Economies of Scale in Cottonseed Processing

By R. McFall Lamm, Jr.*

Over the last two decades, the number of U.S. cottonseed processing plants has declined drastically. The realization of scale economies by expanding plant size is shown to be a basic reason for this decline. The analysis is performed by estimation of a nonhomothetic translog cost function to represent the industry structure. Statistical tests indicate that the translog function is appropriate. Other results indicate that derived demands for inputs in cottonseed processing are inelastic and that economies of scale are greater for smaller plants in the industry.

Keywords:

Cottonseed processing
Translog cost function
Scale economies

INTRODUCTION

Decreasing numbers of firms in food processing industries are frequently considered as evidence of expanding market power. Declining firm numbers generally occur, however, in industries where firms face positive scale economies, or where technical change permits cost reduction. Regulatory economists argue that social benefits from the realization of scale economies or from technical change have been exhausted for most industries and that declining firm numbers result principally from market power. One proposed solution calls for Government intervention to maintain firm numbers at levels greater than would otherwise result. Thus, for industries in which firm numbers are declining, the level of firm scale economies is important in determining whether declines in firm numbers result from natural competitive processes or from the exercise of market power by some firms in the industry.

THE PROBLEM

Over the last 20 years, firm and plant numbers have greatly declined in the domestic cottonseed processing industry. This decline is usually attributed to the realization of positive scale economies. A study by Stuart and Morrison (*9*), indicating positive scale economies for soybean processing plants, is frequently referenced as evidence of scale economies in cottonseed processing because of the similarity of the two production processes.[1] No one has attempted to determine the nature of economies in cottonseed processing, however.

THE OBJECTIVES

The domestic cottonseed processing industry was studied to determine whether scale economies exist, and results are presented in this article. Economies are measured by estimating translog cost functions through use of classified ESCS data for individual firms in the industry. The translog functional form allows estimation of nonhomothetic and nonhomogenous cost functions, and it provides a means for determining derived demand price elasticities.

Characteristics of the cottonseed processing industry are examined and evidence presented on the decline of firm numbers. Next, the postulated decision process assumed to underly the construction of the model is explained, and followed by the model specification. Data used, the results, and conclusions complete the article.

THE INDUSTRY

Cottonseed processing, a multimillion dollar industry in the United States, employs thousands of people and produces four major products—cottonseed linters, cottonseed oil, cottonseed hulls, and cottonseed meal. These products are produced in processing plants generally referred to as cottonseed oil mills because the oil provides the major share of the revenue. Most of these plants are located in the major cotton producing areas of the country.

Uses

Better grades of cottonseed linters form the stuffing material in mattresses, pads, and cushions. Lower grades, referred to as "second cuts," are a source of cellulose. After being refined, cottonseed oil becomes an ingredient in cooking oil, shortening, margarine, and other foods. The hulls are used as a roughage ingredient in animal feeds, and the meal constitutes a protein source in feeds.

---

*The author is an agricultural economist in the Commodity Economics Division, ESCS. Opinions presented to do necessarily represent those of USDA. Assistance of Elmer Purdue and Stan Thurston in obtaining data is gratefully acknowledged, as are Roger Spindler's research assistance and comments of Duane Hacklander and the reviewers.

[1] Italicized numbers in parentheses refer to items in References at the end of this article.

"Although some decline would be expected, given decreasing cotton production, the drop in plant numbers has been much greater than would have been necessary to compensate for decreased production."

Output and Structure

Annual domestic cottonseed production depends on the level of cotton production. Over the last 20 years, production has fallen gradually because of foreign competition and increased use of manmade fibers. The number of firms processing cottonseed has similarly declined (*8*). A substantial drop occurred from 1952 through 1972 (*10*). Although some decline would be expected, given decreasing cotton production, the drop in plant numbers has been much greater than would have been necessary to compensate for decreased production. Consequently, the average size of cottonseed processing plants has increased over time.

From 1954 to 1972, the number of plants in the industry declined from 286 to 115 (table 1). Cottonseed oil output per plant (a measure of plant size) increased from 3.52 million pounds to 11.24 million pounds. Total employment per plant increased slightly, from 47.8 employees in 1954 to 48.7 in 1972.[2]

Table 1—Selected industry statistics, selected years, 1954-72

Year	Number of plants	Per plant		
		Oil output	Employees	Production workers
		Mil. lbs.	*Number*	
1954	286	3.52	47.8	38.7
1958	214	6.71	39.0	29.5
1963	188	10.33	44.5	34.9
1967	150	7.38	36.0	28.0
1972	115	11.24	48.7	39.1

Source: U.S. Department of Commerce. *Census of Manufactures.*

The increase in oil output per plant represents an increase in size of over 300 percent in 18 years, approximately 17 percent per year.

POSTULATED DECISION PROCESS

In the study, cottonseed processors were assumed to maximize profit subject to a multi-output, multi-input production function. Because cottonseed products are jointly produced, one product can be isolated for study to simplify the analysis. With input levels as the relevant decision variables, profit maximization yields rules for optimum input decisions. Substituting these optimal decision rules into the firm profit identity under competition yields the cost function:

$$c = \phi\ (q, r) \qquad (1)$$

where c is total cost, q is scalar output, and r is a vector of input prices.

