a.E.a.(pd)

Staff Contribution 9-1

Miscellaneous Staff Contribution
of the
Department of Agricultural Economics

Purdue University Sch. of agric
Lafayette, Indiana

For information concerning additional available publications write: Librarian, Department of Agricultural Economics

7-1

LINEAR PROGRAMMING APPLIED TO LOCATION OF AND PRODUCT FLOW DETERMINATION IN THE TOMATO PROCESSING INDUSTRY

A. Robert Koch and Milton M. Snodgrass
Purdue University

The authors are, respectively, Assistant Professor of Agricultural Economics at Rutgers University (formerly graduate assistant at Purdue University) and Assistant Professor of Agricultural Economics at Purdue University. This paper is approved as Purdue Journal Paper No. 1378 and is based on research completed under Project #838 in the Agricultural Experiment Station, Purdue University. The authors acknowledge the assistance of Dr. Vernon W. Ruttan who initiated the study and provided generous criticism.

Introduction

At the Philadelphia meetings last year Professors Isard[1] and Stevens[2] presented papers in which they developed some general interregional linear programming models. In these papers, little attempt was made to solve specific problems. Stevens says ". . .we do not create a general spatial model to solve specific problems. We construct it for its value in the understanding of the spatial economic system. . ."[3]

This paper will attempt to use the transportation model which is a special case of the linear programming technique to investigate a few specific interregional problems relating to the tomato processing industry in the United States. The use of the transportation model is well established for tackling various interregional problems. For example, the Koopmans-Hitchcock[4] transportation model was one of the first problems to be solved by linear programming. Samuelson[5] has employed the transportation model in a more generalized form in his work on spatial price equilibrium. In agriculture, Judge[6] has employed the spatial equilibrium model to study interregional competition in the egg industry and the authors used the transportation model of linear programming to study the dairy industry.[7] Beckmann and Marschak[8] have applied the model to firm analysis. Koopmans and Beckmann[9] have more recently attacked plant location problems

To review briefly, the formal characteristics of the transportation model are (1) one unit of any input can be used to produce one unit of any output, (2) the cost of converting one unit of input into one unit of output is constant regardless of the number of units converted, and (3) total inputs must equal total outputs. While these characteristics seem to be rather limiting, various devices can be employed to allow a wide range of problems to be formulated under the transportation model. Additional "artificial" vectors may be added in order to equalize inputs and outputs or to effectively separate different types of inputs or outputs. High conversion costs can be assigned to eliminate impossible or unwanted processes. Other modifications are also possible.

The Transportation Model

The problem usually conceived is one of minimizing the transportation cost of shipping goods from several production points to several warehouses or consumption points. This simple problem involves the minimization of total transport costs denoted by $\sum x_{ij}t_{ij}$ where x_{ij} denotes the amount shipped from production point i to consumption point j and t_{ij} represents the cost of transporting one unit of product from production point i to consumption point j. Constraints of the problem are:

$$x_{11}+x_{12}+x_{13}+x_{14} \ldots + x_{1n} = P_1$$
$$x_{21}+x_{22}+x_{23}+x_{24} \ldots + x_{2n} = P_2$$
$$\ldots$$
$$x_{n1}+x_{n2}+x_{n3}+x_{n4} \ldots + x_{nn} = P_n$$

and

$$x_{11}+x_{21}+x_{31}+x_{41} \ldots + x_{n1} = C_1$$
$$x_{12}+x_{22}+x_{32}+x_{42} \ldots + x_{n2} = C_2$$
$$\ldots$$
$$x_{1n}+x_{2n}+x_{3n}+x_{4n} \ldots + x_{nn} = C_n$$

where P_i equals amount of production at point i
C_j equals amount of consumption at point j

Thus, predetermined variables include the p_i's, C_j's, and t_{ij}'s. The number of x_{ij}'s that need appear for a valid solution is one less than the $\sum P_i$'s + C_j's.

The Model Solution

For purposes of analysis, several aspects of the model solution are pertinent. These include total cost, shipment pattern, marginal values, and equilibrium prices.

The first two are quite simple. Total cost is merely the $\sum x_{ij}t_{ij}$ which indicates the least possible cost. The shipment pattern is merely the analysis of the pertinent x_{ij}'s that appear in the least cost solution.

