates with negative error to exist without being offset by any other activity estimates with positive error. Thus, in the general case, these results are not suggestive of any consistent direction or bias in aggregation error in the individual product estimates. Both positive and negative errors may exist and, when the Z vectors are unrestricted as to sign, no relationship is indicated between the errors in either direction.

A two-product example

It is desirable at this point to consider an extension of the Sheehy and McAlexander example to two products. First, observe what happens in the one-product model discussed by Sheehy and McAlexander when farm A is fixed so that $L_A < C_A$, capital on farm B is fixed at C_B and the amount of labor available on farm B (L_B) is parameterized. As long as $0 \leq L_B \leq C_B$, there is no aggregation error and when $L_B > C_B$, there is positive aggregation error. In general, there is no error if the same resource is absolutely restrictive on both farms and positive error when different resources are absolutely restrictive. In summary, for the one-product case the following combinations are possible:

Absolutely restrictive resource		
Farm A	Farm B	Error
L	L	None
L	C	Positive
C	L	Positive
C	C	None

In this example two cases exhibit positive aggregation error and two cases exhibit no error from aggregating. The situation that Sheehy and McAlexander have chosen for their example exhibited positive aggregation error. The important point is that there also exists an equal possibility of no aggregation error. This fact may have escaped readers of the Sheehy and McAlexander article.

The reasoning of the Sheehy and McAlexander study may be expanded to two products. Assume that individual farms A and B both have the alternatives of producing corn and soybeans with the inputs labor and capital. The aggregate farm, farm D, would then have the sum of the resources of farms A and B. Assume that all of the usual assumptions of linear programming are met. Then assume the programming problems of the two individual farms (i = A, B) and the aggregate farm (i = D) are as follows:

Maximize $\quad Z\, X_i$

$$\begin{bmatrix} 3 & 4 \end{bmatrix} \begin{bmatrix} x_{1i} \\ x_{2i} \end{bmatrix}$$

subject to $\; B\, X_i \leq C_i$

$$\begin{bmatrix} 1/2 & 1 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} x_{1i} \\ x_{2i} \end{bmatrix} \leq C_i$$

and $\qquad X_i \geq 0$

where x_{1i} is the product corn,

x_{2i} is the product soybeans,

c_{1i} is the labor resource and

c_{2i} is the capital resource.

Assume that

$$C_A = \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \quad C_B = \begin{bmatrix} 2 \\ 6 \end{bmatrix} \text{ and } C_D = \begin{bmatrix} 4 \\ 9 \end{bmatrix},$$

such that $C_A + C_B = C_D$. Aggregation error would exist whenever $X_A + X_B \neq X_D$. Notice that these farms all have the same vectors of net returns (Z) and that both the elements of these vectors are positive. Thus, they meet the requirements of Theorem III and its corollary.

Figure 3 shows the graphical presentation and solution of these three problems.[1] Farm A maximizes returns by producing no corn and 1 1/2 units of soybeans, farm B by producing 2 units of corn and 1 unit of soybeans and farm D by producing 1 unit of corn and 3 1/2 units of soybeans. Observe that

$$\begin{bmatrix} 0 \\ 1\frac{1}{2} \end{bmatrix} + \begin{bmatrix} 2 \\ 1 \end{bmatrix} \neq \begin{bmatrix} 1 \\ 3\frac{1}{2} \end{bmatrix}$$

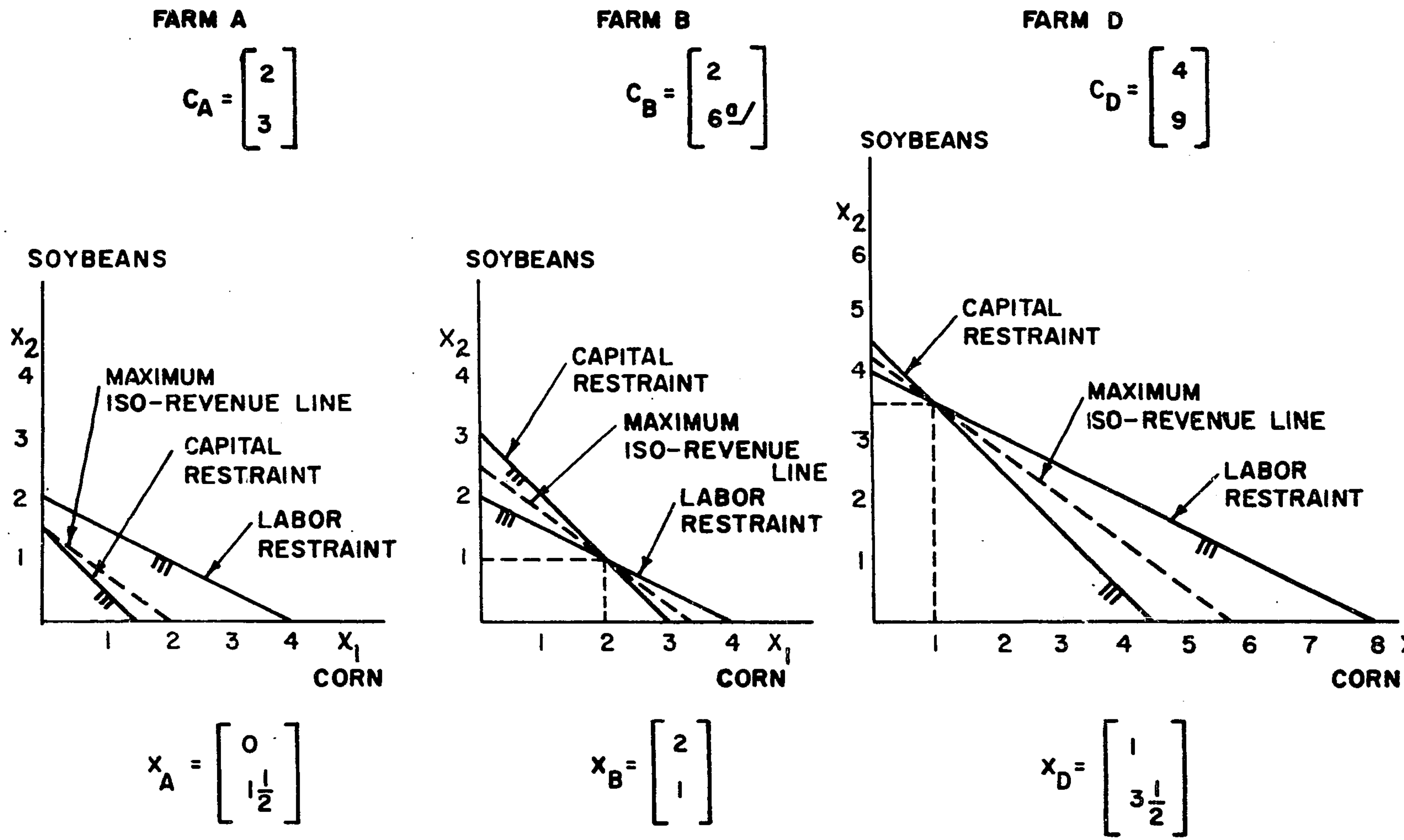
so that error results in the aggregate farm's estimate of both corn and soybeans. For corn, the aggregate farm <u>under</u>estimates the production (negative error), and for soybeans it <u>over</u>estimates the production of farms A and B programmed separately and summed (positive error).

The definition of absolute resource restrictions is relevant for this two-product case, and is quite obvious from the graphical presentation in Figure 3. For farm A, capital is an absolute restriction for both corn and soybeans; for farms B and D, capital is an absolute restriction for corn and labor is an absolute restriction for soybeans. Such an absolute restriction crosses the axis representing the product in question closest to the origin.

The amounts of resources available can be parameterized in the two-product case as was done previously in the one-product case. For example,

---

[1]Graphical solutions to linear programming problems of the type presented in Figure 3 have been discussed by Heady and Candler (21, Chapter 2). Readers not familiar with the technique may want to review this chapter since the technique provides the basis for much of the following discussion.



a/ THE QUANTITY THAT IS PARAMETERIZED IN THE DISCUSSION

Figure 3. Graphical solutions to linear programming problems for three farms

consider varying the amount of capital available on farm B from zero upwards while making appropriate changes in capital available on farm D so that its resources are always the sum of the other two farms. Table 2 presents the results for different possible amounts of capital programmed on farm B.[1] For each of these amounts, the table shows what resources are absolutely restrictive for what products and the direction of the aggregation error.

For two products and two inputs and where farm A is held constant as in Figure 3 and capital on farm B (c_{2B}) is varied, there are three possible combinations of absolute resource restrictions on farms A and B. First, both farms may have both products controlled by the same absolute resource restrictions. This is shown on Table 2 for $0 \leq c_{2B} \leq 4$ where capital is the absolute restriction on both farms for both products. A second possibility is for both farms to have the same absolute resource restriction for one product but have different absolute restrictions for the second product. This is the case in Table 2 for $4 < c_{2B} < 8$. Here capital is absolutely restrictive for corn on both farms and for soybeans on farm A, but labor is absolutely restrictive for soybeans on farm B.

---

[1]Table 2 results are obtained by variable resource programming; Heady and Candler discuss the technique (21, Chapter 7). Capital is varied on farm B (resource c_{2B}) from zero to ten units and optimal solutions are obtained for all capital levels within this range. The amount of capital available on farm D always equals the sum of the amounts available on farms A and B. As a result, corresponding optimal solutions for farm D are obtained for capital levels between three and 13 units. The sum of solutions for farms A and B are then compared with the optimal solutions for farm D for each capital level within the programmed range. The aggregation error for each level of capital is recorded in Table 2 for both products.

Table 2. Relation of amount of capital on farm B to absolute resource restriction and direction of aggregation error

Amount of capital on farm B	Products and farm for which given resource is absolutely restrictive: Farm A, Labor	Farm A, Capital	Farm B, Labor	Farm B, Capital	Direction of aggregation error	
$0 \leq c_{2B} \leq 4$	-	C,S	-	C,S	C S	None None
$4 < c_{2B} < 8$	-	C,S	S	C	C S	-[a] +
$8 \leq c_{2B} < 9$	-	C,S	C,S	-	C S	- +
$c_{2B} = 9$	-	C,S	C,S	-	C S	None +
$9 < c_{2B} < 10$	-	C,S	C,S	-	C S	+ +
$c_{2B} = 10$	-	C,S	C,S	-	C S	+ None
$c_{2B} > 10$	-	C,S	C,S	-	C S	+ -

[a]A negative error occurs when the estimate of the aggregate farm, farm D, is less than the sum of the individual farms A and B.

The third possibility is for the two farms to have differences in the absolute resource restrictions for both products as in Table 2 when $c_{2B} \geq 8$. In this case capital is the absolute resource restriction for both products on farm A, and labor is the absolute resource restriction for both products on farm B.

A relation between these three combinations of absolute resource restrictions and aggregation error is apparent from Table 2. When both farms A and B have the same absolute resource restrictions for both products, no error arises from aggregating. When farms A and B differ in absolute resource restrictions for one product ($4 < c_{2B} < 8$), the aggregation error is negative for corn and positive for soybeans. When the farms differ in absolute resource restrictions for both products ($c_{2B} \geq 8$), the aggregation error may be negative for corn and positive for soybeans, or any of four other combinations, depending on the level of the capital restriction.

By varying in turn each of the resources on farms A and B, it is possible to generate the nine possible combinations of absolute resource restrictions which are presented in Table 3. (Table 2 results now appear as possibilities 1, 2 and 3 in Table 3.) Here again there is no error when both farms have the same absolute resource restrictions for both products as in possibilities 1, 5 and 9. When the farms differ in absolute restrictions for one product (possibilities 2, 4, 6 and 8), the error is either negative for corn and positive for soybeans or positive for corn and negative for soybeans. However, when the farms have different absolute resource restrictions for both products as in possibilities 3 and 7, it is possible to have any of the five different combinations of aggregation error.

Table 3. Relation of possible absolute resource restrictions and direction of aggregation error

	Products and farm for which given resource is absolutely restrictive				Direction of aggregation error	
	Farm A		Farm B			
Possibility	Labor	Capital	Labor	Capital		
1	-	C,S	-	C,S	C S	None None
2	-	C,S	S	C	C S	- +
3	-	C,S	C,S	-	C S	- 0 + + +[a] + + + 0 -
4	S	C	-	C,S	C S	- +
5	S	C	S	C	C S	None None
6	S	C	C,S	-	C S	+ -
7	C,S	-	-	C,S	C S	- 0 + + +[a] + + + 0 -
8	C,S	-	S	C	C S	+ -
9	C,S	-	C,S	-	C S	None None

[a]All of the combinations of error presented may result.

The results of the two-product example summarized in Table 3 are much more enlightening than the original one-product example. Now it is apparent that several combinations of directions of error are possible for the two products. These combinations depend upon the ratios in which the labor and capital resources are held on the two farms, that is, which resources are absolute restrictions. Negative error makes an appearance in the two-product case -- it was not a possibility in the one-product case.

Agreement of the two-product example to theory

The relationships of Theorem III and its corollary are visible in this two-product example. First note in Figure 3 that the point representing the sum of solutions on farms A and B (corn = 2 and soybeans = 2 1/2) is indeed within the restraints and thereby feasible on farm D. However, this production point is not optimal for farm D -- as a result the optimal net revenue from farm D exceeds the sum of the net revenues from farms A and B. This relationship agrees with Theorem III.

The relationship between the corollary of Theorem III and the two-product example is visible from the last column of Table 3. In every possibility containing a negative error for one product, the other product contains a positive error. When the vector of net returns is nonnegative, it is impossible to vary the resources so that the error is negative for both products. On the other hand, it is possible to have both positive errors, as found in small ranges in possibilities 3 and 7 on Table 3. These results agree with the corollary of Theorem III.

These results can also be related back to Theorem I in Chapter III. Farms A and B both have the same coefficient matrices. In the two-product case, the condition that both farms have the same net revenue vectors and the same absolute restrictions for both products is enough to assure that they have qualitatively homogeneous output vectors. Both farms produce either (1) corn and soybeans, (2) corn and dispose of labor, (3) corn and dispose of capital or (4) some other combination of these activities, depending upon the relative amounts of the two resources available. If both farms have the same absolute restrictions for both products, they have the same combination of activities and no aggregation error results. As such, these results are completely compatible with Theorem I.

In addition, a very meaningful interpretation of the requirement of qualitatively homogeneous output vectors (QHOV) may be made with the aid of Figure 3. Farms having QHOV will all have optimum solutions at similar corner points of the convex polyhedrons representing their constraint sets. In Figure 3, the optimum solution for farm A is at the intersection of the X_2 axis and the capital restraint while the optimum solution for farm B is at the intersection of the labor and capital restraints. These two corner points are not similar and aggregation error results. This error would be eliminated if capital on farm A were expanded to the point where the two restraints met -- or by any other change of resources that would allow solutions at similar corner points of the constraint sets of the two farms.

Bias tendencies in the two-product example

What can be said about the predominate direction of error in the estimated activity levels of this two-product example? To answer this question, refer back to Table 3. Since two opposing directions of errors for a single crop have offsetting effects, the loose argument may be made that a positive and a negative error possibility would cancel each other and contribute to an unbiased estimate. Likewise, a large excess of positive error possibilities over negative error possibilities would be consistent with a positive bias in the estimate for that crop.

Referring to Table 3, this reasoning reveals an excess of positive errors for each product. However, this excess arises solely from possibilities 3 and 7; the number of positive errors equals the number of negative errors in the remainder of the possibilities. Reference back to Table 2 (which includes a blow-up of possibility 3, Table 3) shows that only a very small range of resource availability ($9 \leq c_{2B} \leq 10$) is actually responsible for the excess of positive over negative error. Thus, the excess of positive error possibilities results from a relatively improbable set of resource ratios. Such an excess arises from the lack of one situation -- there is no situation in which all products have negative error. Under the conditions of the example, it is not possible for all product estimates to have negative error. However, the number of attainable error combinations would grow as the size of the model increases. As such, the small bias tendency would tend to grow relatively smaller as the number of resources in the model increase. As a result, the tendency for bias in the estimate would be very minor in relation to the total error that may occur in multi-resource, multi-product models.

This small tendency for bias results from the lack of a situation in which all products have negative error. This situation of all negative errors doesn't exist when the net return vector is nonnegative. However, when the net return vector contains negative elements, it is possible to have negative aggregation error in all products. Such a situation is shown in Figure 4. It represents the programming problem of two individual farms (i = M, N) and the aggregate farm (i = P) where the objective is to:

Maximize $\quad Z X_i$

$$\begin{bmatrix} 2 & -3 \end{bmatrix} \begin{bmatrix} x_{1i} \\ x_{2i} \end{bmatrix}$$

subject to $\quad B X_i \le C_i$

$$\begin{bmatrix} 1 & -1 \\ -2 & -2 \end{bmatrix} \begin{bmatrix} x_{1i} \\ x_{2i} \end{bmatrix} \le C_i$$

and $\quad X_i \ge 0$

where
$$C_M = \begin{bmatrix} 1 \\ -6 \end{bmatrix}, \quad C_N = \begin{bmatrix} 2\frac{1}{2} \\ -4 \end{bmatrix} \text{ and } C_P = \begin{bmatrix} 3\frac{1}{2} \\ -10 \end{bmatrix}.$$

Here farms M, N and P have the same coefficient matrices and the same net return vectors containing one negative and one positive value; farm P has the sum of the resources of farms M and N.

The aggregate farm, farm P, estimates both products with negative aggregation error. Thus, when the net return vector is unrestricted in

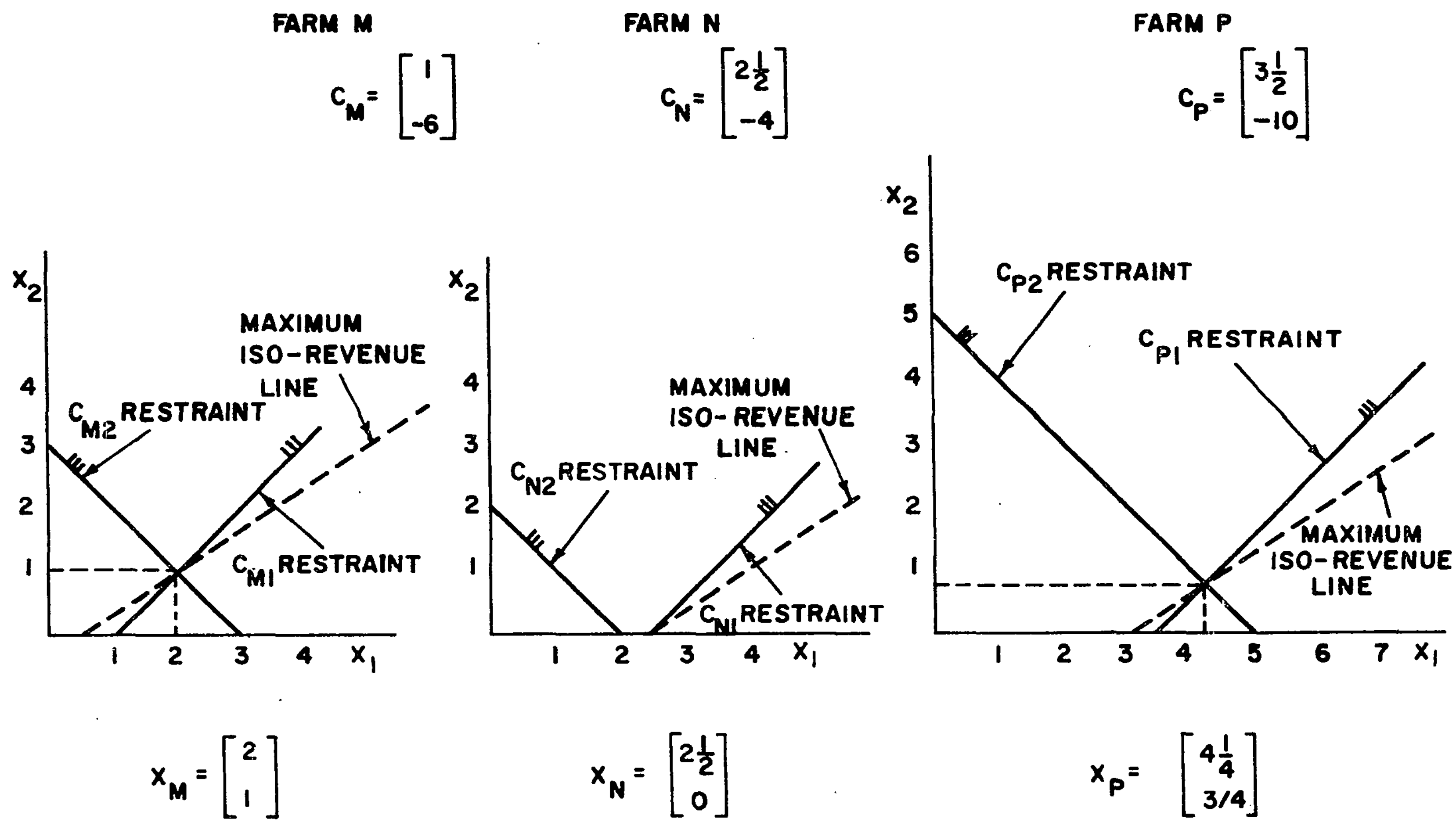


Figure 4. Graphical solution of linear programming problems for farms M, N and P

sign, the observed tendency toward positive error resulting from the absence of a situation with all negative errors is eliminated. No tendency for a predominance of error in any one direction is observable in this example.