For the cottonseed processing industry, q represents cottonseed oil output—the cottonseed product generating the largest percentage of total revenue. Cottonseed (m), labor (w), and capital (k) are the inputs considered.[3] Because the oil generally is sold to regional food processors with specialized needs and because transportation costs are relatively high, q is largely predetermined. In addition, sales offices usually decide on output levels but typically do not make the actual production decisions. Given q, the plant managers attempt to minimize cost by choosing the appropriate input levels. Consequently, the postulate of predetermined q seems reasonable. If q is not predetermined, a solution to the firm's decision problem yields $c = \phi\ (p, r)$, where p is a vector of output prices, which implies that relation (1) is misspecified.

[2] Cottonseed production and crushings reached their lowest level in more than 20 years in 1967, which explains why oil output was low that year.

[3] One could also eliminate cottonseed input as a decision variable, using an identity to relate oil production to oilseed input. This would be equivalent to assuming perfectly inelastic input demand for oilseeds and zero partial elasticities of substitution between oilseeds and other inputs. In the approach used here, a more general one, these restrictions are not imposed arbitrarily.

MODEL SPECIFICATION

To carry out the analysis, relation (1) is approximated as a translog function proposed by Christensen-Jorgenson-Lau (*4*). This function has been used widely in recent studies to approximate nonlinear forms. It is especially appropriate here because it allows the estimation of nonhomothetic and nonhomogeneous cost functions. Translog cost functions were first applied by Binswanger in a study of agricultural production (*2*).

The translog form of relation (1) is written as:

$$lnc = \alpha_o + \alpha_q \ln q + \sum_i \gamma_{qi} ln\, q \ln r_i + \sum_i \alpha_i ln\, r_i$$

$$+ 1/2 \sum_i \sum_j \gamma_{ij} ln\, r_i \ln r_j + 1/2\, \gamma_{qq} (lnq)^2 \quad (2)$$

where q is defined as scalar output of cottonseed oil, the α_i, γ_{ij}'s are structural parameters of the cost function, and the logarithms are natural. Symmetry ($\gamma_{ij} = \gamma_{ji}$) is imposed unconditionally. In addition, homogeneity in prices is imposed unconditionally:

$$\sum_i \alpha_i = 1 \quad (3)$$

$$\sum_i \gamma_{qi} = \sum_i \gamma_{ij} = \sum_j \gamma_{ij} = \sum_i \sum_j \gamma_{ij} = 0 \quad (4)$$

where $i,j = m,k,w$, for cottonseed, capital, and labor input, respectively.[4]

By noting that $c = r'x$, where x is a vector of inputs, it is apparent that $d\, c/d\, r_i = x_i$. In this respect:

$$ln\, c/d \ln r_i x_i/c = s_i \quad (5)$$

where s_i is the cost share of the ith input. The translog cost function yields the cost share equations:

$$s_i = \alpha_i + \gamma_{qi} \ln q + \sum_j \gamma_{ij} ln\, r_j \quad i = m,k,w \quad (6)$$

[4] See Binswanger (*2*, *3*) for a discussion.

Requiring (5) to hold unconditionally implies a representation of the cost structure consisting of relation (2) with three cost-share relations of the type specified in equation (6).

The translog cost function is nonhomothetic and nonhomogeneous.[5] Homotheticity may be imposed by requiring:

$$\gamma_{qi} = 0 \qquad i = m,k,w \quad (7)$$

Homogeneity is imposed by requiring that:

$$\gamma_{qq} = 0 \quad (8)$$

In addition, unitary elasticities of substitution may be imposed by restricting $\gamma_{ij} = 0$, in which case the translog function reduces to a Cobb-Douglas form.

Following Christensen and Green (*5*), the measure of scale economies used in this study is defined as:

$$\sigma = 1 - d\, lnc/d\, ln\, q \quad (9)$$

When $\sigma > 0$, economies of scale exist; when $\sigma < 0$, diseconomies of scale exist. For the general nonhomothetic cost function:

$$\sigma = 1 - (\alpha_q + \gamma_{qq}\ lnq + \sum_i \gamma_{qi} \ln r_i) \quad (10)$$

When homotheticity is imposed:

$$\sigma = 1 - (\alpha_q + \gamma_{qq}\ lnq) \quad (11)$$

When both homotheticity and homogeneity are imposed:

$$\sigma = 1 - \alpha_q \quad (12)$$

[5] In a homothetic cost function, cost curves in input-price space exhibit disproportionate shifts as output varies. In a homogeneous cost function, cost curves in input-price space exhibit proportionate shifts as output varies.

Nonhomogeneous cost functions yield different scale measures for each size of firm, while the homogeneous scale measure (*12*) yields one value for all sizes of firms in the industry.