Marginal values - From any solution, a marginal value (one for each constraining vector) are obtained. If p_i indicates the marginal value for constraints P_i and c_j for C_j, then for each x_{ij} that appears as part of the solution, $p_i + c_j = t_{ij}$. The marginal values (p_i's) indicate the saving or cost of shifting one unit of production among P's and the values (c_j's) indicate the saving or cost incurred in shifting one unit of consumption among C's. For example, if $p_4 = 6$ and $p_8 = 2$, then a shift of one unit of production from p_4 to p_8 would reduce $x_{ij}t_{ij}$ by $4 ($p_4 - p_8$). A shift in the opposite direction (P_8 to P_4) would similarly increase $x_{ij}t_{ij}$ by $4. Similar results occur from the consumption constraints.

Equilibrium prices - From any solution a set of equilibrium prices can be computed. If these prices existed at the various production and consumption points, they would generate the optimum shipment pattern. Equilibrium prices are calculated by choosing a base price in any exporting region. The base price assigned should represent actual prices in the base region. For example, suppose a base price of $2 was established in exporting region P_4. If in the optimum solution, P_4 exported to P_6 and if the cost of transportation was $.30, then the equilibrium price in region P_6 would be $2.30 (($P_4$ price ($2) + t_{46} ($.30)). Continuing on, if exporting region P_2 also exported to P_6 in the optimum solution and t_{26} equaled $.50, then the equilibrium price in region P_2 would be $1.80 (($P_6$ price ($2.30) - t_{26} ($.50)). After the calculation of equilibrium prices, they can be compared to actual prices for purposes of analysis. To ascertain the conformity of the equilibrium price structure with that of the actual, a correlation test between the predicted and actual prices can be performed. For example, if an r^2 equal to .50 resulted, the coefficient would suggest that one half of the variation in the geographical price structure is attributable to considerations other than those allowed for in the model. Therefore, if these other considerations can be quantified and included in the model, and if the r^2 value increases as a result, then it would follow that the model is approaching reality and qualitative factors are successfully being quantified.

Characteristics of the Tomato Processing Industry

The production of tomatoes for processing has changed considerably during the past twenty years. Acreage has been reduced, but increased yields per acre, improved varieties, favorable prices and better control of disease have made possible an overall increase in output. Location shifts of some magnitude have accompanied the acreage and production trends. Raw product production has been shifting into California while other producing regions have declined in relative importance. The number of tomato processing establishments on a national basis has declined; however, in the Midwest and South, the number of processing establishments has increased slightly. The size of processing establishments measured by the average number of employees per plant increased considerably for the United States during the 1939-1957 period. California experienced phenominal growth in the size of processing establishments, while processing establishments located in Indiana have generally decreased in size. The specific location shifts mentioned above are but a few of the more general changes that have occurred within the tomato processing industry.

When an industry undergoes the "pangs" of growth and spatial location shifts, the questions of efficient resource allocation arises. Have resources been relocated within the industry so that production and shipment patterns approach an optimal situation? Interregional product flows and price equilibriums developed through use of the linear programming transportation model form a basis from which to evaluate resource allocation within the tomato processing industry.

Data

The validity of any conclusions pertaining to resource allocation within the tomato processing industry depend in part on the data utilized in the program.

When deciding upon the amount of basic data to include in the model one is torn between the ends of reality of manageability. A rigorous procedure would necessitate the investigation of the spatial relations between each producer and each of his consumers. However, the magnitude of the problem makes this procedure impractical. Working with small aggregates of the data makes the computational burden of the problem manageable. Since the interest in most interregional problems lies in goegraphic flows, aggregation of data should take place on a spatial basis so as to retain a maximum amount of information in regard to any particular state, region, or area. In the end, the optimum degree of aggregation will be a function of the problem to be solved and the resources possessed by the research worker.

Production estimates (P_i's) for canned tomatoes, catsup, and tomato juice were calculated for all states that produced canned tomatoes, catsup or tomato juice. In order to make production estimates that were not biased by the fact that any one state or region experienced abnormal output variation relative to other regions for a given year, average production coefficients were calculated for the period 1951-1956. Production estimates were calculated on a standard unit basis, i.e., #303 can size equivalent for canned tomatoes, 14 ounce bottle equivalent for catsup, and #3 cylinder can size equivalent for tomato juice. Quality differences were included for fancy, extra standard, and standard grades.

Consumption estimates (C_j's) pose a formidable problem. Consumption estimates should reflect differences in population, incomes, tastes, preferences, ethnic influences and other factors that are basic demand determinants.