Figure 4 also tends to dispel an argument for positive aggregation bias in models where one product is strongly predominate, for example, dairy farms where milk is the only salable product. For such a model, the argument has been made that the supply estimate of milk would be positively biased because of the relation of Theorem III. The reasoning is that if only one activity has a positive net return value, then that one activity must have a positive error because the functional overestimates the total net returns. In fact, the studies of Frick and Andrews and Sheehy and McAlexander discussed in Chapter II would tend to substantiate this reasoning. They found only positive error in milk supply functions.

The example of Figure 4 shows that this reasoning is inconclusive. Such reasoning overlooks the possibility that an activity with a negative net return may be negatively biased and in this way cause the expected positive error in the functional. It would seem that some other explanation would be required for the predominance of positive error found in this empirical work with milk supply functions.

Few concrete statements can be made about the relative probability or improbability of these situations in the real world. This is not intended. What these examples point out is that the expectation of negative aggregation error is roughly the same as the expectation for positive aggregation error in general representative farm linear programming models. These examples suggest the hypothesis that in the representative farm linear programming supply estimation model where all coefficients are unrestricted as to sign, the aggregation error in the product estimates is unbiased. This suggestion is in direct contrast to the impression left by earlier researchers. It is unfortunate that the tools that have been developed are not strong enough to conclusively verify or refute this hypothesis.

It is hoped that this discussion of the direction of aggregation error has placed the popularly held concept of bias in representative farm linear programming supply estimates in its proper place. The indicated lack of such bias makes many avenues for handling the aggregation error problem appear more useful. These avenues center around statistical concepts of minimizing error in sample estimates. Some obvious ones are the use of larger numbers of observations (larger numbers of representative farms) and stratification to reduce within-group variance. Such techniques should reduce expected error in unbiased estimators. In addition, elimination of the idea of bias from such errors strips them of many of their formidable aspects. Perhaps this in itself goes some distance towards increasing the apparent validity of the representative farm linear programming supply estimation procedure.

Situations Causing Aggregation Error

These results and Figure 3, page 66, may be evaluated to identify the situations that will lead to aggregation error in representative farm linear programming supply estimates. For example, the labor restriction is redundant on farm A in Figure 3. It is not, however, redundant on farm B nor in the aggregate farm, farm D. Thus, some labor is forced to go unused on the individual farms that is fully used on the aggregate farm. Under this situation, aggregation error occurs. On the other hand, if labor were also redundant on farm B, it would also be redundant on farm D and no aggregation error would arise.

As a result, one cause of aggregation error can be intuitively explained in terms of resource redundancies. If a resource is redundant on one farm, it must be redundant on all other farms in the group before exact aggregation can be achieved. One of the goals of farm stratification should be grouping farms into groups such that all farms within a group will have the same resources limiting and the same resources redundant. Such grouping would eliminate aggregation error from this cause.

Nonhomogeneous output vectors

A more rigorous definition of a situation causing aggregation error can be obtained from Theorem I. Theorem I stated that the condition of qualitatively homogeneous output vectors (QHOV) is a sufficient condition for exact aggregation. Under certain conditions, QHOV is also a necessary condition for exact aggregation.

<u>Theorem IV</u> If all of the elements of the optimal basic vectors are nonzero, the condition of exact aggregation can be achieved only if all farms in the set to be aggregated have qualitatively homogeneous output vectors (QHOV). Restated, QHOV is a necessary condition for exact aggregation under this condition.

<u>Proof</u> Assume the theorem is false and that exact aggregation is accomplished without QHOV. Then n-1 farms of the n farms in the group may have a common set of k basic nonzero variables in their optimum solutions but at least one farm must have a new basic nonzero variable replacing one of the k common to the rest. Under this condition

$$\sum_{g=1}^{n} X_g$$

will have the k basic nonzero variables common to the set of n-1 farms plus the one additional nonzero variable for the nonhomogeneous farm, a total of k + 1 nonzero elements.

However, X has only k basic nonzero variables. Therefore,

$$X \neq \sum_{g=1}^{n} X_g$$

contradicting the basic premise.

Theorem IV and its proof are dependent upon the requirement of nonzero optimal basic vectors. This requirement is necessary because of a technicality of linear programming that has little counterpart in real world production problems. Zero level optimal basic variables are quite common in linear programming problems of the type being considered and, although they are technically a form of degeneracy, their occurrence presents no problem. If, for example, corn is in the optimal solution at a zero level, it is merely ignored in the interpretation of the results. The occurrence of such zero level optimal basic variables allows exact aggregation to occur without QHOV being met, as long as the "nonhomogeneous" activities are all in their respective solutions at zero levels. Farms not meeting the requirement of QHOV may be aggregated without error if all basic activities outside the set of basic activities common to all farms are at zero levels. This problem arises from QHOV being strictly defined as a requirement of optimal basic vectors and not merely one of nonzero activity levels.

A slightly relaxed interpretation of QHOV as vectors defining similar corner points would overcome this difficulty, from an intuitive standpoint at least. A zero level optimum activity would occur in farm A in Figure 3 if capital were expanded to 4 units, at which point the capital and labor restraint would meet at the X_2 axis. At this point, the solution to farm A would be 2 units of soybeans and zero corn or 2 units of soybeans and zero disposal of either labor or capital. In either event, there would be no redundant resource on either farm and farms A and B could be aggregated without error. Thus, zero level activities do not present any real problem in this context. The solutions of the farms would still be at similar corner points of their constraint sets. They may not have QHOV in the strict sense of the definition, but the difference is not too important from a practical standpoint. With this loose interpretation, any situation in the real world which is observed as not meeting the requirement of QHOV would be expected to generate aggregation error.

Variation in net returns

Another situation causing aggregation error is variation in net return vectors among individual farms. In general, such variation becomes a problem when it destroys the condition of QHOV. Variation in net returns within the range in which QHOV is still achieved does not lead to aggregation error. Variation in net returns for a farm changes the slope of its iso-profit lines. Any farm may have a different net return vector and as long as the slope of its iso-revenue line is not changed to the extent that would move the solution to a new corner point, the farms still meet the QHOV requirement and may be aggregated without error.

As a result, Theorem I covers the problem of aggregation of farms with different net return vectors. Theorem I does not include a condition on net return vectors simply because the requirement of QHOV overrides such a condition. As long as QHOV is met, there is no need to be concerned about variation in net returns. Only when net returns vary outside this range do they become a problem. Then additional stratification must be made to eliminate aggregation error.

Sheehy gives a very good discussion of the way variation in costs can cause aggregation error (41, pp. 34-36). He presents a figure similar to Figure 5 showing the hypothetical stepped supply functions for two farms and the actual and estimated aggregate supply functions. The first of these farms is able, through lower costs, to expand output at a lower price (P_3) than the second farm, which doesn't expand production until P_1 is reached. This leads to two steps in the true supply

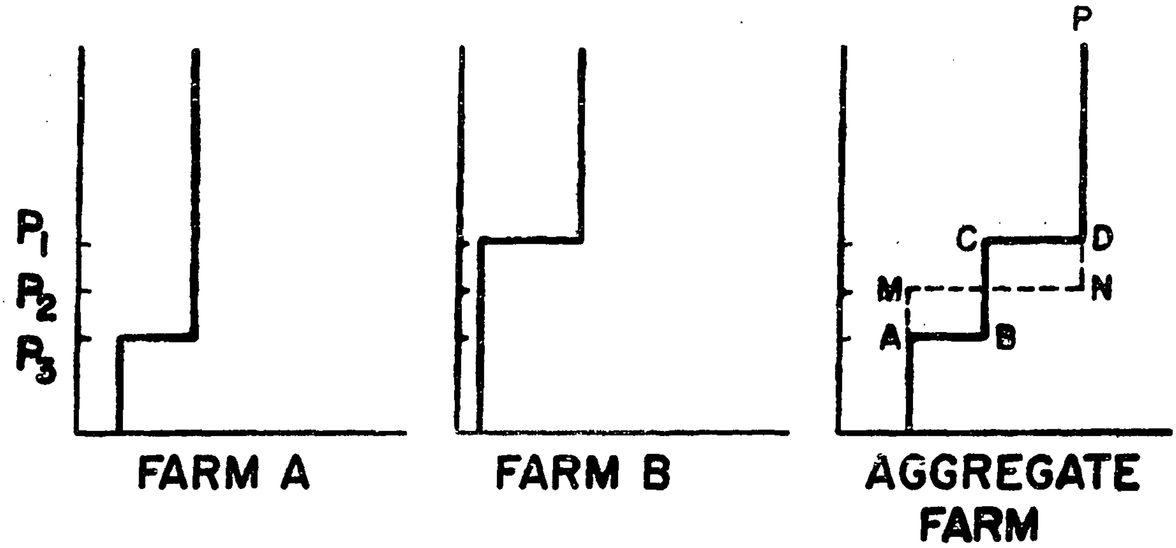


Figure 5. Effect of variation in costs of individual farms on aggregate supply functions

curve found by summation, line OABCDP. The aggregate farm will, however, have only one step, just as the individual farms did. If its costs are the average of the costs on farms A and B, it will expand production at some intermediate price, say P_2. This will make the supply curve estimated by the aggregate farm OMNP. Actual supply will be underestimated between P_3 and P_2 and overestimated between P_2 and P_1.

Figure 5 casts light on two aspects of the aggregation problem. The first is on the possible use of modal rather than average costs for the representative farm. Using the mode, the supply would either be underestimated or overestimated all of the way from P_3 to P_1, depending upon the individual farm that was chosen as the mode. Considering this, using average costs may lead to more desirable estimates in some cases than using modal costs.

Sheehy's discussion also gives a special meaning to the problem previously discussed in reference to Theorem I -- the problem of exact aggregation under varying prices. If the problem is approached from the direction of aggregating the supply of different farms at given and discrete sets of prices, the previous discussion is accurate. There will be no aggregation error if all farms have solutions meeting the QHOV requirement at each point aggregated. However, if the problem is viewed as one of estimating an entire supply curve, the omission of a net return requirement from Theorem I may be misleading, although the theorem is still valid without it. The omission tends to hide the fact that farms may have to be stratified into groups of nearly identical net return vectors in order to obtain QHOV and achieve exact aggregation at all points on the supply curve. At a given price, different net returns and QHOV will give exact aggregation. However, as prices are varied in generating a supply function, at some point a given difference in net returns will force QHOV not to be met. Farm groups will then have to be subdivided. The problem is that the probability of reaching such points is increased by varying prices. Thus, for estimation of a complete supply function, it may be necessary to approach identical net return vectors before QHOV can be obtained at all points.

Variation in coefficient matrices

A variety of different factors may affect the supply response of individual farms and hence become possible candidates for causing aggregation error. These factors have been classified as (41, p. 21):

(1) Physical environment, such as climate and topography.

(2) Institutional restrictions, such as markets and government regulations.

(3) Motivational forces, including risk aversion and demand for leisure.

(4) Management ability.

(5) Technology.

(6) Resource endowments and mobility.

Generally, variance in the first five factors affects the coefficient matrices of the individual farms. The sixth factor, variance in resource endowments, has been the variable receiving major emphasis so far in this chapter.

Achieving exact aggregation for a set of farms with different coefficient matrices is difficult but not impossible. Assuming that a set of farms has different coefficient matrices but still meets the requirement of QHOV, at least three cases can be visualized where exact aggregation could still occur. The first may be thought of as the type of coefficient variance equivalent to row scaling. For any linear programming problem, a given row (including the value in the resource vector for that row) can be multiplied by a constant without affecting the solution. As a result, variation that similarly affects all elements of a row can occur in the coefficient matrices of individual farms without leading to aggregation error. To accomplish this, the coefficient matrix for the representative farm must be defined as the average of the coefficient matrices of the individual farms. Then, for one farm, if one coefficient differs by a certain factor and all other coefficients in the same row differ by the same factor (including the resource vector coefficient), that farm can be included in the set without aggregation error. Restated, once one coefficient in a set of farms differs, then all rows containing that coefficient in the set of farms must be scaler multiples of each other.[1] Such variation may occur in actual data where larger amounts of a resource are offset by decreased productivity.

Another type of variation that could occur in coefficient matrices and not cause aggregation error is more likely but also of little concern

---

[1]The importance of including the resource vector item for that row in the requirement cannot be overemphasized. Aggregation error will occur when the appropriate resource coefficient is omitted from the requirement.

from a practical standpoint. This is (1) variation in coefficients of activities that are not in the optimum solutions of any of the farms and (2) variation of coefficients in resource rows that are not restrictive on any of the farms. Such variation does not lead to aggregation error simply because it does not enter into any of the solutions.

There is a third possible situation where variation in coefficients may occur and still not lead to aggregation error. This is when only one product is to be estimated and the conditions of Sheehy's aggregation principle hold -- that is, when one resource has a proportional relation to the product in question (41, p. 22). Under this condition, coefficients may vary as long as the proportionality is maintained.

These three cases would all appear to be somewhat improbable in actual data. Generally, other types of variation would be expected among the coefficient matrices of individual farms. This variation would lead to aggregation error. As a result, stratification of farms into groups with identical coefficient matrices would be the first step in representative farm identification. This step would be followed by definition of substrata for each of these groups until each farm within each of the substrata would meet the requirement of QHOV. These two steps would assure achieving exact aggregation.

Under widely varying coefficient matrices, the concept of QHOV can become rather abstract in itself, since it depends upon identity of activities. Do two activities with different coefficients have the same identity? Are resources with different productivities the same? Inability to answer this type of question may make a strict interpretation of the QHOV requirement difficult when the coefficient matrices for the individual farms are not equal. This is another way variation in the coefficient matrices confuses the concept of exact aggregation.

CHAPTER V. PROGRAMMED SUPPLY FUNCTIONS FROM DIFFERENT STRATIFICATIONS

The theory discussed in the previous two chapters leaves unanswered questions concerning (1) the magnitude of the aggregation error in actual models and (2) the relative importance of different factors contributing to aggregation error. These are essentially empirical questions with the answers depending upon the area, type of agriculture and type of supply estimates desired. Generally, there are unique answers for each specific research project. Relationships identified as important in one project may not be important in another.

A model was developed to answer these questions concerning aggregation error for an existing research project involving the technique of representative farm linear programming supply estimation. This chapter discusses the development of the model and presents supply functions that were estimated using four different groups of representative farms. Chapter VI analyzes the aggregation error and the possible factors that contribute to it and Chapter VII presents a discussion of what this work reveals in reference to developing stratifications to reduce aggregation error.

The divisions of this chapter follow the steps of the technique of representative farm linear programming supply estimation that have been outlined in Chapter I. These steps may be summarized as (1) definition of the population of interest and the supply estimates desired, (2) collection of sample data, (3) determination of strata and representative farm resources, (4) development of the linear programming models for each representative farm and (5) computations leading to the final estimates. These steps provide a logical order for discussing the development of the model.

The Population Defined

Iowa's contributing project to the North Central regional research project known as NC-54 was chosen as the basis for this study of aggregation error. The title of the regional project was "Supply Response and Adjustments for Hog and Beef Cattle Production." The Iowa contributing project was carried on cooperatively by the Iowa Agricultural Experiment Station and the Economic Research Service, U.S. Department of Agriculture. It represented an important and continuing segment of the research of these two institutions. This project utilized the technique of representative farm linear programming supply estimation. The data had already been collected at the time the aggregation error study was initiated. For these reasons the project represented a relevant research framework to use in exploring the actual problems of aggregation error.

Choice of this research project as a basis for the study of aggregation error automatically provided answers to most of the questions concerning the population of interest and the specific supply function estimates desired. The estimates were desired for the population of commercial farms in Iowa. Specifically, the population was defined using Census of Agriculture definitions and included all livestock, general and cash-grain farms in economic classes I through V in Iowa. This classification included nearly all farms except those receiving over 50 percent of their income from poultry or dairy and farms with less than $2,500 per year gross sales. The 1959 census listed 136,331 farms in the defined population or 78.0 percent of all farms in the state (51, pp. 32-33). However, these farms accounted for 89.4 percent of the total farmland in the state and a similar high percentage of other important resources and major products.

The desired estimates for this population were intermediate-run supply functions for pork and beef. Determination of these supply function estimates was one of the objectives of the regional project. It provided a two-product situation for the study of aggregation error -- a desirable situation since previous empirical studies of aggregation error had involved only one-product models.

The Sampling Procedure and Data

The sampling procedure was developed and the data were collected to meet the requirements of the NC-54 project in Iowa (38). No additional data were collected for the study of aggregation error. A 5 percent sample of 1959 Census of Agriculture data was used as the basis for much of the information on farm resources. These data were supplemented by mailed questionnaires to Iowa county agents concerning livestock facilities available on the representative farms and to Iowa bankers concerning the estimated financial position of typical groups of farmers. Data on costs, returns and input-output coefficients on farms were collected from secondary sources, namely State Experiment Station bulletins. Most of the data for livestock coefficients were developed by the regional NC-54 committee as a means of obtaining comparability among all states participating in the project. The data collection techniques are all described in detail by Sharples in a report on one phase of the Iowa NC-54 work (39, pp. 20-33).

Information obtained on individual farms in the 5 percent sample of census data included all of the farm characteristics found in the published form of the Census of Agriculture. Major sections of information were land use, tenure, land value, type of farm, labor used, cash expenditures, conservation practices, machinery inventory, livestock programs and fertilizer use. This core of data provided the basic information on resources for individual farms.

The Stratification Procedure

The definition of aggregation error makes its exact measurement an expensive and often nearly impossible task because exact measurement implies programming every farm in the population of interest. The alternative used for this study is to make four stratifications involving successively smaller numbers of representative farms. The differences among the state supply functions estimates based on these four groups of representative farms are then due to aggregation error. This error is then analyzed to show how it is affected by the number of representative farms programmed and by the method of stratification. This procedure does not give an exact measurement of aggregation error. Rather, it shows how aggregation error accumulates as smaller numbers of representative farms are used in the estimation of state supply functions.

The four alternative stratifications were developed to include the range in numbers of representative farms that would possibly be considered in future regional adjustment studies involving Iowa. They were a basic stratification which resulted in 36 representative farms and three less detailed stratifications involving ten, three and one representative farms. Aggregation coefficients were defined for each of these four groups of representative farms so that four different sets of Iowa supply function estimates could be obtained.

Basic stratification into 36 representative farms

The basic stratification of individual farms used for the study of aggregation error differed from the one used for the NC-54 project. The stratification for the NC-54 project was made on the basis of (1) ten soil association areas of the state, (2) three sizes of farms in acres and (3) two types of farms (livestock farms and cash grain or general farms on the basis of the Census of Agriculture definitions). This procedure resulted in a total of 63 representative farms for the NC-54 project.

Programming results for the NC-54 project revealed that no significant differences in the identity of activities in the optimal solutions resulted from the third stratification factor, type-of-farm. When considered in view of the requirements of Theorem I, this result would indicate that type-of-farm stratification accomplished little in increasing the accuracy of the estimates. Therefore, it was not used as a stratification factor in the study of aggregation error.

Thus, the two main stratification factors for the study of aggregation error were: (1) soil association area of the state and (2) size of farm. Both previous NC-54 results and knowledge about the agriculture of the state suggested that these two factors played an important role in determining the optimal organization or response of individual farms in the state. Stratifying the sample of 6,800 individual farms by these two factors was designed to delineate strata of individual farms that approximated the conditions of Theorem I -- that is, the strata of farms would have nearly identical coefficient matrices and would approach the condition of qualitatively homogeneous output vectors. Such a stratification would tend to minimize aggregation error in estimates developed using the basic group of 36 representative farms.

The stratification by soil association area of the state was developed with the help of soil scientists and agronomists. The areas used, which followed county lines because of the availability of most other data at the county level, are shown in Figure 6. Land quality, crop yields and fertilizer practices were found to vary greatly among these ten areas. Ten different sets of crop yield coefficients and fertilizer costs were developed to recognize these differences, based on the work of Shrader and others (43, 44).

Sample farms in each of these ten areas of the state were then divided into three strata on the basis of farm size. The three size strata, based on total farmland, were (1) less than 140 acres, (2) from 140 to 240 acres (inclusive) and (3) greater than 240 acres. As with the stratification by area, the stratification by size was designed

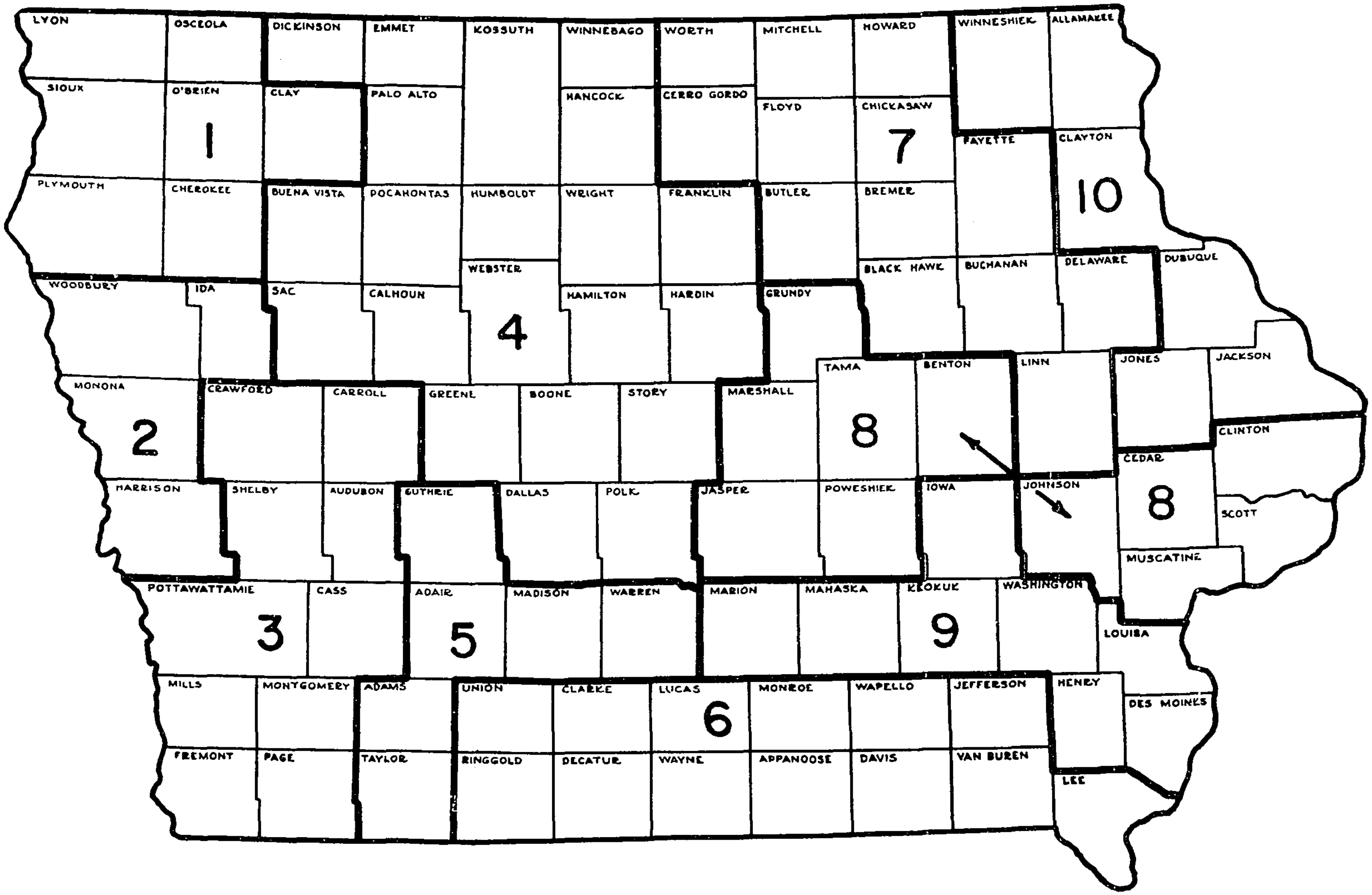


Figure 6. Location of the 10 soil association areas in Iowa

to separate farms into groups with similar coefficient and production response characteristics. These three size strata represented three farm sizes common in Iowa, namely 80 acre, 160 acre and 320 acre farms.

The classification by soil association area and size of farm divided the 6,800 sample farms into 30 strata. One or two representative farms were delineated for each of these strata in the following manner. First, one representative farm was defined for each of the small and medium farm size strata in each of the ten soil-type areas. In areas 1, 3, 5 and 7, one representative farm was defined for the large farm strata. In the remainder of the areas, two representative farms were defined for the large farm strata. These additional farms recognized segments of sample farms having significantly different hired labor availabilities. These steps resulted in 36 representative farms in the basic stratification.

The 36 representative farms were developed to possess the typical bundle of resources of the strata rather than the average resources. This was done primarily because a similar procedure had been followed for the NC-54 work and all of the secondary data on farm resources had been compiled for typical rather than average farms.[1] The size and

---

[1]The typical or modal concept of a representative farm has certain advantages when the results are used for a purpose that has not been mentioned previously -- making recommendations to individual farmers. Generally, the optimum programming results for a modal representative farm would have applicability for a larger number of real world farms than the optimal programming results of an average representative farm. Following such a procedure unfortunately may tend to increase aggregation error in the estimated state supply functions. However, it does not affect the amount of aggregation error found by comparing the four groups of supply functions developed in this study. This is because the modal farm concept is used only in the basic group of 36 representative farms against which the others are judged. The resources on the representative farms in the three subgroups are defined as averages of the 36 farms and not as modes.

location of these representative farms are shown in Table 4. The first part of the farm number identification tells the area in which the farm is located; the second part denotes the size stratum; and where used, the lower case letter denotes the two levels of hired labor available on the farm. As an example, farm number 2Ca designates the large representative farm in area 2 with a lower availability of hired labor.

The aggregation coefficients shown in Table 4 represent the factors necessary to (1) aggregate the 36 representative farm results up to the sample total for the 6,800 sample farms and (2) estimate the state response from the 6,800 farm sample in a single step. The discrepancy between the aggregation coefficient total of 135,375 farms and the population total of 136,331 farms arose from using the modal rather than the average representative farm. Generally, the size of the modal farm did not equal the average size of farm in a particular strata. As a result, the aggregation coefficients were defined as the total cropland acreage in each strata divided by the cropland acreage of the respective representative farms. Following this procedure assured that the total cropland figure for the population would be equaled by the aggregation of the representative farms -- a desirable characteristic in view of the primary importance of cropland in determining the amount of production for the representative farms.

This basic group of 36 representative farms was developed to provide population supply estimates that were relatively free of aggregation error. The stratification of sample farms was carried out with the objective of approximating the conditions of Theorem I. As was brought

Table 4. Average size and aggregation coefficients for the group of 36 representative farms

Farm number	Cropland	Total farmland	Aggregation coefficient
	(acres)	(acres)	
1A	75.4	87.9	2,104
1B	151.7	177.0	5,521
1C	282.4	329.4	4,134
2A	58.0	77.0	598
2B	134.0	169.0	2,571
2Ca	275.5	364.8	2,372
2Cb	474.0	513.0	256
3A	73.8	93.1	2,775
3B	149.0	177.8	5,118
3C	303.7	374.3	5,731
4A	76.0	93.0	3,108
4B	156.7	176.1	14,650
4Ca	268.9	311.2	7,771
4Cb	316.0	362.2	4,608
5A	58.0	83.4	1,544
5B	124.2	175.7	2,767
5C	245.1	401.3	2,855
6A	58.0	97.0	1,491
6B	122.9	191.7	3,621
6Ca	204.2	365.6	4,716
6Cb	256.0	410.0	996
7A	70.5	90.7	4,683
7B	148.6	178.3	7,847
7C	286.4	348.2	4,926
8A	76.8	95.4	4,250
8B	153.0	178.4	8,474
8Ca	244.8	317.0	3,763
8Cb	279.0	355.0	1,320
9A	33.8	57.5	4,917
9B	137.4	180.8	4,201
9Ca	259.7	355.8	3,319
9Cb	310.0	423.0	328
10A	62.0	94.0	1,815
10B	121.7	178.0	4,042
10Ca	202.0	342.0	1,469
10Cb	265.0	378.0	714
			135,375

out in Chapter III, the number of representative farms required to assure exact aggregation is extremely large for the size of model being considered; however, it appears possible to achieve reasonably accurate aggregation with much smaller numbers of representative farms.

Substratifications

After the basic determination of 36 representative farms was completed, smaller groups of representative farms were developed by computing weighted averages of the resources of the original 36. Data for the three subgroups of representative farms developed in this manner are presented in Table 5.

The first subgroup of ten representative farms consisted of an average farm in each of the ten soil association areas of the state. For example, farm 1BB, the average farm in area 1, is a weighted average of resources on farms 1A, 1B and 1C with the aggregation coefficients used as the weights. The aggregation coefficient for farm 1BB is then the sum of the coefficients of the other three farms. As a result the aggregation coefficients for the ten representative farms are the ones necessary to obtain population estimates for the state using these ten farms.

A second subgroup of three representative farms was delineated to represent the small, medium and large farms in the population. These three representative farms, StA, StB and StC in Table 5, were determined by weighted averages of resources on the ten small farms, the ten medium farms and the 16 large farms, respectively. Here again, the aggregation coefficients sum to the same state total and may be used to obtain population estimates based on this group of three representative farms.

Table 5. Average size and aggregation coefficients for the three subgroups of representative farms

Farm number	Cropland	Total farmland	Aggregation coefficient
	(acres)	(acres)	
Ten-farm subgroup			
1BB	184.0	214.6	11,759
2BB	199.1	254.4	5,797
3BB	198.8	243.2	13,624
4BB	201.7	230.8	30,137
5BB	158.1	245.7	7,166
6BB	161.6	274.5	10,824
7BB	166.6	202.7	17,456
8BB	163.6	201.0	17,807
9BB	133.7	185.0	12,765
10BB	135.6	206.8	8,040
			135,375
Three-farm subgroup			
StA	64.0	85.6	27,285
StB	145.6	178.0	58,812
StC	269.8	351.8	49,278
			135,375
One-farm subgroup			
StBB	174.4	222.7	135,375

The final stratification considered was one representative farm, StBB, which could be used alone in estimating production for the population of interest. Farm StBB was the weighted average of resources on all the 36 basic representative farms -- its aggregation coefficient was the total of the 36 basic aggregation coefficients, 135,375.

Due to the manner in which the four groups of representative farms were developed and their aggregation coefficients determined, they all represent use of exactly the same total amounts of resources in the population. For example the labor available in the 36 farm group multiplied by the respective aggregation coefficients sums to the same total as the labor available in the three farm group multiplied by the respective aggregation coefficients. As a result, the four sets of supply estimates developed are free of differences that would arise from using different amounts of resources.

The Representative Farm Models

A linear programming model was developed for each of the 50 representative farms. The models for all farms had the same number of restrictions and activities; however, the value of many coefficients varied from farm to farm. This section explains the purpose of the restrictions and activities included in the model. Appendix Table 1 shows the matrix of coefficients used for one of the representative farms.

The resource restrictions

The 36 restrictions of the models are identified in Table 6, along with the quantities of resources available on the medium sized representative

Table 6. Identification of restrictions and a typical resource vector for the linear programming model

Row number	Item	Unit	Amounts available farm 9B[a]
C_j	Net returns over variable costs	dollar	(maximized)
1	Land with 25% row crop capability	acre	20.7
2	Land with 50% row crop capability	"	51.2
3	Land with 100% row crop capability	"	65.4
4	Pasture (noncropland)	ton AHY[b]	50.5
5	Meadow to be harvested	ton	0
6	Corn to be harvested	bu.	0
7	Central farrowing facilities	sows	17.6
8	Portable farrowing facilities	"	0
9	Confinement feeding facilities	pigs	0
10	Portable feeding facilities	"	166.0
11	Beef housing - period 1	a.u.[c]	26.4
12	Beef housing - period 2	"	26.4
13	Low beef mechanization - period 1	head	8.6
14	Low beef mechanization - period 2	"	8.6
15	High beef mechanization - period 1	"	42.4
16	High beef mechanization - period 2	"	42.4
17	Corn equivalents	cwt.	0
18	Corn silage	"	0
19	Hay equivalents	"	0
20	Purchased yearlings - one period	head	0
21	Purchased yearlings - both periods	"	0
22	Beef calves	"	0
23	Cash account	$10	1,448.4
24	Chattel mortgage	"	346.5
25	Beef for sale	cwt.	0
26	Hogs for sale	"	0
27	Total operator and family labor	hour	2,368.0
28	Dec., Jan., Feb., March labor	"	786.4
29	April labor	"	246.6
30	May labor	"	271.6
31	June labor	"	321.6
32	July labor	"	321.6
33	August, September labor	"	568.2
34	October labor	"	246.6
35	November labor	"	221.6
36	Total hired labor limit	"	124.4

[a]This is the resource vector for the medium sized representative farm in area 9.

[b]Tons of anticipated hay yield.

[c]Animal units.

farm in area 9, farm 9B. All of these restrictions comprise the usual upper bounds to the activities of the solution with the exception of the first three -- these are strict equalities.

The first three restrictions represent the three classes of cropland available and sum to the total cropland on the farm. These restrictions reflect agronomic restraints on cropping intensity. The relative proportion of land in each of these three classes is different for each of the ten soil association areas studied, but the proportions are the same for all farms within an area.

Restriction 4 represents noncropland pasture available for grazing on the farms and restriction 5 represents cropland planted to alfalfa and grass which may be used for hay or grazing.

Restriction 6 is an accounting row for the intermediate product, unharvested corn, which is produced by the different crop rotations. It may be harvested either as grain or as silage.

Hog facilities is the subject of the next four restrictions, 7 through 10. These are central and portable farrowing facilities and confinement and portable feeding facilities.

The next six restrictions are on beef facilities. These are beef housing capacity and beef feeding facilities involving a low and a high level of mechanization. These are divided into two use periods of the year, November through April and May through October. These two periods are defined so that enterprises using the facilities at different times in the year will not compete for the same facilities. The same facility is represented in each of the two periods; thus, the amount available for both of the periods in the resource vector is the same.

Restrictions 17, 18 and 19 are accounting rows for intermediate livestock feed products. The corn equivalent row collects corn purchased and corn and oats harvested as grain from the rotations in corn equivalent units. It is available for feed to both the hog and beef enterprises, or for sale. The corn silage and hay equivalent rows make feed available to the beef enterprises.

The next three restrictions make purchased and farm raised feeder calves and yearlings available to the beef feeding enterprises.

Restriction 23 is the operating capital available on the farm in \$10.00 units. The amount available in the resource vector includes cash on hand and the farm value of feed and livestock inventories less short-term liabilities at the beginning of the year. Feed and livestock inventories were converted to cash and as a result were not included elsewhere in the resources available.

Restriction 24 limits chattle credit that may be obtained without providing any additional collateral. The amount available in the resource vector represents 50 percent of the owned machinery inventory less current intermediate term liabilities. Only 15 percent of the collateral required for livestock loans was required to come from this source -- the remainder was provided by the livestock itself.

The next two restrictions, numbers 25 and 26, accumulate all of pork and beef produced by the respective enterprises and allow its sale through two selling activities. The number of selling activities was minimized to accommodate the variable pricing technique used with the model.

Restrictions 27 through 35 are the operator and family labor restrictions. Restriction 27 is the total annual operator and family labor. The total amount available is less than the sum of the amounts available during the other periods of the year because the fixed labor requirements for the farm (which may be performed at any slack time during the year) have been deducted. The restrictions 28 through 35 represent eight potentially restrictive labor periods of the year.

The last restriction, number 36, sets a limit on the number of hours of labor that may be hired. Each representative farm's labor hiring was restricted to its historical level to prevent aggregate labor hiring in the state from exceeding the amounts of farm labor available.

The activities considered

Table 7 summarizes the 73 activities considered in the linear programming models of the representative farms. Again, the reader should refer to Appendix Table 1 for additional information about the structure of the matrix.

Activities P_1 through P_6 are the variable pricing section of the model. Four prices for both pork and beef or a total of 16 price combinations were programmed to generate supply functions for each representative farm. Pork prices programmed were \$10.50, \$11.00, \$12.00 and \$13.00 per hundredweight; beef prices programmed were \$14.00, \$15.50, \$17.00 and \$19.00 per hundredweight. Experience with previous NC-54 work in Iowa suggested that aggregate production forthcoming at these prices would bracket historical state production levels for both pork and beef. Both the functional value (C_j) of these six activities and selected

Table 7. Identification of activities for the linear programming models

Activity number	Explanation	Unit
P_1	Sell pork[a]	1,000 lb.
P_2	Sell beef[a]	cwt.
P_3	Buy feeder calves[a]	head
P_4	Purchase yearlings - one period program only[a]	"
P_5	Purchase yearlings - two period program[a]	2 head
P_6	Sell calves[a]	head
	Rotations for land with 100% row crop capability	
P_7	Corn	acre
P_8	Corn - soybeans	"
P_9	$CCOM_1$	"
P_{10}	$CSOM_1$	"
P_{11}	CSSOM	"
	Rotations for land with 50% row crop capability	
P_{12}	$CCOM_2$	"
P_{13}	$CSOM_2$	"
P_{14}	CSSOMM	"
P_{15}	$Meadow_2$	"
	Rotations for land with 25% row crop capability	
P_{16}	COMM	"
P_{17}	$Meadow_3$	"
	Harvest corn	
P_{18}	As grain	10 bu.
P_{19}	As silage	ton
	Harvest meadow	
P_{20}	As hay	"
P_{21}	By grazing	ton AHY[b]
	Hog enterprises	
P_{22}	One sow - one litter - portable farrow and feed	1 litter
P_{23}	One sow - two litters - portable farrow and feed	2 litters
P_{24}	One sow - two litters - central farrow and portable feed	"
P_{25}	One sow - two litters - central farrow and feed	"
P_{26}	Two sows - four litters - central farrow and portable feed	4 litters
P_{27}	Two sows - four litters - central farrow and feed	"
P_{28}	Three sows - six litters - central farrow and portable feed	6 litters
P_{29}	Three sows - six litters - central farrow and feed	"

[a]These activities were variable priced to generate the supply curves.

[b]Tons of anticipated hay yield.

Table 7. (Continued)

Activity number	Explanation	Unit
	Beef feeding enterprises	
	No silage - low mechanization feeding	
	Calves	
P_{30}	Drylot	head
P_{31}	Pasture	"
	Yearlings	
P_{32}	Period 1[c]	"
P_{33}	Period 2[d]	"
P_{34}	Periods 1 and 2[e]	2 head
	No silage - high mechanization feeding	
	Calves	
P_{35}	Drylot	head
P_{36}	Pasture	"
	Yearlings	
P_{37}	Period 1	"
P_{38}	Period 2	"
P_{39}	Periods 1 and 2	2 head
	Silage - low mechanization feeding	
	Calves	
P_{40}	Drylot	head
P_{41}	Pasture	"
	Yearlings	
P_{42}	Period 1	"
P_{43}	Period 2	"
P_{44}	Periods 1 and 2	2 head
	Silage - high mechanization feeding	
	Calves	
P_{45}	Drylot	head
P_{46}	Pasture	"
	Yearlings	
P_{47}	Period 1	"
P_{48}	Period 2	"
P_{49}	Periods 1 and 2	2 head
	Beef cows	
P_{50}	Hay	1 cow unit
P_{51}	Silage	"
P_{52}	Borrow chattel credit	$100
P_{53}	Invest cash off farm	"

[c]Fed October to April.

[d]Fed April to October.

[e]One steer fed October to April and another steer fed April to October.

Table 7. (Continued)

Activity number	Explanation	Unit
P_{54}	Reservation price on labor	hour
P_{55}	Sell corn	cwt.
P_{56}	Buy corn	"
	Investment in additional hog facilities	
P_{57}	Central farrowing	1 litter
P_{58}	Portable farrowing	"
P_{59}	Confinement feeding	"
P_{60}	Portable feeding	"
P_{61}	Central farrowing and feeding	"
	Investment in beef facilities	
P_{62}	Housing	a.u.[f]
P_{63}	Low mechanization feeding equipment	head
P_{64}	High mechanization feeding equipment	"
P_{65}	Conversion of low mechanization to high mechanization feeding equipment	"
	Labor hiring activities	
P_{66}	Dec., Jan., Feb., March	hour
P_{67}	April	"
P_{68}	May	"
P_{69}	June	"
P_{70}	July	"
P_{71}	August, September	"
P_{72}	October	"
P_{73}	November	"

[f] Animal units.

internal matrix values varied simultaneously as the different prices were programmed. Coefficients which varied with price are denoted by an "a" superscript in Appendix Table 1. The variable pricing section of the model was designed to provide the required supply estimates and to recognize changes in the purchase price and credit available on feeder calves and yearlings as the selling price of beef changed.

The quantities of pork and beef produced on the representative farms were expressed on a liveweight basis and were determined as follows. For pork, the levels of the selling activity P_1 were used directly. For beef, the quantity produced was defined as net beef produced on the farm and computed as the level of the selling activity P_2 plus the level of activity P_6 at 430 pounds per head minus the level of activity P_3 at 440 pounds per head and minus the level of activities P_4 and P_5 at 715 pounds per head. When all 16 price combinations were programmed, these quantities allowed determination of four discrete points on each of four supply functions for pork and four supply functions for beef.

The next section of the model, comprised of activities P_7 through P_{17}, contains the alternative crop rotations considered. These were divided into three groups on the basis of the three qualities of cropland considered; the more intensive rotations were limited to the higher quality land. Each activity was expressed on an acre basis -- for example, the $CSOM_1$ activity has coefficients which reflect a rotation consisting of 1/4 acre each of corn, soybeans, oats and meadow on the highest quality land.

Activities P_{18} through P_{21} allow for the consideration of alternative harvesting methods. P_{18} and P_{19} allow harvesting corn produced and accumulated in row 6 either as grain or as silage. P_{20} and P_{21} allow harvesting cropland meadow (row 5) either as hay or transferring it into row 4 where it may be utilized as pasture by the livestock enterprises.

The eight hog enterprises considered are activities P_{22} through P_{29}. These consist of selected combinations of litters per year and feeding and farrowing facilities. The unit of these activities is the size required to utilize one unit of farrowing capacity at any one time. Thus, P_{28}, with three sows each producing two litters per year, provides for six non-overlapping farrowings and requires the same farrowing capacity as P_{24}, with one sow and two litters. The choice between the eight activities is primarily one of (1) type of hog facility to use and (2) intensity of use for the central farrowing facilities.