Estimates of national canned tomato, catsup, and tomato juice consumption are readily obtainable for the United States. However, these data do not reflect state or regional consumption differences. A study conducted by the U.S.D.A. in 1955[10] analyzed weekly regional household food consumption in the United States by income group. Though the study did not encompass total regional consumption, differences among regions were indicated. Using per capita disposable income as a primary demand shifter in conjunction with the U.S.D.A. weekly data, per capita regional consumption, per capita state consumption, and total state consumption were calculated for canned tomatoes, catsup, and tomato juice. The methodology used to calculate state consumption estimates for canned tomatoes, catsup, and tomato juice should be useful to calculate the consumption of other commodities where meaningful state consumption differences are desired.

Transportation cost data are obviously important in any interregional study. There were assumed to be no transportation costs for output consumed intra-state. Transport costs between states or regions were computed on a per case basis for canned tomatoes, catsup, and tomato juice shipped by rail. Since most of the interregional shipments of canned tomatoes and tomato products are by rail, the transportation cost matrix (t_{ij}'s) does not deviate greatly from reality. A complete structure of transportation costs could be incorporated into the model by including vectors for water transportation from the West and truck rates for short hauls. The inclusion of these vectors would depend on the importance of the modes of transportation for the particular product.

The matrix formulation may be abstract or have varying degrees of reality depending upon the data included. The degree of aggregation will depend on the problem to be answered and the resources that are available. The first model concerning the tomato processing industry included all forty-eight states. The same model was also solved with the forty-eight states aggregated into twenty regions. Both the non-aggregated and aggregated models yielded the same general solution, both with regard to cost and shipment pattern. Thus, the aggregation of the data included in the matrix not only reduced computational time but also served as a check on the non-aggregated solution.

Modifying the Transportation Model to Approach an Imperfect Market

Previous studies pertaining to the tomato processing industry have indicated imperfect competition in the market structure.11/ A realistic model dealing with interregional competition within the tomato processing industry should reflect elements of imperfect competition. The transportation model as generally used yields answers for a perfectly competitive situation. Generally this means that (1) products are homogeneous, (2) there are such a large number of buyers and sellers that no one can influence price, (3) there is a complete resource mobility, (4) there is complete knowledge on the part of buyers and sellers, and (5) the price of a given product in importing regions common to a specific exporter will vary only by the cost of transportation, from exporter to importer. The transportation model can be solved for a perfectly competitive situation and the solution compared with exisiting distribution patterns. Conclusions made from this type of comparison would point up the divergence of the existing distribution patterns from a highly abstract market situation. Since the market structure of the tomato processing industry is not perfectly competitive, it would be desirable for the formulated model to approach the actual situation.

Comparisons made between the model solution and actual situation would in this case, be more meaningful to determine if resources have been effecitently allocated within the industry.

In order to make the formulated model as realistic as possible, the following factors affecting production, consumption, and the market structure of the tomato processing industry were quantified and included: (1) processing costs, (2) consumer preference, (3) product differentiation, and (4) pricing inefficiencies.

Processing Costs - Computing cost data for a given product is often a laborous endeavor. Also, cost data are seldom complete enough to make meaningful comparisons. Since the problems of assembling reliable cost data have at times outweighed the value of such computations, an implicit assumption of many models dealing with interregional competition has been that production or processing costs are equal between producing regions. Production costs refer to those costs incurred in the production of tomatoes for processing whereas processing costs refer to those costs incurred in the conversion of raw tomatoes into canned tomatoes, catsup, and tomato juice. The calculation of production and processing costs is a study in itself.[12/] This paper will merely show how production costs and processing costs can be included in the transportation model.

It may seem that the inclusion of cost data in the transportation model is an unnecessary refinement; however, both for the sake of reality and the belief of the authors that regional production and processing costs are important, they have been included in the model formulation. The importance of regional production and processing costs lies in the magnitude of the cost differences between regions. An exporting region with a relative transportation cost disadvantage may overcome this disadvantage if the processing costs of the exporting region are relatively low as compared to other exporting regions.

Several ways to measure the differences between regional processing costs are possible. Regional processing costs may be represented by the regional f.o.b. price of the product. This method of representing processing costs assumes that the regional cost of processing canned tomatoes is reflected by the regional f.o.b. price and that the magnitude of regional differences in f.o.b. price is equal to the magnitude of regional processing cost differences. Another method by which to measure the differences between regional processing costs is to calculate actual processing costs for each state or region. Processing costs calculated in this manner would necessarily be average costs but should show the magnitude of cost differences between each region and be representative of the regional costs. The effect of future regional cost differences can be determined by modifying the regional cost estimates. For instance if it is believed that the cost structure for an enterprise within a given region will increase or decrease relative to the cost structure of other regions the cost structure in the given region can be increased or decreased relative to the other regions and the model solution analyzed to determine the probable effect on interregional product flows. The inclusion of processing costs in the matrix is accomplished by modifying the t_{ij}'s. For example the cost of moving a unit of input from region *i* to region *j* now includes two cost elements (transportation and processing). The cost of processing a unit in region *i* is merely added to the transportation cost t_{ij}.