Activities P_{30} through P_{49} are the beef feeding enterprises considered. These alternative enterprises involve (1) feeding hay or feeding a combination of hay and silage, (2) use of a low or a high level of mechanization in feeding, (3) feeding calves or yearlings, (4) feeding the calves on drylot or on drylot with summer pasture and (5) the choice of two different periods of the year for yearling feeding. The information given in Table 7 is self-explanatory with the possible exception of the two-period yearling feeding enterprises. These were developed because of the lower per head annual capital requirements when yearlings are fed in two non-overlapping feeding periods.

The beef cow enterprises are activities P_{50} and P_{51} -- the first with hay as the only roughage and the second utilizing a combination of hay and silage. These activities include all of the requirements and returns of the cow and calf up to weaning time plus the cow's share of replacement heifers and bull cost.

The chattel credit borrowing activity is P_{52}. It requires chattel credit capacity or collateral from row 24 and makes the money available for expenses in the model in row 23 at a cost of 7 percent annual interest.

Activity P_{53} allows investment of unused operating capital off the farm at 5 percent annual interest.

A reservation price on operator and family labor of \$.50 per hour is included as activity P_{54}.

Activities P_{55} and P_{56} are the corn selling and buying activities. The selling price of corn was \$.85 per bushel and the purchase price was \$1.00 per bushel.

Investment in additional hog and beef facilities is allowed by activities P_{57} through P_{65}. Annual costs are included in the functional (C_j) and the capital for the investment is drawn from available or borrowed operating capital.

Activities P_{66} through P_{73} allow labor hiring in each of the eight labor periods. As was previously stated, the total labor hired in these activities was limited by row 36 to the historical labor hiring practices of the particular representative farm.

The coefficients

Appendix Table 1 presents the actual coefficients used in the linear programming model for the medium sized representative farm in area 9. The coefficients in this model have been divided into four groups: (1) values with an "a" superscript which change with the different pork and beef prices as discussed earlier; (2) values with a "b" superscript which represent crop yields and costs that are different in each of the ten soil association areas of the state -- these values vary due to differences in crop yields and fertilization rates; (3) coefficients which are different on each of the three sizes of farms and have a "c" superscript -- these are primarily costs, capital requirements and labor requirements; and (4) the remaining unsubscripted coefficients which are the same on all farms programmed. Thus, for the basic group of 36 representative farms, a coefficient with a "bc" superscript may take on 30 possible values depending on the size and the area in which the representative farm is located. A coefficient with a "c" superscript may take on three possible values depending on the size of farm.

The coefficients shown in Appendix Table I were not averaged in the models used for the three smaller subgroups of representative farms. All representative farms in the ten farm subgroup were given coefficients for the medium sized farms in the respective areas. A separate set of state average yields and fertilizer costs was developed for the three farm subgroup and the one farm subgroup. The three farm subgroup used coefficients for the three respective sizes of farms and the one farm subgroup used coefficients for the medium sized farms.

The complete coefficient matrix presented in Appendix Table 1 contains 2,628 nonzero elements -- a density of about 27 percent. Equation 3.12 and Table 1, page 49, suggest that nearly every individual farm in the population must be programmed to assure achieving exact aggregation for all activity levels with a coefficient matrix of this size and density. This relation emphasizes the difficulty of obtaining exact aggregation in the estimates of a research project of typical scope. The same difficulty is, of course, encountered in obtaining an exact measurement of aggregation error in such estimates.

This section has considered the size, density and complexity of the matrix -- factors that have been found to contribute to aggregation error. The assumptions behind the individual coefficients and their values were the same as those used for the NC-54 regional project work (38, 39). The reader should refer to these sources for the basic information. Generally the individual values of the coefficients may affect the location of the estimated supply curves but should have little effect on the findings of this study -- that is, the values should not greatly affect the amount of aggregation error among the four estimated population supply functions.

The Resulting Supply Functions

Optimum linear programming solutions were obtained for all 50 of the defined representative farms under each of the 16 price combinations previously discussed. These representative farm optimum solutions were then used to obtain four different population supply estimates using the aggregation coefficients presented in Tables 4 and 5. The optimal solutions for the basic group of 36 representative farms were aggregated to obtain one set of population supply estimates -- the same procedure was repeated using the optimal solutions for each of the three smaller groups of representative farms.

The four sets of estimated supply functions for beef and pork are presented in Figures 7 and 8. Four functions are given on each graph because changes in the price of pork cause shifts in the beef supply functions; similarly, changes in the price of beef cause shifts in the pork supply functions. These shifts are a result of the familiar cross-elasticity relationships of supply. Each of the supply functions is drawn from estimation of four discrete points.[1] The numerical production estimates used as a basis for drawing Figures 7 and 8 are presented in Tables 8 and 9.

A detailed analysis of these supply estimates and the amount of aggregation error between them will be made in the next chapter. However, at this stage it is interesting to observe the similarity among the four sets of supply functions obtained using the different numbers of representative farms. Aside from a few points on the three and one representative farm estimates, the supply curves have much in common. Both the location and the elasticities appear to be similar. The differences

---

[1]This is in contrast to the more usual "stepped" supply functions estimated by linear programming which result when the price is continuously varied within a given range. In the two-product case, varying two prices continuously within even a small range results in a multitude of different solutions -- for this reason only 16 discrete price combinations were programmed.

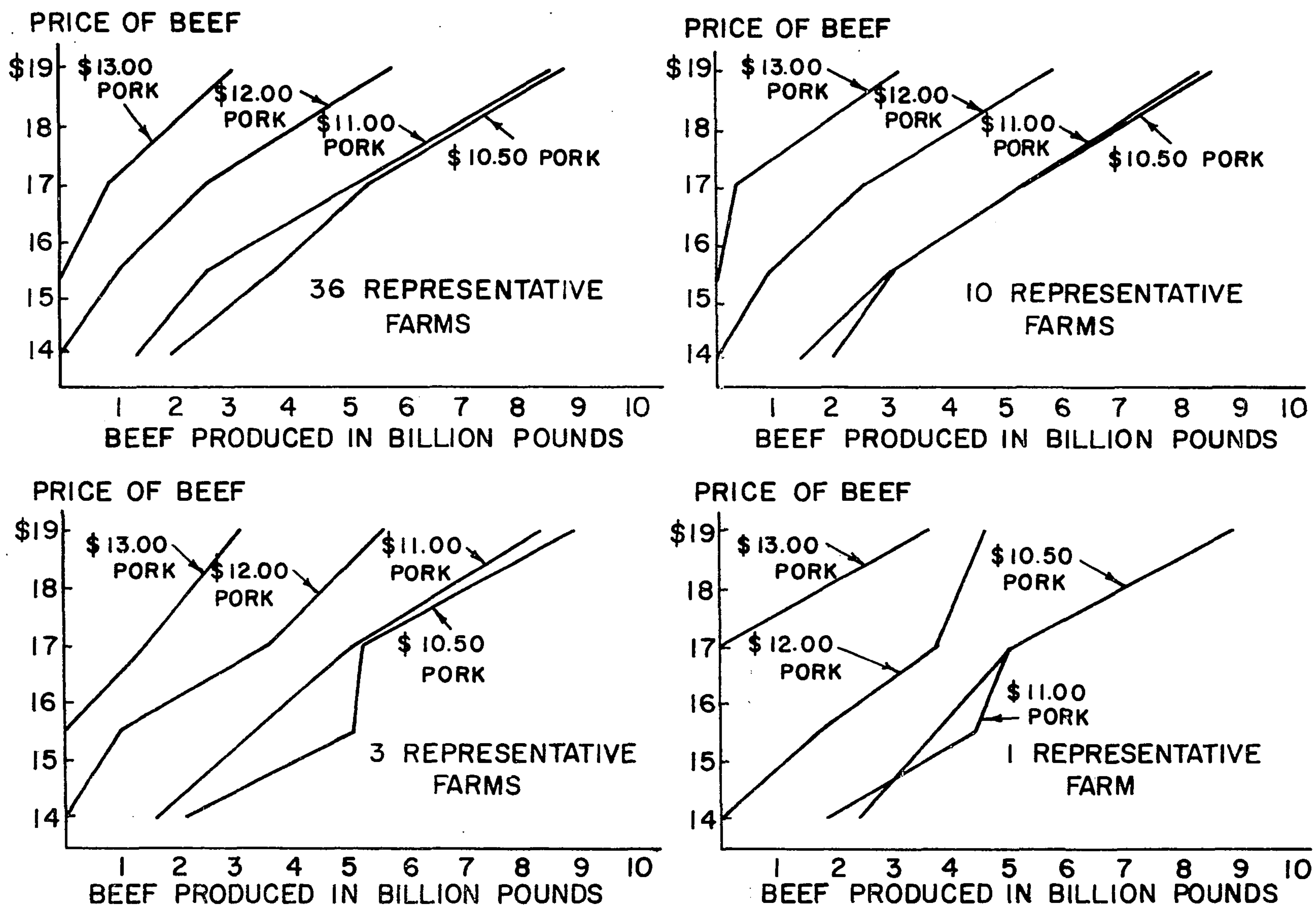


Figure 7. Iowa beef supply functions estimated from four groups of representative farms

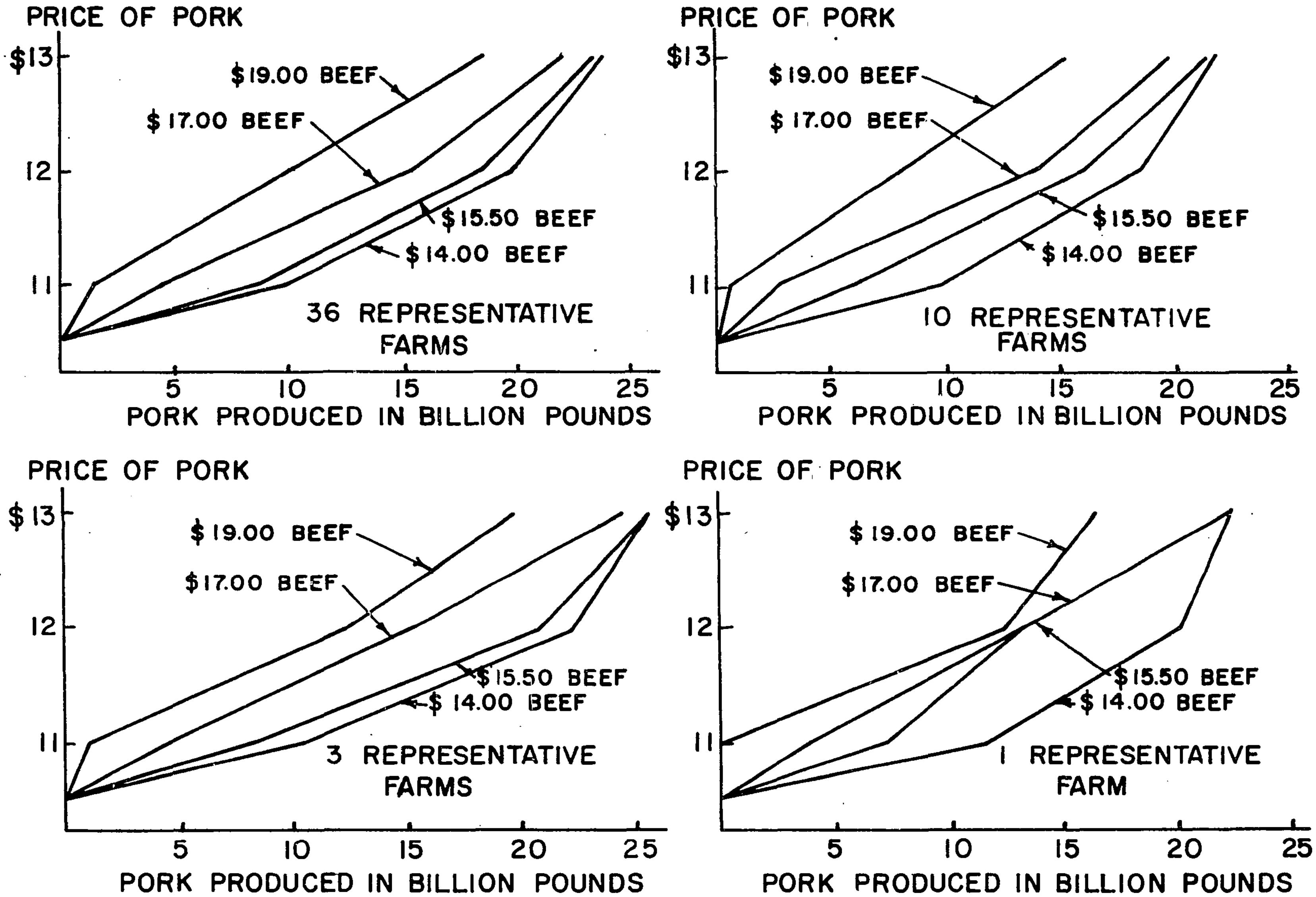


Figure 8. Iowa pork supply functions estimated from four groups of representative farms

Table 8. Production of Iowa beef under different price combinations estimated from four groups of representative farms

Pork price	Beef price	Number of representative farms used			
		36	10	3	1
		(million pounds)			
$10.50	$14.00	1,909	2,028	2,171	2,354
10.50	15.50	3,762	3,030	5,093	3,750
10.50	17.00	5,454	5,405	5,213	5,093
10.50	19.00	8,802	8,508	8,939	9,012
11.00	14.00	1,316	1,508	1,686	1,766
11.00	15.50	2,542	2,978	3,335	4,431
11.00	17.00	5,343	5,375	5,103	5,093
11.00	19.00	8,625	8,461	8,439	9,012
12.00	14.00	195	57	--	--
12.00	15.50	1,026	936	981	1,725
12.00	17.00	2,566	2,626	3,613	3,712
12.00	19.00	5,752	5,855	5,667	4,638
13.00	14.00	54	--	--	--
13.00	15.50	155	18	--	--
13.00	17.00	845	365	1,458	--
13.00	19.00	3,110	3,247	3,130	3,589

Table 9. Production of Iowa pork under different price combinations estimated from four groups of representative farms

Pork price	Beef price	Number of representative farms used 36	10	3	1
		(million pounds)			
$10.50	$14.00	344	--	--	--
10.50	15.50	--	--	--	--
10.50	17.00	--	--	--	--
10.50	19.00	--	--	--	--
11.00	14.00	9,923	9,659	11,041	10,726
11.00	15.50	8,473	5,817	8,461	6,884
11.00	17.00	4,623	2,837	4,496	3,769
11.00	19.00	1,567	413	1,043	--
12.00	14.00	20,055	18,277	22,118	19,981
12.00	15.50	18,365	15,838	20,942	13,235
12.00	17.00	15,577	14,220	15,539	13,398
12.00	19.00	10,269	8,387	12,285	12,535
13.00	14.00	23,743	20,824	26,119	22,520
13.00	15.50	23,582	20,803	26,119	22,520
13.00	17.00	22,057	19,672	24,301	22,520
13.00	19.00	18,634	15,294	19,868	16,560

are quite small when considered in terms of the differences in costs involved in the different estimates.

The costs of the four different estimates were roughly proportional to the number of representative farms involved. Programming and aggregation costs on the IBM 7074 computer for this study were about $41 per farm. Thus, there was a computing cost difference of about $1,066 between the 36 and 10 representative farm estimates and about $369 between the 10 and 1 representative farm estimates. Including other normal research costs would naturally raise these figures substantially.

CHAPTER VI. EVALUATION OF PROGRAMMED SUPPLY FUNCTIONS

The central objective of the empirical programming work was to determine the amount of aggregation error in Iowa beef and pork supply estimates. The beef and pork supply aggregation errors are evaluated in this chapter and the state estimates for other major farm products derived from the four groups of representative farms are also presented and appraised. Representative farm optimum solutions for one soil-type area of the state are presented in detail to provide an understanding of the complexity of the relation between representative farm data, optimum solutions and aggregation error.

State Supply Estimates

Differences among beef and pork supply estimates

Figures 7 and 8 and the data in Tables 8 and 9 of Chapter V provide bases for a more detailed analysis of the aggregation error among the different estimated beef and pork supply functions for the state. As was discussed in the last chapter, the differences among the state estimates based on the four groups of representative farms are due to aggregation error.

This aggregation error could be analyzed at all of the 16 price combinations programmed. However, all of the 16 sets of results are not equally realistic by real world standards. Price combinations resulting in Iowa beef and pork production in the neighborhood of present actual production levels appear to be the most realistic because such production levels are most consistent with the input prices, market structures and input-output relations built into the linear programming models (40). In addition, production levels in this neighborhood and their associated prices are good estimates of equilibrium market conditions under the assumptions of the model. For these reasons solutions approximating historical production levels are given major emphasis in the analysis of the aggregation error.

In 1965, Iowa produced 2.8 billion pounds of beef and 4.4 billion pounds of pork (50, pp. 30, 34). Reference to Figure 8 reveals that this level of pork production is most consistent with an $11.00 pork price. With a pork price of $11.00, reference to the beef supply curves on Figure 7 shows that current levels of beef production are achieved in the model at approximately a $15.50 beef price. This $15.50 beef price is also consistent with the $11.00 pork price in obtaining near current levels of pork production in Figure 8.

Beef supply functions estimated with the price of pork held at $11.00 and the pork supply functions estimated with the price of beef held at $15.50 thus assume relatively more importance. The analysis of aggregation error in this chapter is based primarily on these supply functions. The four such functions for beef are superimposed in Figure 9. In a similar manner, the four pork supply functions estimated with a beef price of $15.50 are superimposed in Figure 10.

In Figure 9, the four Iowa beef supply functions estimated by the four different groups of representative farms are very similar. Except for the one representative farm estimate being relatively higher at the

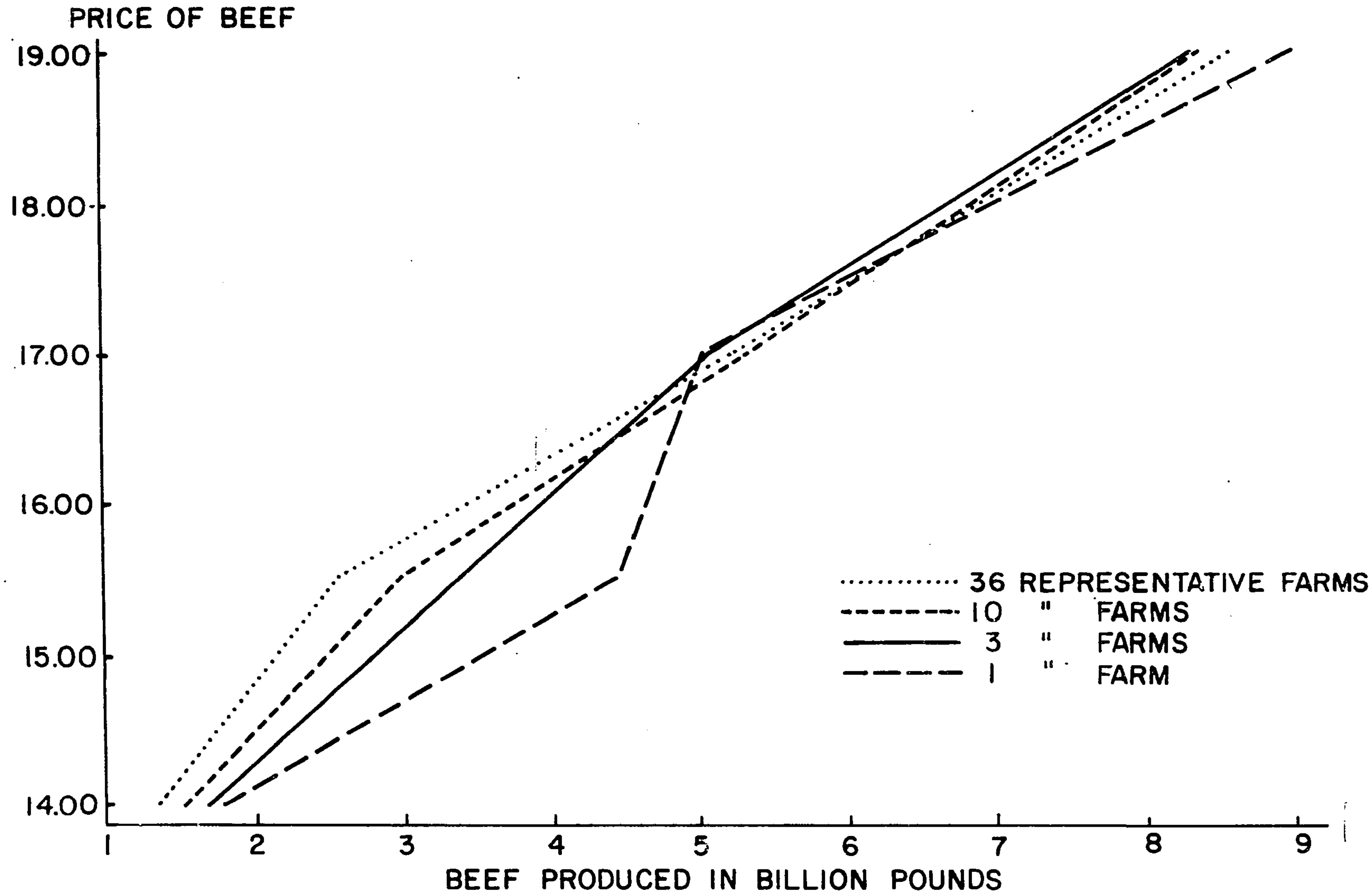


Figure 9. Aggregation error in Iowa beef supply functions at the $11.00 pork price