In the models formulated for the tomato processing industry, the inclusion of processing costs did not affect the optimal distribution patterns developed in models where processing costs differences were not included. Evidently, regional processing cost differences were not large enough to offset relative transportation cost disadvantages.

Consumer Preference -- If it is known that a certain form of a given product is preferred by consumer to other forms of the same product, it would be desirable for the formulated model to show that the preferred product could move more readily into interregional trade as compared to the nonpreferred product. In this case, the preference of one form of the product over another may be looked upon as a cost to the supplier of a nonpreferred product. Viewing preference as a cost to the supplier of a nonpreferred product makes possible the inclusion of consumer choice, which is generally qualitative in nature in the transportation model. Either the cost (t_{ij}) of the preferred product can be reduced by some cost factor to represent the desire for the preferred product or the transportation cost of the nonpreferred product can be increased by some factor to represent the dislike for the nonpreferred product. In the study dealing with the tomato processing industry, the transportation costs (t_{ij}'s) for canned tomatoes moving from producing regions exclusive of California were increased by 10 cents to represent the willingness to purchase California canned tomatoes over the canned tomatoes from the other producing regions.

Product Differentiation -- When qualitative information pertaining to the production of given grades or qualities of a product is available, the production estimates (P_i's) can be modified to reflect product differentiation. Likewise if information is available which shows consumption differences by grade or qualtiy a more realistic model can be developed through modification of the consumption estimates (C_j's). Product differentiation in the form of advertising can be incorporated into the model by also modifying the (t_{ij}'s). The advertising cost per case would be added to the transportation costs. Upon solving the matrix the restrictive nature of advertising on the distribution pattern would be indicated.

Pricing Inefficiencies -- Third degree price discrimination where markets are spatially separated and different prices are paid for the product depending upon the elasticity of the market demand curves is one form of pricing inefficiency that can be built into the transportation model. If estimates of market area boundaries can be made and if differences in price between the markets are indicated, then the transportation costs (t_{ij}'s) from a producing region that practices discriminatory pricing can be modified to include price differences in the separated markets.

Modifying the Model to Study Production Location

To study production location, the P_i constraints are modified and "artifical" vectors may be introduced to absorb excess production (idle capacity). Optimum location of producing areas for tomato production was determined under three different situations: (1) unlimited production in every region, (2) specification of a maximum production in each region, and (3) specification of a minimum production in each region. In the first situation, since the constraint of the P_i's is removed, the problem is no longer a programming problem. Consumption is merely satisfied from a producing region that minimizes total cost. The latter two situations impose modified restraints on the P_i's and necessitate the introduction of "artifical" vectors. All tomato products are aggregated to a common raw product equivalent base. Once the problem is solved, the raw product equivalents are reconverted for each product.

Summary and Conclusions

The three major areas in the tomato processing industry are California, the Midwest (Indiana, Illinois, and Ohio) and the Tri-States (New Jersey, Delaware, and Maryland). Major products considered were canned tomatoes, tomato juice, and catsup since these products make up approximately 80 percent of the output in the tomato processing industry. Three grades of product were considered: fancy, extra standard, and standard.

To summarize all of the findings of the study would be too extensive for this paper; however, the more important generalizations should be mentioned. The results of the model variations indicated:

(1) Actual Midwest catsup and juice market areas should be contracted in the West and expanded South and East.

(2) The actual Midwest canned tomato market area should be expanded in the South and contracted in the East.

(3) The TriStates should generally follow the current distribution patterns for canned tomatoes, catsup and juice.

(4) California should dominate actual shipments of the three products west of the Mississippi river and also ship juice and catsup into the Gulf States and New England. Depending on the grade of tomatoes California canned tomato shipments should reach all segments of the country.

The production marginal values for the catsup distribution patterns indicate that if production were to shift between regions an East to West production shift would reduce aggregate cost. If juice production were to shift between regions, the production marginal values indicate a West to East production shift would reduce aggregate cost. If a re-alignment of canned tomato production w re to occur between regions, fancy canned tomato production should shift from the Tri States to Indiana while extra standard and standard canned tomato production should shift from the Tri States to the West. When transportation costs were alone relevant in the cost structure, the production marginal values reflected the relative transportation disadvantage of the West. However, when actual processing costs and transportation costs were relevant in the cost structure, the relatively lower processing costs in the West overcame the relative transportation cost disadvantage to a considerable degree as shown by the narrower range of the marginal values.

Throughout the analysis, the actual prices received by California canned tomatoes, catsup, and juice processors were much higher than the equilibrium prices developed in the models. This difference between the California actual and equilibrium prices indicates possible pricing inefficiencies in the tomato processing industry. The Western region dominates production so conclusively that the price in the other producing regions appears dependent upon the surplus production moving out of the West. "Responsible" pricing policies also seem to influence the pricing mechanism in the catsup phase of the industry. The fear of "breaking the market" and the beliefs that price increases will not be followed, but that price decreases will be followed are probable causes of the catsup pricing policy.

When the regional production of canned tomatoes, catsup, and juice was not a predetermined variable (production vector constraints removed), the actual production patterns changed considerably. Assuming unlimited or maximum specified production, the Midwest became the most important region as a supplier of tomatoes for processing, canned tomatoes, catsup and tomato juice. California production was decreased substantially. When minimum regional production was specified the production patterns developed from the model solution generally agreed with existing production patterns. However, some important production changes were noted. They are: California catsup production increased to three-quarters of the national output, while juice production decreased. The Tri-States became the most important juice supplier but ceased the production of catsup entirely.

The equilibrium prices were correlated with the actual prices for all models to determine if imperfect elements in the market had been quantified as these elements were included in the models. When perfect competition was assumed there was no correlation between the actual and equilibrium prices, but when the models were varied the equilibrium prices and actual prices for each product were correlated in varying degrees. For instance the equilibrium price correlation for catsup (a perfectly competitive situation) showed an r^2 or .0415 but when product differentiation was included in the model the r^2 for fancy catsup increased to .890 which signified that approximately 11 percent of the fancy catsup price variation is explained by factors other than those included in the model. The extra standard catsup equilibrium and actual prices were correlated to a lesser degree. Thus, the authors believe that the transportation model can be varied to approach a realistic situation and the validity of the imposed modification can be ascertained from the model solution and the existing regional prices. In general, the usefulness of the results obtained in attacking certain inter-regional problems using programming methods are probably more limited by in-adequate data than by the analytical model.

Footnotes

1/ Walter Isard, "Interregional Linear Programming: An Elementary Presentation and a General Model", *Journal of Regional Science*, Vol. 1, No. 1, Summer, 1958, pp. 1-59.

2/ Benjamin H. Stevens, "An Interregional Linear Programming Model", *Journal of Regional Science*, Vol. 1, No. 1, Summer 1958, pp. 60-98.

3/ Ibid, p. 89.

4/ F. L. Hitchcock, "The Distribution of a Product from Several Sources to Numerous Localities", Journal of Mathematics and Physics, M.I.T., Vol 20, 1941, pp. 224-230; and T. C. Koopmans, "Optimum Utilization of the Transportation System", proceedings of the International Statistical Conference, Washington, D.C., 1947 Vol. 5, (reprinted as a Supplement to Econometrica, Vol. 17, 1949).

5/ P. A. Samuelson, "Spatial Price Equilibrium and Linear Programming," *American Economic Review*, Vol. 42, June 1952, pp. 283-303.

6/ G. G. Judge, "Competition Position of the Connecticut Poultry Industry — 7. A Spatial Equilibrium Model for Eggs", Bulletin 318, Storrs Agricultural Experiment Station, University of Connecticut, January 1956.

7/ H. H. Snodgrass and C. E. French, "Linear Programming Approach to Interregional Competition in Dairying," Station Bulletin #637, Agricultural Experiment Station, Purdue University, May 1958.

8/ M. Beckmann and T. Marschak, "An Activity Analysis Approach to Location Theory", *Kyklos*, Fasc. 2, 1955.

9/ T. C. Koopmans and M. Beckmann, "Assignment Problems and the Location of Economic Activities", *Econometrica*, Vol. 25, January 1957, pp. 53-76.

10/ Household Food Consumption Survey in the United States (Reports 2-5) U.S.D.A., Washington, D.C., 1955.

11/ J. F. Ritchey, The Marketing of Indiana Tomatoes and Tomato Products, Unpublished M.S. Thesis, Purdue University, January 1958, pp. 100-132.

12/ W. L. Fishel, "Cost and Efficiency for Indiana Tomato Canning Plants", Unpublished M.S. Thesis, Purdue University, Janauary 1958. Also see V. W. Ruttan and W. L. Fishel, "Cost and Efficiency in Indiana Tomato Canning Plants", Research Mimeo ID-25, Agricultural Experiment Station, Purdue University, June 1958.