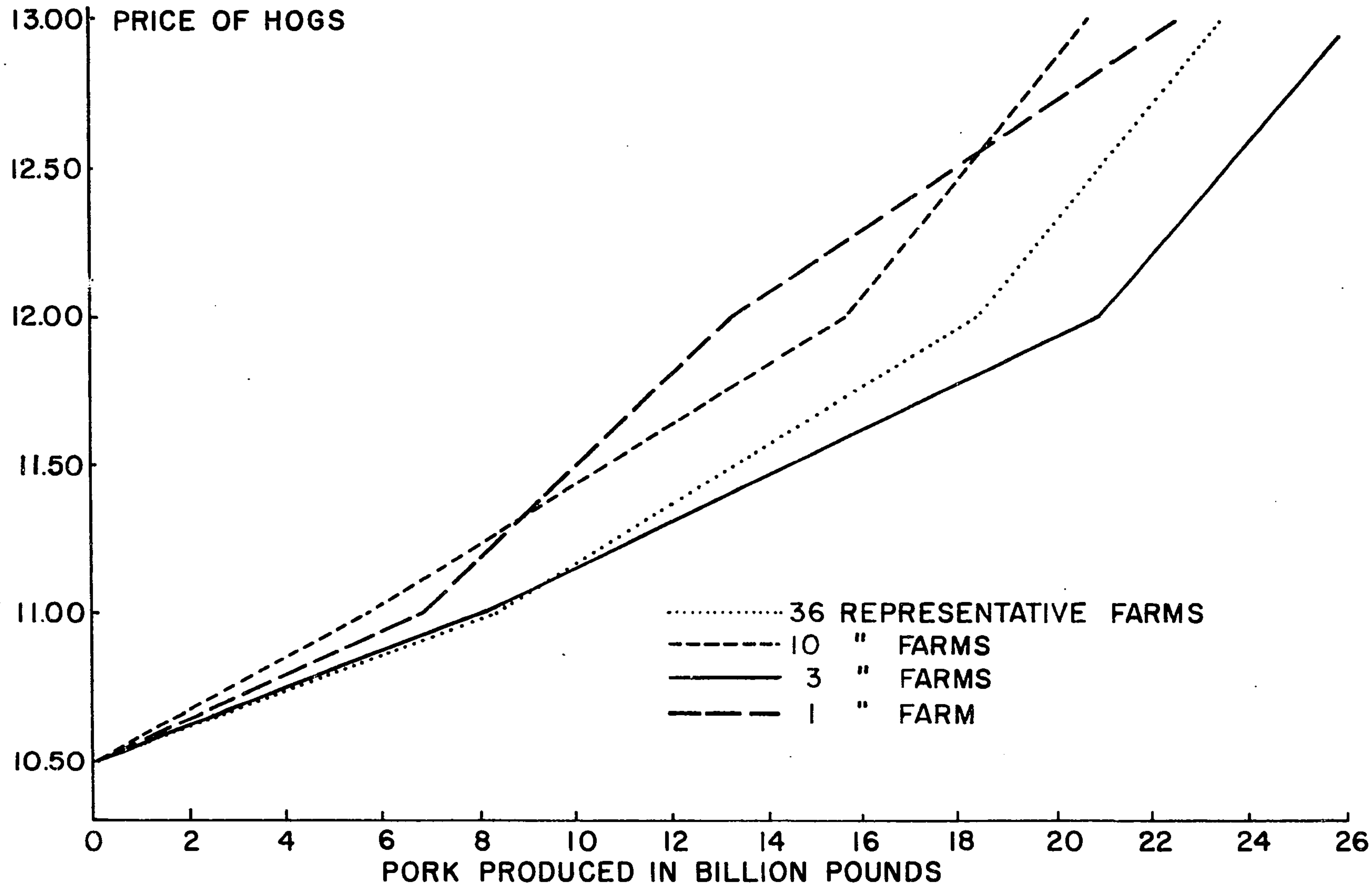


Figure 10. Aggregation error in Iowa pork supply functions at the $15.50 beef price

$15.50 price, the remainder of the estimated points are quite similar. The over-all slopes of the functions are in agreement and differences between them are small -- in fact these differences virtually disappear at the $17.00 and $19.00 beef price levels. As was mentioned in the last chapter, these functions are strikingly uniform considering the large differences in the underlying research costs involved.

In Figure 10, the four Iowa pork supply functions exhibit somewhat larger differences than are true for the beef supply functions, especially at the two higher pork prices. Pork production at these higher prices is from three to five times greater than current production levels. The linear programming models would be expected to provide a rather poor simulation of actual farm conditions at these levels of production (40). As a result, the aggregation error at these levels of production is probably increased by the specification problem involved. In the neighborhood of historical production levels, the four estimated pork supply functions are much closer, at least in an absolute sense.

The four programmed supply functions for both pork and beef are in agreement with the hypothesized lack of consistent direction of the aggregation error discussed in Chapter IV. Figures 9 and 10 reveal no significant direction of error as fewer representative farms are used in the supply estimation process. In Table 8, the beef production estimates based on the three subgroups of representative farms are less than the 36 farm estimate 25 times and greater 23 times. In a similar manner, the pork production estimates in Table 9 based on the three subgroups of representative farms are less than the 36 farm estimate 28 times and greater 11 times. Considering both beef and pork, the aggregation error is positive 34 times and negative 53 times. These results appear consistent with the hypothesis of lack of bias in aggregation error. If anything, the slight excess of negative error in pork estimates tends to refute the historical notion of a tendency toward a positive bias.

Number of representative farms and amount of error

A comparison was made of the relative amounts of aggregation error among the four different supply estimates presented in Figures 9 and 10. Table 10 presents an index of the absolute error with the 36 representative farm estimate used as the base of 100. In other words, the ten farm beef supply estimate at the $19.00 beef price has an absolute error index of 101.9 -- this value indicates it has a 1.9 percent difference from the 36 farm estimate at that price. The average error along the beef supply function estimated by the ten representative farms is 8.6 percent as shown in Table 10. Eleven of the 24 individual index values for beef and pork in the table are below 105 and 18 are below 120.

An understanding of the effect of using different numbers of representative farms in the estimation process is provided by Figure 11, where the Table 10 averages are plotted. For beef, the error decreases as larger numbers of representative farms are programmed, with the marginal contribution of each representative farm declining as more representative farms are included in the model. This decline suggests that both of the factors used in the more detailed stratifications,

Table 10. Index of absolute aggregation error of points along state beef and pork supply functions estimated by three subgroups of representative farms (36 farm estimate = 100)

	Price	Number of representative farms 10	3	1
Beef supply estimates ($11.00 pork price)				
	14.00	114.6	128.1	134.2
	15.50	117.2	131.2	174.3
	17.00	100.6	104.5	104.7
	19.00	101.9	102.2	104.5
Average absolute error		108.6	116.5	129.4
Pork supply estimates ($15.50 beef price)				
	10.50	100.0	100.0	100.0
	11.00	131.3	100.1	118.8
	12.00	113.8	114.0	127.9
	13.00	111.8	110.8	104.5
Average absolute error		114.2	106.2	112.8

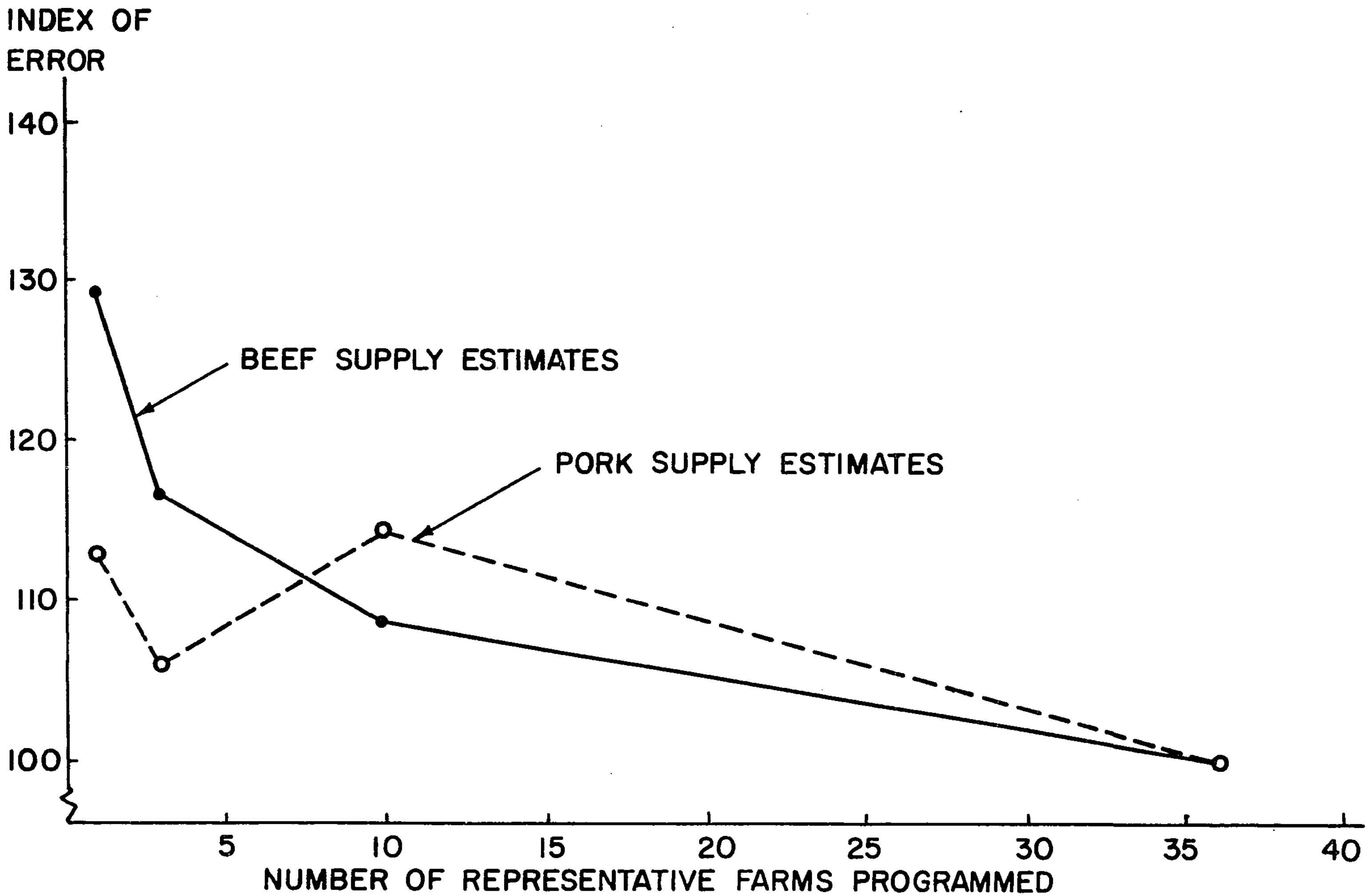


Figure 11. Effect of programming different numbers of representative farms on aggregation error in state beef and pork supply estimates

soil-type area and size of farm, were useful in obtaining accuracy in the beef supply estimates.

For the pork supply estimates, the error does not always decrease as more representative farms are used -- the three farm estimated supply function has less error than the ten farm estimate. Since the three representative farms resulted from a size of farm stratification and the ten representative farms were based on soil-type areas, the results suggest that size of farm is a much more important factor in influencing pork production than is soil type area. A review of the individual farm solutions substantiates the hypothesis. Size of farm is quite important in influencing whether or not a farm produces pork -- area of the state is relatively less important. Thus, it appears that omitting the size of farm classification results in more error in the pork estimates when ten representative farms were used, even though the number of representative farms programmed is increased from three to ten.

The relationships of Figure 11 substantiate the theory developed in Chapter III: (1) that more detailed stratification increases accuracy only when it recognizes factors that influence the existence of certain enterprises on the farms and (2) that different stratification factors may be required in the estimation of different products. When a factor such as soil type area has only a small effect on the existence of pork production enterprises on the farms, additional stratification based on it provides relatively small gains in accuracy. Likewise, overlooking a factor that does affect the existence of the hog enterprise on the farms results in a substantial increase in error. Finally, a factor that is useful for controlling error in the estimation of one product may not have the same effect on estimates of another product.

Error in elasticity estimates

Price elasticities of supply are often computed for supply functions as a measure of the degree of responsiveness of output to changes in product prices. Many questions of agricultural policy have answers depending on the price elasticities of supply for the products in question. In addition, the representative farm linear programming technique may provide more accurate estimates of the elasticity of supply than of the actual level of supply -- for example, overestimating all resources on a farm would affect the level of the supply estimates but not the elasticity. Thus, elasticity estimates may have usefulness even when the actual level of the estimated supplies is inaccurate.

The estimated elasticities and cross-elasticities of supply between the prices of beef and pork and the quantities of beef and pork produced are presented in Tables 11 and 12. These elasticities are computed by the equation

$$e = \frac{Q_2 - Q_1}{Q_2 + Q_1} \bigg/ \frac{P_2 - P_1}{P_2 + P_1}$$

and represent the percent change in quantity that results from a 1 percent change in price. They are arc elasticities and represent the average elasticity between the two points programmed on the supply functions. Four estimates are available for each elasticity figure, one from each of the four sets of programmed supply functions.

Table 11. Elasticities and cross-elasticities between price of beef and quantities of beef and pork estimated by four groups of representative farms

		Percent change in:							
		Beef production				Pork production			
Pork price	Arc range of beef price	36 farms	10 farms	3 farms	1 farm	36 farms	10 farms	3 farms	1 farm
$10.50									
	$14.00-15.50	6.43	3.90	7.91	4.50	--	--	--	--
	15.50-17.00	3.98	6.10	0.25	3.29	--	--	--	--
	17.00-19.00	4.23	4.01	4.74	5.00	--	--	--	--
$11.00									
	$14.00-15.50	6.25	6.44	6.46	8.46	-1.55	-4.88	-2.60	-4.29
	15.50-17.00	7.70	6.22	4.54	1.51	-6.35	-7.46	-6.63	-6.34
	17.00-19.00	4.23	4.01	4.44	5.00	-8.89	-13.43	-11.22	0
$12.00									
	$14.00-15.50	13.38	17.41	19.67	19.67	-0.87	-1.41	-0.54	-3.99
	15.50-17.00	9.29	10.28	12.41	7.92	-1.78	-1.17	-3.21	0.13
	17.00-19.00	6.89	6.85	3.98	2.00	-3.70	-4.64	-2.11	-0.60
$13.00									
	$14.00-15.50	9.46	19.67	--	--	-0.07	-0.01	0	0
	15.50-17.00	14.99	19.63	21.67	--	-0.72	-0.61	-0.78	0
	17.00-19.00	10.31	14.36	6.56	18.00	-1.51	-2.25	-1.81	-2.75

Table 12. Elasticities and cross-elasticities between price of pork and quantities of pork and beef estimated by four groups of representative farms

Beef price	Arc range of pork price	Percent change in: Pork production				Beef production			
		36 farms	10 farms	3 farms	1 farm	36 farms	10 farms	3 farms	1 farm
$14.00									
	$10.50-11.00	40.12	43.00	43.00	43.00	-7.91	-6.32	-5.41	-6.14
	11.00-12.00	7.77	7.10	7.68	6.93	-17.06	-21.32	-23.00	-23.00
	12.00-13.00	2.11	1.63	2.07	1.49	-14.16	-25.00	--	--
$15.50									
	$10.50-11.00	43.00	43.00	43.00	43.00	-8.32	-0.37	-8.97	3.58
	11.00-12.00	8.48	10.64	9.76	7.26	-9.77	-12.00	-12.54	-10.11
	12.00-13.00	3.11	3.39	2.75	6.49	-18.47	-24.06	-25.00	-25.00
$17.00									
	$10.50-11.00	43.00	43.00	43.00	43.00	-0.44	-0.12	-0.46	0
	11.00-12.00	12.47	15.35	12.68	12.90	-8.08	-7.90	-3.93	-3.61
	12.00-13.00	4.30	4.02	5.50	6.35	-12.61	-18.90	-10.62	-25.00
$19.00									
	$10.50-11.00	43.00	43.00	43.00	43.00	-0.45	-0.12	-1.24	0
	11.00-12.00	16.91	20.85	19.40	23.00	-4.60	-4.19	-4.52	-7.37
	12.00-13.00	7.24	7.29	5.90	3.46	-7.45	-7.16	-7.21	-3.19

A few generalizations can be made about these estimated price elasticities of supply. In Table 11, which expresses the effect of a change in the price of beef, the 36 and 10 farm estimates generally agree. This comparability is especially high at the $11.00 pork price. In Table 12, the effect of a change in the price of pork is generally quite similar as estimated by each of the four groups of representative farms. Even the estimates developed by programming one representative farm provide a reasonably accurate estimate of the changes in the quantities of pork and beef produced that result from changes in pork prices.

Error in other major estimates

The severity of the aggregation error problem and the number and type of representative farms required are extremely dependent upon the particular estimate desired of the model. Table 13 shows the amount of aggregation error present in various state estimates developed using the four groups of representative farms. The percent errors in beef and pork production estimates are the same as the ones found in Table 10 for the appropriate price combination. Other estimates shown in the table are hog and cattle numbers in the state, major crop acreages and production, hired labor use and the maximum net returns (linear programming functional value) for the state.

Table 13 shows a wide range in the amount of aggregation error in the different estimates. Estimated net feeder calf purchases and net corn sales have the highest amounts of aggregation error. At the other extreme are the extremely low errors in estimates of corn and

Table 13. Comparison of major state estimates developed using four groups of representative farms ($15.50 beef price and $11.00 pork price)

		Number of representative farms						
		36	10		3		1	
State estimate	Unit	Estimate	Estimate	Percent error	Estimate	Percent error	Estimate	Percent error
Litters farrowed	1,000 lit.	4,475.2	3,063.7	-31.5	4,456.4	-0.4	3,614.9	-19.2
Total pork produced	mil. lbs.	8,472.8	5,817.0	-31.3	8,461.3	-0.1	6,884.2	-18.8
Beef cows	1,000 head	1,273.3	1,382.7	+8.6	865.4	-32.7	37.9	-97.0
Net feeder calves purchased	1,000 head	2,009.1	2,510.8	+25.0	3,800.8	+89.2	6,658.9	+231.4
Calves fed	1,000 head	3,015.1	3,603.2	+19.5	4,484.5	+48.7	6,688.8	+121.8
Total beef produced	mil. lbs.	2,542.1	2,977.7	+17.2	3,335.0	+31.2	4,431.1	+74.3
Corn acreage	1,000 ac.	9,007.1	8,935.9	-0.8	8,748.5	-2.9	8,748.5	-2.9
Corn production	mil. bu.	772.2	771.8	-0.1	755.8	-2.1	755.8	-2.1
Net corn sold in state	mil. bu.	123.5	246.4	+99.5	24.3	-80.3	--	-100.0
Oat acreage	1,000 ac.	1,218.8	1,054.3	-13.5	997.6	-18.1	997.6	-18.1
Oat production (in corn equivalents)	mil. bu.	32.3	28.8	-10.8	27.5	-14.9	27.5	-14.9
Soybean acreage	1,000 ac.	9,702.1	9,615.3	-0.9	9,746.1	+0.5	9,746.1	+0.5
Soybean production	mil. bu.	318.9	316.0	-0.9	320.6	+0.5	320.6	+0.5
Meadow acres (cropland)	1,000 ac.	3,674.8	3,997.1	+8.8	4,110.5	+11.9	4,110.5	+11.9
Hay production	1,000 tons	3,930.1	4,488.2	+15.0	4,302.7	+10.2	4,538.4	+16.3
Total labor hired	1,000 hrs.	12,032.9	8,040.0	-33.2	9,346.3	-22.3	6,856.5	-43.0
Maximum net returns	mil. dol.	930,755	853,773	-8.3	944,389	+1.5	868,282	-6.7

soybean acreage and production. It would be difficult to state a stronger case for delineating different numbers of representative farms to use in making different estimates. If an estimate of net feeder calf purchases for the state is desired, a considerable number of representative farms of the types developed in this study would be required. On the other hand, if the objective is to estimate corn and soybean acreages, much smaller numbers of representative farms would suffice. In fact, one representative farm for the entire state does an effective job of estimating major crop acreages with an aggregation error of only 0.5 percent for soybeans and 2.9 percent for corn and with an absolute error of less than 259,000 acres for corn and less than 436,000 acres for meadow.

This finding has relatively important implications for models designed solely for estimating Iowa crop acreages and production when there is no desire for livestock production estimates. Aggregation error appears to be much less important in crop production estimates and may be held to allowable levels with small numbers of representative farms. The distribution of livestock production on different farms in the state has a relatively minor effect on profitable cropping systems. The choice of optimal cropping systems is somewhat independent of desired or profitable combinations of livestock enterprises. Rather, cropping systems are dependent on land capabilities and agronomic restraints, factors that can be handled quite well with even one representative farm.[1]

---

[1]Remember there is no need to define an additional representative farm solely to recognize a different ratio of resources. For example, different ratios of 50 percent vs. 100 percent capability land may be covered by one representative farm as long as the same two rotations are always optimal on each of the respective classes of land over the entire range of the ratio.

It is interesting to note that the value of the optimal functional (maximum net returns) for the state from the ten and one farm groups is less than the optimal functional estimated by the 36 representative farms, as shown on the last line of Table 13. On the surface, this may appear a contradiction of Theorem III in Chapter IV. However, Theorem III specifies farms with identical cost and coefficient matrices, a requirement not met by the groups of farms programmed for developing the state estimates. Thus, the estimate of the functional as shown in Table 13 may have negative error rather than the positive error implied by Theorem III. There is no inconsistency in these results.

Area 3 in Detail

This section analyzes in more detail the aggregation error found at one price combination in one of the ten soil-type areas in an attempt to ascertain (1) the individual farm characteristics that lead to aggregation error and (2) the interrelationships among the errors in the various estimates that may be derived from the programming results. The analysis is based on comparison of the optimal solutions of farms 3A, 3B and 3C with the optimal solutions of farm 3BB. These programming results may be aggregated into two estimates of area 3 production. The differences between these two area production estimates are due to aggregation error in the same manner as differences between the various state estimates are due to aggregation error.

Characteristics of area 3

Soil-type area 3, composed of ten counties in southwestern Iowa (see Figure 6, p. 94) was chosen for this detailed analysis because the relations and solutions in this area appeared less complex than those in other areas -- at the same time data and results of this area portrayed many of the characteristics found throughout the state. Only one solution was analyzed in detail in area 3. This was the solution obtained from the $11.00 pork and the $15.50 beef price combination. As stated previously, this price combination yielded the most reasonable results of any of the 16 combinations programmed.

For the 36 representative farm group, three farms, 3A, 3B and 3C, were used to represent area 3. These represented the small, medium and large farms in the area. For the ten representative farm group, one farm, 3BB, was programmed for this area. As was discussed in Chapter V, the resources for farm 3BB were the weighted averages of resources available on farms 3A, 3B and 3C. The rest of the net return and input-output coefficients for farm 3BB were the same as the coefficients for farm 3B, since both farms were in the medium size range.

The first step in analyzing the aggregation error in area 3 is to review the four optimum representative farm linear programming solutions in some detail. Table 14 shows amounts of resources available on each of these four representative farms and identifies resources that are restrictive under the optimal solutions. All livestock facility resources are included in the table because of their primary importance in determining the optimum solutions at the $11.00 pork and $15.50 beef price combination. Other resources that are not restrictive and

Table 14. Available resources and the extent of their use on area 3 representative farms ($15.50 beef price and $11.00 hog price)

Resource[a]	Unit	Representative farm number							
		3A		3B		3C		3BB	
Number of farms represented	No.	2,775		5,118		5,731		13,624	
Land with 25% row crop capability	acre	6.9	R[b]	14.0	R	28.5	R	18.7	R
Land with 50% row crop capability	"	34.6	R	69.9	R	142.4	R	93.2	R
Land with 100% row crop capability	"	32.3	R	65.1	R	132.7	R	86.9	R
Total cropland	"	73.8		149.0		303.6		198.8	
Pasture (noncropland)	ton AHY[c]	15.1	R	19.7	R	54.4	R	33.4	R
Central farrowing facilities	sow	--		9.6	-	12.1	-	8.7	-
Portable farrowing facilities	"	14.1	-	1.8	R	9.4	R	7.5	R
Confinement feeding facilities	pig	--		--	-	--	-	--	-
Portable feeding facilities	"	92.6	-	156.7	-	225.7	R	172.7	-
Beef housing - both periods	a.u.[d]	23.1	-	28.6	-	39.8	R	32.2	R
Low beef mechanization - both periods	head	48.3	-	11.4	-	19.2	-	22.2	-
High beef mechanization - both periods	"	--	-	89.5	-	127.9	-	87.4	-
April operator and family labor	hour	227.7	-	246.6	R	285.4	P	259.1	P
October operator and family labor	"	227.7	-	246.6	-	285.4	P	259.1	-
November operator and family labor	"	202.7	-	221.6	-	255.6	P	232.1	-

[a]Resources not listed had no effect on determination of optimal solutions.

[b]R = restrictive in optimal solution.
P = restrictive in optimal solution and additional quantities purchased or hired.

[c]Tons of anticipated hay yield.

[d]Animal units.

play no part in influencing the final solutions are not shown on the table. Restrictive resources at this price combination are cropland, pasture (noncropland), selected hog and beef facilities and April, October and November operator and family labor, as shown in Table 14. Generally, more resources become restrictive as farm size increases, with farm 3A having only four restrictive resources and farm 3C having ten restrictive resources. This increase in number of restrictive resources is a result of more complex solutions including greater numbers of activities on the larger farms.

A summary of the optimum solutions for each of the four representative farms in area 3 is presented in Table 15. The optimum solution for farm 3A includes two crop rotations (corn-soybeans and CSSOMM), meadow on the low capability cropland, feeding 26.8 purchased calves to utilize the pasture available and selling the surplus corn produced.

The solution for the medium size farm, farm 3B, has the same cropping pattern as farm 3A. However, this solution also includes one and four litter hog systems and beef cows. These three enterprises and additional purchased feeder calves are combined to utilize all of the corn, pasture, portable farrowing facilities and April operator and family labor available on the farm. No corn is sold and no labor is hired. Reference to Table 14 shows that none of the beef housing or beef feeding facilities are fully utilized by this combination of enterprises.

The solution for farm 3C in Table 15 is more complex than the other solutions. One additional rotation, COMM, is added to the crop system on part of the low capability cropland. One, two and four litter

Table 15. Optimum solutions on area 3 representative farms ($15.50 beef and $11.00 hogs)

	Unit	Representative farm number			
		3A	3B	3C	3BB
Meadow on 25% row crop capability land	acre	6.9	14.0	6.4	18.7
COMM rotation on 25% row crop capability land	"	--	--	22.1	--
CSSOMM rotation on 50% row crop capability land	"	34.6	69.9	142.4	93.2
Corn-soybean rotation on 100% row crop capability land	"	32.3	65.1	132.7	86.9
Harvest corn as grain	bu.	1,615	3,260	6,730	4,348.8
Sell corn not fed	"	238.5	--	--	--
Hay meadow	ton	18.0	26.5	53	38.0
1 litter sow system (portable farrow and feed)	litter	--	1.8	3.9	7.5
2 litter sow system (portable farrow and feed)	"	--	--	11.1	--
4 litter sow system (central farrow and portable feed)	"	--	21.0	37.6	16.5
Pork produced	lbs.	--	43,100	99,400	45,800
Calves on pasture (low mechanization feeding)	head	26.8	--	--	--
Calves on pasture (high mechanization feeding)	"	--	15.5	20.8	33.7
Beef cows	"	--	10.7	26.3	10.3
Purchase beef calves	head	26.8	7.1	--	25.6
Beef produced	lbs.	17,700	14,900	25,100	26,700
Hire April labor	hour	--	--	130.5	56.2
Hire October labor	"	--	--	76.7	--
Hire November labor	"	--	--	14.3	--
Cash invested off farm	dol.	7,068.9	6,625.2	15,585	9,693

hog systems, beef cows and a feeding enterprise for farm raised beef calves comprise the livestock enterprises. These enterprises utilize all of the corn, pasture, portable hog farrowing and feeding facilities and beef housing available on the farm. No corn is sold and additional labor is hired for the months of April, October and November.

The average representative farm for area 3, farm 3BB, has a solution much like the medium sized farm, farm 3B. The same cropping pattern and livestock enterprises are present. However, beef housing is restrictive on farm 3BB and April labor is hired -- two things that did not occur on farm 3B.

Beef and pork production in the optimum solutions are made up of many related components and are determined by a number of relationships. Hence, it is desirable to look at aggregation error over the range of estimates derived from the linear programming solutions as a means of gaining an understanding of the interrelationships among various individual activities of the models. Aggregation errors in area 3 activity levels and production estimates are shown in Table 16. At the \$15.50 beef price and \$11.00 pork price, the one representative farm underestimates pork production for area 3 by 21.1 percent (line 12) and overestimates beef production by 35.1 percent (line 17). The magnitude of these errors is comparable to the size of errors found in the state supply estimates of pork and beef (see Table 10, page 125).

The levels of the corn-soybean and CSSOMM rotations are the only two activities estimated without error by farm 3BB. Since these two rotations produce all of the soybeans found in the optimal solutions,

Table 16. Aggregation error in activity levels and production estimates for area 3 ($11.00 pork price and $15.50 beef price)

Line	Activity	Unit	Three farm estimate	Farm 3BB estimate	Amount of error	Percent error
1	Meadow on 25% row crop capability land	1,000 acre	217.6	254.5	+36.9	+17.0
2	COMM rotation	"	36.9	--	-36.9	-100.0
3	CSSOMM rotation	"	1,270.0	1,270.0	--	--
4	Corn-soybean rotation	"	1,183.3	1,183.3	--	--
5	Harvest corn as grain	1,000 bu.	59,737	59,239	-498	-0.8
6	Corn sold	"	661.8	--	-661.8	-100.0
7	Soybean production[a]	"	30,344	30,344	--	--
8	Hay harvested on meadow	1,000 ton	492	518	+26	+5.3
9	1 litter sow system (portable farrow and feed)	1,000 litter	31.4	102.2	+70.8	+225.5
10	2 litter sow system (portable farrow and feed)	"	63.5	--	-63.5	-100.0
11	4 litter sow system (central farrow and portable feed)	"	322.9	224.9	-98.0	-30.3
12	Pork produced	mil. lbs.	790.5	623.6	-166.9	-21.1
13	Calves on pasture - no silage (low mechanization feeding)	1,000 head	74.5	--	-74.5	-100.0
14	Calves on pasture - no silage - (high mechanization feeding)	"	198.6	459.5	+260.9	+131.4
15	Beef cows - no silage	"	205.7	140.0	-65.7	-31.9
16	Purchase beef calves	"	110.5	348.9	+238.4	+215.7
17	Beef produced	mil. lbs.	269.4	364.0	+94.6	+35.1
18	Hire April labor	1,000 hour	748.2	765.9	+17.7	+2.4
19	Hire October labor	"	439.7	--	-439.7	-100.0
20	Hire November labor	"	81.8	--	-81.8	-100.0
21	Total labor hired[a]	"	1,269.7	765.9	-503.8	-39.7
22	Cash surplus and invested off farm	mil. dol.	142.8	132.1	-10.7	-7.5
23	Maximum net returns (functional value)[a]	"	88.1	76.1	-12.0	-13.6

[a]A function of activity levels.

soybean acreage and production are also estimated without error (line 7). Table 16 shows three items estimated without error, seven items estimated with positive error and 13 items estimated with negative error.

Relations between solutions and aggregation error

It is possible to explain the aggregation error in some of the activity levels in Table 16 by the makeup of the optimal solutions for the representative farms summarized in Table 15 and by the amounts of the resources available as shown in Table 14.

First, exact aggregation is achieved in three of the estimates listed on Table 16. Comparing the acres of the corn-soybeans and CSSOMM rotations on Table 15 with the amounts of land available in the 50 and 100 percent row crop capability classes on Table 13 shows that these two resources are each completely exhausted on all farms by the same two respective rotations. When a given resource is exhausted by a given activity on all representative farms in the three farm group, it is exhausted by that activity on the one representative farm. No error results in the estimates based on these activity levels. Hence, corn-soybeans rotation acres, CSSOMM rotation acres, and soybean acres (since they are a function only of these two activities) are estimated without error by farm 3BB.

Next, it is possible to identify some characteristics of the optimum solutions that lead to error in the estimates. In Table 15 farms 3A, 3B and 3BB use all of the 25 percent row crop capability land for meadow, whereas farm 3C uses 22.1 acres of it for a COMM rotation and only 6.4 acres of it for meadow. Hence, farm 3BB overestimates meadow and underestimates COMM rotation acres -- which in turn leads to an underestimation of corn production. This difference in land use accounts for all of the crop production aggregation errors found in Table 16.

From this point on, it becomes extremely difficult to determine the individual solution characteristics that lead to aggregation error. Within the linear programming models, subsets of activities and resources interact as simultaneous equation systems. The optimum level of the activities within these systems equals the solution to the system of simultaneous equations. Thus, several characteristics of an optimal solution interact to cause aggregation error in several activity levels. For example, it is tempting to explain error in estimated beef cow numbers in terms of an unused quantity of beef housing in either the three farm group or on farm 3BB. However, the relationship is always confused by the fact that cows are not the only activity using beef housing and that cows also combine with one or two other activities in using the remainder of the resources required by cows. This interrelationship follows automatically from the fact that the activity levels being considered are determined by a system of simultaneous equations. As a result, no clear explanation of such aggregation error can be made.

Even when it is possible to observe an optimum solution characteristic that leads to an aggregation error, it is much more difficult to determine the underlying data characteristic of the representative farm that lead to that solution. Differences in solutions are generally due to (1) differences in resource availabilities, (2) differences in costs and returns and (3) differences in input requirements. It is difficult to assign the responsibility for a particular aggregation error because these three factors vary simultaneously among the four representative farms in area 3. It may be possible to see where the error originates (for example, the corn production aggregation error) but impossible to determine what data characteristic caused it.

As a result of the complexity of the relations, the two objectives of the second section of Chapter VI are only partially achieved. The individual farm characteristics that lead to aggregation error may sometimes be defined in terms of solutions but are much more difficult to define in terms of the representative farm data. Interrelationships among the various aggregation errors in the model are not always visible because the individual activity levels themselves are determined by a rather complex process. The detailed analysis of area 3 serves mainly to show the complexity of the aggregation error problem in a typical representative farm linear programming model.

As was discussed in Chapter V, the coefficient matrices of the representative farms in the three subgroups were not averages of coefficient matrices of the original 36 representative farms. Only three sets of coefficients were developed for different sizes of farms and these were used for all representative farms in the three respective size groups discussed earlier. Thus in area 3, farm 3BB had the same coefficient matrix as farm 3B, even though it was 50 acres larger. Possibly the use of average or weighted average coefficients would have decreased aggregation error.[1] However, there are many unanswered questions about the

---

[1]Sheehy discusses some of the problems connected with coefficient averages in his thesis (41).

advantages of using such coefficients, especially when the input-output relations involved are curvilinear. For example, if weighted averages are used, questions arise about choice of the relevant weights. For the purpose of this study, only three sets of coefficients were used because of the computational simplicity.

CHAPTER VII. GUIDES TO NEW STRATIFICATIONS

The linear programming results provide some insights about the more efficient control of aggregation error. The potentials for (1) reducing the number of representative farms programmed with a minimum build-up of aggregation error and (2) achieving further reduction in aggregation error are explored in this chapter. The optimum solutions of the 36 representative farms are first analyzed to determine combinations of these farms that would be consistent with the theory developed in Chapters III and IV. Then the possibilities for using alternative stratification systems are discussed. Such new systems could be based on the original stratification factors or on new factors postulated to be more closely associated with individual farm response patterns.

Postprogramming Analysis of the 36 Farms

Chapter V discusses how the original 36 representative farms were combined into groups of ten, three and one representative farms for making the aggregation error comparisons. These three preprogramming groupings were based primarily on data characteristics of the original 36 representative farms. Optimal solutions have now been obtained for the original 36 representative farms. This information may be used to determine what groupings are implied by the optimal solutions themselves and what similarities exist between the optimal solution groups and the preprogramming groups.

The theory discussed in this thesis provides two basic criteria for use in grouping farms to control aggregation error. First, farms may be grouped by restrictive resources. This procedure was followed by Sheehy and McAlexander (42) and Frick and Andrews (15) in their empirical work. A representative farm may substitute for a group of individual farms with a minimum of aggregation error if the individual farms all have the same limiting resources for the product in question. Second, farms may be grouped by response patterns. This procedure groups farms so that they approach the condition of qualitatively homogeneous output vectors. Theorem I in Chapter III has suggested the value of this procedure in farm stratification. Both these criteria require information about the optimal solutions of the farms to be grouped.

Grouping restrictive resources

Figure 12 shows the original 36 farms grouped by restrictive resources.[1] This grouping is developed from the optimum linear programming solutions for the $11.00 hog price and the $15.50 beef price. The main restrictive resources at this price combination are (1) cropland and pasture, which are restrictive on all farms, (2) operator and family labor in 3 months, (3) hog facilities and (4) beef facilities. The classification presented in Figure 12 is obtained by dividing hog facility restrictions into four categories and beef restrictions into two categories.

---

[1]The first part of a representative farm number tells the area in which the farm is located, the second part denotes the size grouping and, where used, the lower case letter denotes the two levels of hired labor available on the farm. The development of these farms is discussed in Chapter V.

Hog facilities restrictive	Labor restrictive: None — No restrictive beef fac.	None — Beef fac. res.	April — No restrictive beef fac.	April — Beef fac. res.	April and October — No restrictive beef fac.	April and October — Beef fac. res.	Apr., Oct. and Nov. — No restrictive beef fac.	Apr., Oct. and Nov. — Beef fac. res.
None	1A* 2A 3A 4A* 7A* 8A*	5A* 5B 6A 6B 9A* 10A		6Cb 9B				
Farrowing			3B	6Ca		8B		5C 9Ca
Feeding				10B	1B	7B		2Cb 9Cb
Farrowing and feeding			2B	10Ca		4B 10Cb	1C 4Ca 4Cb 7C 8Ca 8Cb	2Ca 3C

Figure 12. Stratification of the 36 representative farms by resources restrictive in their optimal solutions at the $11.00 pork price and $15.50 beef price
(Note -- cropland and pasture are restrictive on all farms)

Three groups in the figure contain 18 of the original 36 representative farms. Cropland and pasture are the only restrictive resources on six of the small farms -- 1A, 2A, 3A, 4A, 7A and 8A. Cropland, pasture and beef facilities are restrictive on six additional farms in the small and medium size groups. The remaining major group consists of six large farms on which cropland, pasture, hog farrowing and feeding facilities and April, October and November operator and family labor are restrictive. Larger farms tend to have more restrictive resources than small farms, as was found in the detailed analysis of area 3 in Chapter VI.

The 36 farms are grouped into 15 cells in Figure 12. The original 36 farms could be represented by these 15 representative farms (one for each occupied cell in Figure 12) with a small amount of additional aggregation error. Using a similar procedure, Frick and Andrews stratified 51 farms into five representative farms and added only 6.6 percent to the aggregation error (15, p. 698). The 15 representative farms implied by Figure 12 should yield similar accuracy compared to the original stratification of 36 representative farms.

Grouping by response patterns

To the extent that the data are available for its use, grouping farms by response patterns (or more specifically, stratification so that farms meet the qualitatively homogeneous output vector requirement of Theorem I) should yield better estimates than grouping by restrictive resources. Groups of farms with similar response patterns are subsets of groups of farms with the same restrictive resources.

Two farms may have the same restrictive resources and still not meet the qualitatively homogeneous output vector condition of Theorem I. On the other hand, farms which meet the conditions of Theorem I will have the same restrictive resources.

Figure 13 shows the original 36 farms grouped by major response patterns. All 36 of the representative farms sell fed beef and all but three of them have hog enterprises at the $11.00 pork price and $15.50 beef price. The major group of farms in Figure 13 includes seven farms which hire labor, sell corn, hogs and fed beef and buy feeder calves in addition to those raised on the farm. Another group of five farms uses hired labor and sells hogs and fed beef but not corn. These five farms also purchase additional feeder calves and must purchase additional facilities for their livestock enterprises. A total of 22 of the representative farms purchase hired labor while 14 do not.

Some further division of these cells could be made -- that is, more cells could be defined using more detail with the same factors. Possibilities include stratification by type of hog enterprise, by type of calf feeding system, by whether or not farm had beef cows, by the month labor was hired and by types of rotations. However, the divisions in Figure 13 are intended to delineate major differences in response patterns. The divisions also result in a workable number of groups.

Thirteen of the cells in Figure 13 contain one or more representative farms. These 13 groups represent the major different types of response patterns or farm organizations found on the original 36 farms. Programming new representative farms for each of these 13 groups should

		Sell fed beef							
		Sell corn				Feed own corn			
		Buy calves		Feed own calves		Buy calves		Feed own calves	
		Buy fac.	Do not buy fac.	Buy fac.	Do not buy fac.	Buy fac.	Do not buy fac.	Buy fac.	Do not buy fac.
Sell hogs	Hire labor		4B 4Ca 4Cb 7B 7C 8B 8Cb	2Cb 9Cb		2Ca 5C 6Cb 10Ca 10Cb	1C 8Ca 9B 9Ca	6Ca 10B	3C
	No hired labor		1B 4A* 8A*		10A		3B 5A* 9A*	6B	2B 5B 6A
No hogs	No hired labor		1A* 3A 7A*				2A		

Figure 13. Stratification of the 36 representative farms by response patterns at the $11.00 pork price and $15.50 beef price

yield supply estimates that are nearly as accurate as the original 36 farm estimates.

Three pairs of farms out of the 36 meet the strict qualitatively homogeneous output vector requirements of Theorem I. These pairs, denoted by (*) in Figures 12 and 13, are 4A and 8A, 1A and 7A, and 5A and 9A. These three pairs could be combined, reducing the number of representative farms to 33, with little loss of accuracy in the aggregate estimates.[1] There are a few other similarities between the two figures. Farms 4Ca, 4Cb, 7C and 8Cb are grouped together on both figures. Farm 3A is grouped with the 1A - 7A pair on both figures. Other paired farms under both criteria are 2Cb and 9Cb, 6B and 6A, and 1C and 8Ca. The remainder of the farms are grouped differently under the restrictive resource criteria than they are under the criteria of similar response patterns. This divergence emphasizes the differences in the outcome of these two methods of farm stratification.

Grouping areas or sizes

The combinations of representative farms suggested by either stratification by restrictive resources or stratification by response patterns often stretch across both size and area classifications. However, due to differences in yields, costs and input-output relations, some of these combinations may be impractical because of the coefficient

---

[1]These pairs of farms do not meet the identical coefficient matrix requirements of Theorem I because of differences in crop yields and costs between areas. Thus, there is no guarantee that exact aggregation can be achieved if the pairs were combined.

problems involved. For example, it may be difficult to determine the correct crop yields for the representative farm covering farms 6Ca and 10B in Figure 13. In view of this problem, it may be more meaningful for some research projects to consider mainly combinations of strata within the existing classification, rather than the completely new classifications of Figures 12 and 13.

The central question is then one of determining the areas and/or sizes of the original 36 cell stratification that may be combined without large increases in aggregation error. Figures 12 and 13 offer some clues to the answer of this question. Considering combinations of areas, the figures suggest that areas 4 and 8 could be combined with less aggregation error than the other areas. The small farms in these areas, farms 4A and 8A, could be combined with almost no aggregation error since these farms have qualitatively homogeneous output vectors as discussed previously. In Figure 13, farms 4B and 8B are in the same group and farms 4Cb and 8Cb are in the same group. In Figure 12, farms 4Ca and 8Ca and farms 4Cb and 8Cb are in the same groups. Thus, all sizes of pairs of farms in these two areas are grouped together at least once by the two procedures. These results imply that areas 4 and 8 could be combined with a minimum of aggregation error.

Additional decisions for area combinations could be made from Figures 12 and 13 in a like manner until the desired number of areas is obtained. Of course, aggregation error would accumulate as more areas are combined until the situation portrayed by the three representative farm subgroup of Chapter V is reached. The amount of aggregation error resulting from this grouping is discussed in Chapter VI.

Possibilities for size combinations within areas may be approached similarly, although the problem is less complex. Different sizes of farms in the same area are generally not grouped together in Figures 12 and 13. Thus, the three original sizes of representative farms appear to be a meaningful classification that is reflected in the response patterns and thereby tends to reduce aggregation error. The two figures provide some suggestion that categories Ca and Cb, which represent different amounts of available hired labor, could be grouped together. The same coefficient matrices are used for both of these two categories and apparently the resource ratios are generally not different enough to cause differences in the response patterns of the farms.

New Delineations of Representative Farms

Using the same stratification factors

The last section has discussed combining groups within the original stratification of 36 representative farms. It is also possible to regroup the individual sampled farms by the same factors of soil-type area and farm size, either (1) holding the number of representative farms constant and changing their coverage or (2) increasing the number of representative farms. For example, delineation of representative farms to recognize more extreme resource ratios may identify characteristics that lead to differences in response patterns among farms and thereby decrease aggregation error. On the other hand, a straightforward increase in the number of sizes of representative farms may decrease aggregation error if it defines new groups of farms having different response patterns.

The empirical work that has been done offers little hint of the effect on aggregation error of restratification by the size of farm and soil-type area factors. A large amount of additional data on items such as subarea crop yields and costs and labor requirements on intermediate sizes of farms would be necessary before the impact of new stratifications could be ascertained. Even if the effect of adding ten more representative farms were determined, little could be said about increasing the number still further. Generally, Figure 11, page 126, shows a relatively small decrease in aggregation error between the ten and 36 representative farm estimates. Projections of such trends are uncertain. However, there is a weak suggestion that defining more than 36 representative farms using the same factors would result in relatively small decreases in aggregation error.

Additional stratification factors

Additional attempts to develop new and more effective stratifications must be turned toward consideration of new factors. The empirical work that has been done provides almost no information about the success such stratifications may have in controlling aggregation error. There is always the possibility that recognizing some other factors in the stratification procedure would result in more differences in response patterns among the representative farms programmed and thereby would reduce aggregation error in the supply estimates.

In general, the choice of stratification factors should be aimed at delineating groups of farms that approach the conditions of Theorem I -- identical coefficient matrices and qualitatively homogeneous.

output vectors. The list of such factors is long because of the complexity of the relations involved and the variety of different factors that may influence the realization of the conditions of Theorem I. Expanding the categories discussed by Sheehy (41), possible stratification factors for Iowa farms would include:

(1) Physical environment factors, such as climate, topography and soil-type, which affect crop yields and production costs. Cropland-pasture ratios and the row crop capability of the cropland are also possible classification factors within this group.

(2) Institutional restrictions, such as markets, land tenure patterns and government regulations. These factors are candidates for grouping farms when they have an unequal effect on individual farms.

(3) Motivational forces, including risk aversion, demand for leisure and preferences for certain enterprises. The age of farm operators and the availability of off-farm work may influence motives.

(4) Management ability, including both the differences between farmers and the differences between enterprises for the same farmer.

(5) Technology, which may have an unequal impact on different farms.

(6) Resource endowments, including cropland, labor, livestock facilities and capital. In some cases, ratios of such resources may be considered as classification factors.

The first five classes of factors generally affect the coefficient matrices of the linear programming models while the resource vectors of the models are primarily determined by the sixth class of items.

Decisions made in reference to the specification problem affect the solutions to the stratification and aggregation error problem (45). Questions about the length of run and which factors to hold constant and which to vary should be answered before the representative farms are delineated. Some factors may be relevant to the classification because they affect the existence of alternative enterprises for the farms in question. For example, a classification of cash grain and livestock farms may be made solely on the alternatives considered for the two groups of farms -- i.e., the assumption that for the time span considered by the model, cash grain farmers will not consider livestock. Such a classification would obviously not be warranted in a more normative model where sufficient time is assumed to exist for all farmers to make the most profitable adjustment.

The choice of stratification factors to use in a particular research project is always difficult. It requires a thorough knowledge of the agriculture in the population and of the main relationships that determine the optimum solutions of the linear programming models being used. The condition of qualitatively homogeneous output vectors is a requirement of optimum solutions of individual farms. Therefore, a considerable amount of prestratification programming may be invaluable in identifying the important variables affecting aggregation error in a specific project.

A basic problem facing economists using the representative farm linear programming supply estimation technique should not be overlooked. It is the problem of balancing the desire to reduce aggregation error against the limited resources available for research. This problem is analogous to the producing unit or firm aggregation problem discussed by Heady (19) and is especially acute during the stratification step. The problem of optimal use of scarce resources is certainly not new to economists. However, its application to research may be overlooked in the quest for bigger and more accurate models. In simple terms, each stratification factor should be selected considering its ability to control aggregation error against the relative cost of implementing it. The final group of factors selected should be the unique group that will minimize aggregation error given the research resources available for this task.

CHAPTER VIII. SUMMARY

The Problem and Objectives

Aggregate supply estimates are often desired for a population of individual farms under alternative demand, price and farm program assumptions. Using the single farm as the unit of analysis in making these aggregate estimates has potential for increasing both the usefulness and the accuracy of the estimates. It is common, however, for the population of interest to be too large to allow estimation of the supply for every individual farm and summation of these estimates into the desired aggregate estimate. The technique of representative farm linear programming supply estimation is commonly used in this situation. The steps of this technique are (1) delineation of a small number of representative farms within the population, (2) determining supply estimates of these representative farms by linear programming techniques, (3) weighting the supply estimates of the representative farms and (4) summing these weighted estimates into the desired aggregate estimates.

The aggregation error problem is inherent in this technique. Aggregation error is defined as the difference between (1) the aggregate supply estimate developed as the sum of the linear programming solutions for each individual farm in the population and (2) the aggregate supply estimated by a small number of representative farms. This thesis studies the problem of aggregation error in representative farm linear programming supply estimates. The specific objectives were as follows:

(1) To review the current state of literature and theory relating to the problem of aggregation error and to evaluate these theoretical concepts for practical usefulness.

(2) To explore the theoretical aspects of developing error-free or minimum-error aggregates and if possible to develop additional theory in this area.

(3) To develop empirical supply estimates of Iowa pork and beef based on different stratifications of representative farms and to determine the relative magnitude of the aggregation error and possible factors that contribute to it by comparing these different estimates.

(4) To utilize the results of the first three objectives to recommend practical procedures that may be followed by research workers in controlling aggregation error.

The Theory of Aggregation Error

Two theorems stating conditions sufficient to assure exact aggregation are developed in Chapter III. Exact aggregation describes the situation in which the total supply estimated as the sum of the individual farms is identical to the total supply estimated by one representative farm. Theorem I states that sufficient conditions for exact aggregation are that all of the individual farms have (1) identical coefficient matrices and (2) qualitatively homogeneous output vectors (QHOV). The condition of QHOV requires that the identity of the activities in the optimum solutions must be the same for all farms. Theorem I thus allows the relative levels of the optimal activities to vary among individual farms but requires that all individual farms have the same set of activities in their optimal solutions. When a group of farms meets the conditions of Theorem I, programming only one representative farm will provide exact aggregation.

The condition of QHOV expressed in Theorem I is in terms of the solutions of the individual farms. A link between the solutions and the coefficients of the individual farms is provided by Theorem II which states that sufficient conditions for exact aggregation are that all of the individual farms have (1) identical coefficient matrices, (2) identical net return vectors and (3) identical dual solution vectors. Using Theorem II, it is technically possible to determine strata of individual farms sufficient for exact aggregation from information found in the coefficient matrix of the individual farms. However, the number of representative farms required for exact aggregation under the conditions of Theorem II is extremely high for larger coefficient matrices. This fact makes achieving exact aggregation an unrealistic goal in many research projects. The alternative goal is to develop a procedure for selecting representative farms that will minimize aggregation error. The individual researcher can then choose the number of representative farms he wants to analyze based on comparisons between the computing costs of including each representative farm and the amount of aggregation error it is expected to eliminate.

A proposal for a formal mathematical procedure to minimize aggregation error has been made (27). However, it is fairly complex from a technical standpoint and appears to be computationally difficult. Therefore, an alternative procedure is worthy of consideration -- merely utilizing Theorem I on an intuitive basis to develop representative farms that would be expected to reduce aggregation error. An intuitive interpretation of the implication of Theorem I to the representative farm development problem is summarized later in this chapter.

The direction of aggregation error was discussed in Chapter IV -- a problem alluded to by the popular use of the term "aggregation bias." It was found that in linear programming models of many activities and resources, very little tendency exists for a predominate direction of error. As a result, the use of the term "bias" in reference to representative farm linear programming supply estimates appears to be misleading if not inaccurate.

The conditions stated in Theorems I and II are sufficiently binding to assure exact aggregation; the theorems say nothing about the necessity of meeting these conditions to achieve exact aggregation. The possibility is thus left open for development of less binding sufficient conditions for exact aggregation. It is suggested in Chapter IV, however, that this possibility is very improbable. Except for a technicality of linear programming, the Theorem I condition of QHOV is also necessary for exact aggregation. Achieving exact aggregation for all products is highly improbable if the QHOV condition is not met. This fact leads to delineation of the main causes of aggregation error as (1) not meeting the requirement of QHOV, (2) extreme variation in expected net returns and (3) variation in coefficient matrices. Stratification of farms must recognize these factors if aggregation error is to be controlled.

Aggregation Error in Empirical Work

The discussion of theory leaves unanswered questions concerning (1) the magnitude of the aggregation error in actual models and (2) the relative importance of different factors contributing to aggregation error. A model was developed to answer these questions for a typical research project utilizing the technique of representative farm linear programming.

Development of the model

The representative farm model was developed to estimate aggregate pork and beef supply functions for the population of commercial farms in Iowa. Development of this model is discussed in Chapter V. Four stratifications of representative farms were made involving successively smaller numbers of representative farms. Differences among the estimated population supply functions based on these four groups of representative farms are due to aggregation error. The four groups of representative farms were as follows:

(1) Thirty-six representative farms classified by ten soil-type areas, three sizes of farms and high and low hired labor availabilities,

(2) ten representative farms, one for each soil-type area of Iowa,

(3) three representative farms, one for each of the three size strata of farms in Iowa and

(4) one representative farm to represent the entire population.

Basically, the resources on representative farms in the three smaller groups were weighted averages of the resources in the 36 farm group.

A linear programming model was developed for each of the representative farms in the four groups. The coefficients in the linear programming models of the 36 farm group were not averaged for the three smaller groups -- rather typical or modal coefficients were used for each of the smaller groups. The supply functions were estimated by variable pricing techniques. Sixteen price combinations for pork and beef were programmed. These 16 optimum solutions allowed determination of a beef supply function for each of four pork prices and a pork supply function for each of four beef prices. Four points were estimated on each of these supply functions.

Results of the empirical work

Programming results from the four groups of representative farms were aggregated into four sets of population supply functions for beef and four sets for pork. These estimated supply functions are presented in Figures 7 and 8 and Tables 8 and 9 of Chapter V. The four sets of estimated supply functions are generally quite similar. The over-all slopes or elasticities of the functions are in agreement and differences between them are small.

The programmed supply functions are in agreement with the hypothesized lack of predominate direction of aggregation error. When compared to the 36 farm aggregate estimate, the three smaller groups of representative farms overestimate production about as many times as they underestimate production. Thus, the programming results do not indicate any significant bias in aggregation error.

The difference between the 36 farm estimate and the estimates determined from the three smaller groups of representative farms is measured for one of the beef supply functions and one of the pork supply functions. The relationship between the number of representative farms used and an index of absolute aggregation error (36 farm estimate = 100) is as follows:

Number of representative farms	Beef estimate error	Pork estimate error
1	129.4	112.8
3	116.5	106.2
10	108.6	114.2

For beef, the error decreases as larger numbers of representative farms are programmed. For pork, the three farm estimate has less error than the ten farm estimate. Since the three farm estimate resulted from a size of farm classification and the ten farm estimate resulted from a soil-type area classification, this result indicates that the size of farm classification is much more important in controlling pork aggregation error than is soil-type area. Both classification factors appear to be important in controlling aggregation error in beef supply estimates.

The severity of the aggregation error problem is extremely dependent upon the particular estimate desired. Larger amounts of aggregation error are found in livestock production estimates than in crop production estimates. Using one price combination as an example, the 36 and ten farm population production estimates differ by 31.3 percent for pork, 17.2 percent for beef, 0.8 percent for corn and 0.9 percent for soybeans. These results indicate that aggregation error in crop production estimates may be held to low levels with fewer representative farms than is true for livestock estimates.

It is difficult to determine the data characteristics in the four groups of representative farms that lead to specific aggregation errors. Optimum linear programming solutions are determined by complex interrelationships and it is sometimes difficult to relate optimum solution attributes to specific data characteristics. In addition, the four groups of representative farms differ from each other in numerous ways -- this situation makes it difficult to appraise the effect of a difference in a specific factor. These two problems limit the possibility of explaining aggregation error in terms of data characteristics of the representative farms.

Representative Farm Linear Programming Models

Farm stratification procedures

The theory and the empirical work both contribute to rather basic principles for controlling aggregation error. Theorem I suggests that individual farms should be grouped into homogeneous groups on the basis of their coefficient matrices and then subdivided so that each subgroup has identical sets of optimum production activities. The key idea is that the individual farms be stratified to account for (1) differences in their coefficient matrices and (2) differences in their expected response patterns or adjustments. Grouping farms on the basis of factors that do not influence differences in coefficients or response patterns does not reduce aggregation error. For example, grouping farms by size if size does not affect the type of adjustment or the coefficient matrix is of no value in reducing aggregation error.

Different delineations of representative farms may be required for estimating different products. The empirical work reveals large differences in the amounts of aggregation error found in different product supply estimates. The theory shows that, in some cases, exact aggregation is more easily achieved for one product supply estimates than when estimates are required for several products. These results indicate that a unique choice of stratification factors may be best for each specific research project.

The grouping implied by Theorem I goes a step further than the practice of grouping farms by restrictive resources. Grouping farms by restrictive resources is a move in the right direction since differences in restrictive resources tend to be reflected in differences in the response patterns of the individual farms. However, it is not aimed directly at the basic objective of identifying farms that have different response patterns. The requirements of Theorem I provide a more basic principle to follow in minimizing aggregation error.

Determination of the best factors to control or minimize aggregation error does not provide a complete solution to the problem of representative farm selection. There is still the problem of determining how many representative farms to use in making the estimates in cases where it is not feasible to eliminate aggregation error. The final selection of representative farms must be made considering (1) the additional research costs of including more representative farms versus (2) larger amounts of aggregation error. The problem is basically one of deciding for each research project how much of the aggregation error it is economically feasible to eliminate.

<u>Other types of error</u>

There are three sources of error in representative farm linear programming supply estimation models. These are (1) specification error, (2) sampling error and (3) aggregation error. This study has been primarily interested in aggregation error. However, the emphasis on aggregation error does not mean that the other two types of errors should be neglected or that model builders should not allocate their effort toward holding all three types of error within tolerable limits. Certainly, some balanced attempt to control all three types of errors should be made.

When considered in light of the possible specification and sampling errors involved, the aggregation errors found in the empirical work are diminished in importance. It is, of course, beyond the scope of this thesis to appraise the magnitude of the specification and sampling errors -- obviously the type of analysis followed in the empirical work does not lend itself to such measurement. However, research people who use the technique of supply estimation by representative farm linear programming should still keep all possible ways to reduce errors in the proper perspective.

ACKNOWLEDGEMENTS

The author expresses appreciation to Dr. Earl O. Heady, Professor of Economics, for his inspiration and leadership throughout this research project.

This project was sponsored cooperatively by the Farm Production Economics Division, Economic Research Service, U.S. Department of Agriculture, and the Iowa Agricultural and Home Economics Experiment Station. Special appreciation is extended to Dr. Wesley B. Sundquist, C. W. Crickman and John Lee of the Farm Production Economics Division for their suggestions and interest at all stages of the project.

The author is grateful to Jerry Sharples for his suggestions and criticisms and for the unpublished data and computer algorithms he made available. Association with Dr. Ronald Krenz and Gaylord Worden also facilitated the research work.

BIBLIOGRAPHY

1. Allen, R. H., Erling Hole and R. H. Mighell. Supply responses in milk production in the Cabot-Marshfield area, Vermont. U.S. Department of Agriculture Technical Bulletin 709. 1940.

2. Barker, Randolph and Bernard F. Stanton. Estimation and aggregation of firm supply functions. Journal of Farm Economics 47: 701-712. 1965.

3. Black, John D. Interregional competition in agricultural production. Social Science Research Council Bulletin 13: 87-101. 1932.

4. Boulton, Bill. Variable resource programming for appraising farm adjustment opportunities. Agricultural Economics Research 16: 12-22. 1964.

5. Carter, Harold O. Representative farms: guides for decision making. Journal of Farm Economics 45: 1448-1457. 1963.

6. Christensen, Raymond P. and Ronald L. Mighell. Supply response in milk production in Dodge and Barron Counties, Wisconsin. U.S. Department of Agriculture Technical Bulletin 750. 1941.

7. Cochrane, Willard W. Conceptualizing the supply relation in agriculture. Journal of Farm Economics 37: 1161-1176. 1955.

8. Day, Lee M. Use of representative farms in studies of interregional competition and production response. Journal of Farm Economics 45: 1438-1445. 1963.

9. Day, Richard H. On aggregating linear programming models of production. Journal of Farm Economics 45: 797-813. 1963.

10. Dorfman, R. D., P. A. Samuelson and R. M. Solow. Linear programming and economic analysis. New York, New York, McGraw-Hill Book Co., Inc. 1958.

11. Egbert, Alvin C. and Earl O. Heady. Regional adjustments in grain production: a linear programming analysis. U.S. Department of Agriculture Technical Bulletin 1241 (with supplement). 1961.

12. Egbert, Alvin C., Earl O. Heady and Ray F. Brokken. Regional changes in grain production: an application of spatial linear programming. Iowa Agricultural Experiment Station Research Bulletin 521. 1964.

13. Elliott, F. F. The "representative firm" idea applied to research and extension in agricultural economics. Journal of Farm Economics 10: 483-498. 1928.

14. Elliott, F. F., Jesse W. Tapp and Rex E. Willard. Types of farming in North Dakota. U.S. Department of Agriculture Technical Bulletin 102. 1928.

15. Frick, George E. and Richard A. Andrews. Aggregation bias and four methods of summing farm supply functions. Journal of Farm Economics 47: 696-700. 1965.

16. Gilson, J. C. Discussion: Use of representative farms in studies of interregional competition and production response. Journal of Farm Economics 45: 1445-1447. 1963.

17. Goodwin, John W., James S. Plaxico and William F. Lagrone. Aggregation of normative microsupply relationships for dryland crop farms in the rolling plains of Oklahoma and Texas. Oklahoma Agricultural Experiment Station Technical Bulletin T-103. 1963.

18. Hartley, H. O. Total supply functions estimated from farm surveys. Unpublished mimeographed paper presented before the North Central Farm Management Research Committee, Chicago, Illinois, 1962. Ames, Iowa, Department of Statistics, Iowa State University of Science and Technology. ca. 1962.

19. Heady, Earl O. Aggregation and related problems in models for analysis of interregional competition. In King, Richard A., ed. Interregional competition. pp. 129-145. Raleigh, North Carolina, The Agricultural Policy Institute, University of North Carolina. 1963.

20. Heady, Earl O. Economics of agricultural production and resource use. Englewood Cliffs, New Jersey, Prentice Hall, Inc. 1952.

21. Heady, Earl O. and Wilfred Candler. Linear programming methods. Ames, Iowa, Iowa State College Press. 1958.

22. Heady, Earl O. and Norman K. Whittlesey. A programming analysis of interregional competition and surplus capacity of American agriculture. Iowa Agricultural Experiment Station Research Bulletin 538. 1965.

23. Holmes, C. L. Types of farming in Iowa. Iowa Agricultural Experiment Station Bulletin 256. 1929.

24. Klein, L. R. Macroeconomics and the theory of rational behavior. Econometrica 14: 93-108. 1946.

25. Krenz, Ronald D., Earl O. Heady and Ross V. Baumann. Profit-maximizing plans and static supply schedules for fluid milk in the Des Moines milkshed. Iowa Agricultural Experiment Station Research Bulletin 486. 1960.

26. Lee, John E., Jr. Exact aggregation: a discussion of Miller's theorem. Agricultural Economics Research 18: 58-61. 1966.

27. Lee, John E., Jr. The minimum sufficient conditions for exact aggregation in linear programming models: a non-mathematical exposition. Unpublished mimeographed working materials prepared in Farm Production Economics Division, Economic Research Service, U.S. Department of Agriculture. Washington, D.C., Economic Research Service, U.S. Department of Agriculture. 1966.

28. McKee, Dean E. and Laurel D. Loftsgard. Programming intra-farm normative supply functions. In Heady, Earl O., C. B. Baker, H. O. Diesslin, S. W. Kehrberg and S. D. Staniforth, eds. Agricultural supply functions: estimating techniques and interpretations. pp. 152-166. Ames, Iowa, Iowa State University Press. 1961.

29. Marshall, Alfred. Principles of economics. 8th ed. New York, New York, The Macmillan Company. 1952.

30. May, Kenneth. The aggregation problem for a one-industry model. Econometrica 14: 285-298. 1946.

31. Mighell, Ronald L. and John D. Black. Interregional competition in agriculture. Cambridge, Massachusetts, Harvard University Press. 1951.

32. Miller, Thomas A. Sufficient conditions for exact aggregation in linear programming models. Agricultural Economics Research 18: 52-57. 1966.

33. Nerlove, Marc and Kenneth L. Bachman. The analysis of changes in agricultural supply: problems and approaches. Journal of Farm Economics 42: 531-554. 1960.

34. Plaxico, James S. Aggregation of supply concepts and firm supply functions. Southern Cooperative Series Bulletin 56: 76-91. 1958.

35. Plaxico, James S. and Luther G. Tweeten. Representative farms for policy and projection research. Journal of Farm Economics 45: 1458-1465. 1963.

36. Schuh, George E. The supply of fluid milk in the Detroit milkshed as affected by cost of production. Michigan Agricultural Experiment Station Technical Bulletin 259. 1957.

37. Schultz, Theodore W. Reflections on agricultural production, output and supply. Journal of Farm Economics 38: 748-762. 1956.

38. Sharples, Jerry A. Iowa contributing project to the NC-54 Regional Adjustment Study. Unpublished data. Ames, Iowa, Department of Economics and Sociology, Iowa State University of Science and Technology. 1966.

39. Sharples, Jerry A. Normative production of hogs, beef cattle and other farm products in Iowa. Unpublished Ph.D. thesis. Ames, Iowa, Library, Iowa State University of Science and Technology. 1967.

40. Sharples, Jerry A., Thomas A. Miller and Lee M. Day. An evaluation of a firm model in estimating aggregate supply response. Unpublished preliminary draft of an NC-54 regional publication. [To be published by the Iowa Agricultural Experiment Station, Ames, Iowa. ca. 1968.]

41. Sheehy, Seamus J. Selection of representative benchmark farms in synthetic supply estimation. Microfilm copy, unpublished Ph.D. thesis. State College, Pennsylvania, Library, Pennsylvania State University. 1964.

42. Sheehy, Seamus J. and R. H. McAlexander. Selection of representative benchmark farms for supply estimation. Journal of Farm Economics 47: 681-695. 1965.

43. Shrader, W. D., F. W. Schaller, J. T. Pesek, D. F. Slusher and F. F. Riecken. Estimated crop yields on Iowa soils. Iowa Agricultural Experiment Station Special Report 25. 1960.

44. Shrader, W. D., David Slusher and F. F. Riecken. Soil management groups of Iowa soils. Mimeographed. Ames, Iowa, Department of Agronomy, Iowa State University of Science and Technology. 1962.

45. Stovall, John G. Sources of error in aggregate supply estimates. Journal of Farm Economics 48: 477-480. 1966.

46. Taussig, F. W. Principles of economics. 4th ed. Vol. 1. New York, New York, The Macmillan Company. 1939.

47. Technical Committee of Lake States Dairy Adjustment Study. Equilibrium analysis of income-improving adjustments on farms in the Lake States dairy region, 1965. Minnesota Agricultural Experiment Station Technical Bulletin 246. 1963.

48. Thiel, H. Linear aggregation of economic relations. Amsterdam, Holland, North Holland Publishing Company. 1954.

49. Thompson, James F. Defining typical resource situations. Southern Cooperative Series Bulletin 56: 32-43. 1958.

50. U.S. Department of Agriculture. Consumer and Marketing Service. Livestock and meat statistics. Supplement for 1965 to Statistical Bulletin 333. 1966.

51. U.S. Department of Commerce. Bureau of the Census. Census of Agriculture, 1959. Vol. 1, Part 16. 1961.

52. Wilcox, Walter W. Types of farming research and farm management. Journal of Farm Economics 20: 417-429. 1938.

APPENDIX

Appendix Table 1. Linear programming model used in study of aggregation error in Iowa

	P_1	P_2	P_3	P_4	P_5	P_6
C_j	110.00[a]	15.50[a]	-73.17[a]	-110.82[a]	-221.64[a]	69.51[a]
1						
2						
3						
4						
5						
6						
7						
8						
9						
10						
11						
12						
13						
14						
15						
16						
17						
18						
19						
20				-1.0		
21					-2.0	
22			-1.0			1.0
23			7.317[a]	11.082[a]	11.082[a]	
24			-6.219[a]	-9.420[a]	-9.420[a]	
25		1.0				
26	10.0					
27						
28						
29						
30						
31						
32						
33						
34						
35						
36						

[a]Coefficients that vary with the price level of pork and beef programmed. The coefficients shown are for a pork price of $11.00 and a beef price of $15.50.

Appendix Table 1. (Continued)

	P_7	P_8	P_9	P_{10}	P_{11}	P_{12}
C_j	-24.30^{bc}	7.94^{bc}	-19.44^{bc}	-2.74^{bc}	4.41^{bc}	-21.42^{bc}
1						
2						1.0
3	1.0	1.0	1.0	1.0	1.0	
4						
5			$-.82^{b}$	$-.82^{b}$	$-.66^{b}$	$.78^{b}$
6	-92.0^{b}	-46.0^{b}	-47.5^{b}	-24.0^{b}	-19.2^{b}	-40.25^{b}
7						
8						
9						
10						
11						
12						
13						
14						
15						
16						
17			-4.69^{b}	-4.69^{b}	-3.75^{b}	-4.47^{b}
18						
19						
20						
21						
22						
23	1.721^{bc}	1.972^{bc}	1.336^{bc}	1.463^{bc}	1.615^{bc}	1.534^{bc}
24						
25						
26						
27	3.26^{c}	3.53^{c}	2.82^{c}	2.94^{c}	3.12^{c}	2.82^{c}
28						
29	1.25^{c}	1.20^{c}	1.15^{c}	1.12^{c}	1.13^{c}	1.15^{c}
30	0.80^{c}	0.70^{c}	0.40^{c}	0.35^{c}	0.40^{c}	0.40^{c}
31	0.50^{c}	0.73^{c}	0.25^{c}	0.36^{c}	0.48^{c}	0.25^{c}
32	0.36^{c}	0.18^{c}	0.84^{c}	0.75^{c}	0.60^{c}	0.84^{c}
33		0.12^{c}		0.06^{c}	0.10^{c}	
34		0.25^{c}		0.12^{c}	0.20^{c}	
35	0.35^{c}	0.35^{c}	0.18^{c}	0.18^{c}	0.21^{c}	0.18^{c}
36						

[b]Coefficients that vary according to the area of the state in which the farm is located. The coefficients shown are for area 9.

[c]Coefficients that vary according to the three size groups of farms. The coefficients shown are for a medium-sized farm.

Appendix Table 1. (Continued)

	P_{13}	P_{14}	P_{15}	P_{16}	P_{17}	P_{18}
C_j	-6.59[bc]	-0.14[bc]	-3.80[bc]	-16.42[bc]	-4.00[bc]	-1.511[c]
1				1.0	1.0	
2	1.0	1.0	1.0			
3						
4						
5	-.78[b]	-1.04[b]	-3.1[b]	-1.1[b]	-2.2[b]	
6	-20.25[b]	-13.50[b]		-13.5[b]		10.0
7						
8						
9						
10						
11						
12						
13						
14						
15						
16						
17	-4.47[b]	-2.98[b]		-3.13[b]		-5.60
18						
19						
20						
21						
22						
23	1.614[bc]	1.478[bc]	0.268[bc]	1.187[bc]	0.288[bc]	0.062[c]
24						
25						
26						
27	2.94[c]	2.63[c]	0.30[c]	2.05[c]	0.30[c]	0.365[c]
28						
29	1.12[c]	0.98[c]	0.30[c]	0.89[c]	0.30[c]	
30	0.35[c]	0.33[c]		0.20[c]		
31	0.36[c]	0.40[c]		0.12[c]		
32	0.75[c]	0.50[c]		0.75[c]		
33	0.06[c]	0.08[c]				
34	0.12[c]	0.16[c]				0.274[c]
35	0.18[c]	0.18[c]		0.09[c]		0.091[c]
36						

Appendix Table 1. (Continued)

	P_{19}	P_{20}	P_{21}	P_{22}	P_{23}	P_{24}
C_j	-4.31^c	-9.80^c	0	-84.87^c	-175.53^c	-179.53^c
1						
2						
3						
4			-1.0	1.65	2.25	1.82
5		1.0	1.0			
6	6.0					
7						1.0
8				1.0	1.0	
9						
10				8.0	8.0	8.0
11						
12						
13						
14						
15						
16						
17				65.25	128.25	128.25
18	-20.0					
19		-20.0				
20						
21						
22						
23	1.854^c	0.543^c		11.831^c	20.241^c	20.641^c
24						
25						
26				-19.4779	-37.75	-37.75
27	0.167^c	2.617^c		31.36^c	47.52^c	45.89^c
28				5.23^c	20.43^c	19.50^c
29				$.95^c$	2.86^c	2.86^c
30				2.38^c	2.85^c	2.85^c
31		1.145^c		6.65^c	2.85^c	2.85^c
32		$.782^c$		4.75^c	2.85^c	2.85^c
33	0.167^c	$.690^c$		5.70^c	9.98^c	9.28^c
34				2.85^c	2.85^c	2.85^c
35				2.85^c	2.85^c	2.85^c
36						

Appendix Table 1. (Continued)

	P_{25}	P_{26}	P_{27}	P_{28}	P_{29}	P_{30}
c_j	-186.28[c]	-356.78[c]	-372.57[c]	-535.16[c]	-558.85[c]	-24.29[c]
1						
2						
3						
4		3.35		4.68		
5						
6						
7	1.0	1.0	1.0	1.0	1.0	
8						
9	8.0		16.0		24.0	
10		16.0		24.0		
11						0.65
12						0.65
13						1.0
14						1.0
15						
16						
17	124.42	254.30	248.84	381.45	373.26	30.13
18						
19						16.18
20						
21						
22						1.0
23	21.316[c]	41.054[c]	42.633[c]	61.580[c]	63.949[c]	2.429[c]
24						
25						-10.5
26	-37.75	-75.50	-75.50	-113.25	-113.25	
27	45.89[c]	91.75[c]	91.75[c]	137.61[c]	137.61[c]	15.42[c]
28	19.50[c]	32.47[c]	32.47[c]	48.13[c]	48.13[c]	4.23[c]
29	2.86[c]	6.19[c]	6.19[c]	12.34[c]	12.34[c]	1.67[c]
30	2.85[c]	9.48[c]	9.48[c]	10.60[c]	10.60[c]	2.19[c]
31	2.85[c]	7.27[c]	7.27[c]	12.33[c]	12.33[c]	2.19[c]
32	2.85[c]	5.70[c]	5.70[c]	10.12[c]	10.12[c]	1.85[c]
33	9.28[c]	14.98[c]	14.98[c]	21.16[c]	21.16[c]	2.27[c]
34	2.85[c]	6.18[c]	6.18[c]	12.33[c]	12.33[c]	
35	2.85[c]	9.48[c]	9.48[c]	10.60[c]	10.60[c]	1.02[c]
36						

Appendix Table 1. (Continued)

	P_{31}	P_{32}	P_{33}	P_{34}	P_{35}	P_{36}
C_j	-21.44[c]	-16.04[c]	-16.04[c]	-32.08[c]	-25.13[c]	-22.33[c]
1						
2						
3						
4	1.65					1.65
5						
6						
7						
8						
9						
10						
11	0.65	0.65		0.65	0.65	0.65
12	0.65		0.65	0.65	0.65	0.65
13	1.0	1.0		1.0		
14	1.0		1.0	1.0		
15					1.0	1.0
16					1.0	1.0
17	31.36	27.20	27.20	54.40	30.13	31.36
18						
19	13.40	7.20	7.20	14.40	16.18	13.40
20		1.0	1.0			
21				2.0		
22	1.0				1.0	1.0
23	2.144[c]	1.604[c]	1.604[c]	1.604[c]	2.513[c]	2.233[c]
24						
25	-11.0	-11.0	-11.0	-22.0	-10.5	-11.0
26						
27	14.04[c]	8.50[c]	8.50[c]	17.00[c]	13.11[c]	11.93[c]
28	5.35[c]	5.77[c]		5.77[c]	3.60[c]	4.55[c]
29	1.40[c]		1.06[c]	1.06[c]	1.42[c]	1.19[c]
30	0.94[c]		1.62[c]	1.62[c]	1.86[c]	0.80[c]
31	0.93[c]		1.62[c]	1.62[c]	1.86[c]	0.79[c]
32	0.93[c]		1.65[c]	1.65[c]	1.57[c]	0.79[c]
33	2.06[c]		2.55[c]	2.55[c]	1.93[c]	1.75[c]
34	1.12[c]	1.11[c]		1.11[c]		0.95[c]
35	1.31[c]	1.62[c]		1.62[c]	.87[c]	1.11[c]
36						

Appendix Table 1. (Continued)

	P_{37}	P_{38}	P_{39}	P_{40}	P_{41}	P_{42}
C_j	-16.57^c	-16.57^c	-33.14^c	-26.12^c	-22.81^c	-21.87^c
1						
2						
3						
4					1.65	
5						
6						
7						
8						
9						
10						
11	0.65		0.65	0.65	0.65	0.65
12		0.65	0.65	0.65	0.65	
13				1.0	1.0	1.0
14				1.0	1.0	
15	1.0		1.0			
16		1.0	1.0			
17	27.20	27.20	54.40	25.13	27.66	22.40
18				30.00	22.00	30.0
19	7.20	7.20	14.40	12.18	11.75	3.20
20	1.0	1.0				1.0
21			2.0			
22				1.0	1.0	
23	1.657^c	1.657^c	1.657^c	2.612^c	2.281^c	2.187^c
24						
25	-11.0	-11.0	-22.0	-10.5	-11.0	-11.0
26						
27	7.22^c	7.22^c	14.44^c	15.42^c	14.04^c	8.50^c
28	4.90^c		4.90^c	4.23^c	5.35^c	5.77^c
29		0.90^c	0.90^c	1.67^c	1.40^c	
30		1.37^c	1.37^c	2.19^c	0.94^c	
31		1.38^c	1.38^c	2.19^c	0.93^c	
32		1.40^c	1.40^c	1.85^c	0.93^c	
33		2.17^c	2.17^c	2.27^c	2.06^c	
34	0.95^c		0.95^c		1.12^c	1.11^c
35	1.37^c		1.37^c	1.02^c	1.31^c	1.62^c
36						

Appendix Table 1. (Continued)

	P_{43}	P_{44}	P_{45}	P_{46}	P_{47}	P_{48}
C_j	-21.87^c	-43.74^c	-26.96^c	-23.70^c	-22.40^c	-22.40^c
1						
2						
3						
4				1.65		
5						
6						
7						
8						
9						
10						
11		0.65	0.65	0.65	0.65	
12	0.65	0.65	0.65	0.65		0.65
13		1.0				
14	1.0	1.0				
15			1.0	1.0	1.0	
16			1.0	1.0		1.0
17	22.40	44.80	25.13	27.66	22.40	22.40
18	30.0	60.0	30.00	22.00	30.0	30.0
19	3.20	6.40	12.18	11.75	3.20	3.20
20	1.0				1.0	1.0
21		2.0				
22			1.0	1.0		
23	2.187^c	2.187^c	2.696^c	2.370^c	2.240^c	2.240^c
24						
25	-11.0	-22.0	-10.5	-11.0	-11.0	-11.0
26						
27	8.50^c	17.00^c	13.11^c	11.93^c	7.22^c	7.22^c
28		5.77^c	3.60^c	4.55^c	4.90^c	
29	1.06^c	1.06^c	1.42^c	1.19^c		0.90^c
30	1.62^c	1.62^c	1.86^c	0.80^c		1.37^c
31	1.62^c	1.62^c	1.86^c	0.79^c		1.38^c
32	1.65^c	1.65^c	1.57^c	0.79^c		1.40^c
33	2.55^c	2.55^c	1.93^c	1.75^c		2.17^c
34		1.11^c		0.95^c	0.95^c	
35		1.62^c	0.87^c	1.11^c	1.37^c	
36						

Appendix Table 1. (Continued)

	P_{49}	P_{50}	P_{51}	P_{52}	P_{53}	P_{54}
c_j	-44.80[c]	-9.25	-9.25	-7.00	5.00	0.50
1						
2						
3						
4		3.93	3.93			
5						
6						
7						
8						
9						
10						
11	0.65	1.0	1.0			
12	0.65	1.0	1.0			
13						
14						
15	1.0					
16	1.0					
17	44.80	2.69	2.69			
18	60.0		30.0			
19	6.40	30.0	20.0			
20						
21	2.0					
22		-0.79	-0.79			
23	2.240[c]	16.925	16.925	-10.0	10.0	
24		-12.0	-12.0	10.0		
25	-22.0	-0.86	-0.86			
26						
27	14.44[c]	30.00[c]	30.00[c]			1.0
28	4.90[c]	15.66[c]	15.66[c]			
29	0.90[c]	4.02[c]	4.02[c]			
30	1.37[c]	1.77[c]	1.77[c]			
31	1.38[c]	0.84[c]	0.84[c]			
32	1.40[c]	0.72[c]	0.72[c]			
33	2.17[c]	1.95[c]	1.95[c]			
34	0.95[c]	1.95[c]	1.95[c]			
35	1.37[c]	3.09[c]	3.09[c]			
36						

Appendix Table 1. (Continued)

	P_{55}	P_{56}	P_{57}	P_{58}	P_{59}	P_{60}
C_j	1.518	-1.786	-8.68	-6.90	-10.36	-6.90
1						
2						
3						
4						
5						
6						
7			-1.0			
8				-1.0		
9					-8.0	
10						-8.0
11						
12						
13						
14						
15						
16						
17	1.0	-1.0				
18						
19						
20						
21						
22						
23		.1786	25.048	9.090	29.896	9.090
24						
25						
26						
27						
28						
29						
30						
31						
32						
33						
34						
35						
36						

Appendix Table 1. (Continued)

	P_{61}	P_{62}	P_{63}	P_{64}	P_{65}	P_{66}
C_j	-13.16	-4.24	-0.88	-5.81	-4.95	-1.50
1						
2						
3						
4						
5						
6						
7	-1.0					
8						
9	-8.0					
10						
11		-1.0				
12		-1.0				
13			-1.0		1.0	
14			-1.0		1.0	
15				-1.0	-1.0	
16				-1.0	-1.0	
17						
18						
19						
20						
21						
22						
23	37.976	7.131	0.808	5.328	4.995	0.150
24			-0.404	-2.664	-2.498	
25						
26						
27						-1.0
28						-1.0
29						
30						
31						
32						
33						
34						
35						
36						1.0

Appendix Table 1. (Continued)

	P_{67}	P_{68}	P_{69}	P_{70}	P_{71}	P_{72}	P_{73}
C_j	-1.50	-1.50	-1.50	-1.50	-1.50	-1.50	-1.50
1							
2							
3							
4							
5							
6							
7							
8							
9							
10							
11							
12							
13							
14							
15							
16							
17							
18							
19							
20							
21							
22							
23	0.150	0.150	0.150	0.150	0.150	0.150	0.150
24							
25							
26							
27	-1.0	-1.0	-1.0	-1.0	-1.0	-1.0	-1.0
28							
29	-1.0						
30		-1.0					
31			-1.0				
32				-1.0			
33					-1.0		
34						-1.0	
35							-1.0
36	1.0	1.0	1.0	1.0	1.0	1.0	1.0