THE DATA AND ESTIMATION PROCEDURE

The classified ESCS data used in the study came from a combined time-series, cross-section sample collected over a 20-year period. Because scale economies must be distinguished from technical change, an intertemporal phenomena, cross-sectional data must be used to estimate the industry structure. Unfortunately, degree of freedom limitations prohibit the use of any single cross section from the sample.

As an alternative, two combined time-series, cross-sectional data sets were selected from the sample. The first, data set I, consists of 44 firm observations from 1960, 1961, and 1962. Close inspection of detailed data on these firms indicated that they had neither added to existing facilities nor replaced existing technologies with new technologies during 1960-62. For practical purposes then, little or no identifiable technical change occurred in the data set. The second, data set II, consists of 39 firm observations from 1972, 1973, and 1974.[6] Again, careful examination of the firms indicated little evidence of technical change.

Except for capital costs, all data varied cross sectionally by firm. Cottonseed prices (dollars per ton) were obtained directly from the classified ESCS source. The price of capital was proxied simply as the rate on corporate AAA bonds. Wage rates (dollars per year) were obtained from census data for each State. Total cost (dollars per year), cottonseed oil output (pounds), and the input data necessary to compute cost shares, also came directly from ESCS classified sources.

The cost structure system was estimated with the iterative Zellner procedure used by Christensen-Green (*5*), Griffin (*6, 7*), and Berndt-Christensen (*1) to* estimate similar systems. This procedure, which converges to maximum likelihood estimates, requires that one cost-share equation be dropped from the system to obtain nonsingularity in the variance-covariance disturbance matrix.[7]

Because the iterative Zellner procedure converges to maximum likelihood estimates, the likelihood ratio test may be used to test for homotheticity and homogeneity (functional form). The relevant test statistic, as described by Christensen-Green (*5*), is:

$$\lambda = (|\Omega_c|/|\Omega_u|)^{-n/2} \qquad (13)$$

where λ is the value of the likelihood ratio, $|\Omega_c|$ is the determinant of the conditionally constrained (homothetic or homogeneous) variance-covariance matrix, $|\Omega_u|$ is the determinant of the unconstrained (nonhomothetic) variance-covariance matrix, and n is the number of observations. Tests were performed by using the fact that $-2\ ln\ \lambda$ has a chi-square distribution with degrees of freedom equal to the number of conditional constraints imposed on the system.

RESULTS OF ESTIMATION

Three forms of the translog cost function were estimated for each data set. The first, model A, consists of the system represented by relations (2) and (6), the nonhomothetic form. The second, model B, consists of the system represented by relations (2) and (6) with conditional constraint (7) imposed. This is the homothetic form. The third cost function estimated, homogeneous in output and referenced as model C, consists of the system represented by relations (2) and (6) with conditional constraints (7) and (8) imposed. The cost-share equation for labor is omitted to allow estimation of each model.

Iterative Zellner estimates of all three models for each data set appear in table 2. The α_i parameters are somewhat sensitive to model specification, much more so than the γ_{ij} substitution parameters. The α_q is especially

[6] These years were chosen based on the number of firms included in each year's data.

[7] Pooled data are treated as independent observations.

Table 2—Estimates of cost function parameters

Parameter	Data set I			Data set II		
	A	B	C	A	B	C
α_0	1.244 (5.747)	2.244 (7.035)	-5.041 (.551)	-3.967 (4.838)	-3.240 (4.855)	-4.202 (.835)
α_q	.215 (.666)	.113 (.817)	.967 (.021)	.883 (.576)	.872 (.578)	.985 (.029)
α_m	.664 (.097)	.794 (.078)	.796 (.079)	.612 (.124)	.761 (.044)	.762 (.044)
α_k	.184 (.785)	.116 (.890)	.116 (.811)	.356 (.207)	.458 (.214)	.456 (.215)
α_w	.152 (.062)	.090 (.059)	.088 (.059)	.031 (.193)	-.219 (.211)	-.217 (.211)
γ_{qq}	.023 (.019)	.024 (.023)		.004 (.017)	.003 (.017)	
γ_{qm}	.009 (.004)			.009 (.005)		
γ_{qk}	-.003 (.003)			.001 (.005)		
γ_{qw}	-.007 (.002)			-.010 (.003)		
γ_{mm}	.102 (.014)	.103 (.016)	.104 (.016)	.061 (.008)	.061 (.008)	.061 (.008)
γ_{mk}	-.122 (.023)	-.129 (.024)	-.130 (.024)	-.080 (.012)	-.080 (.012)	-.080 (.012)
γ_{mw}	-.082 (.013)	-.078 (.015)	-.078 (.015)	-.043 (.007)	-.042 (.009)	-.042 (.009)
γ_{kk}	.038 (.349)	.038 (.346)	.038 (.596)	.042 (.014)	.048 (.016)	.048 (.017)
γ_{kw}	.047 (.011)	.052 (.012)	.053 (.012)	-.004 (.027)	-.016 (.032)	-.016 (.032)
γ_{ww}	.017 (.005)	.013 (.005)	.013 (.005)	.024 (.013)	.029 (.015)	.029 (.016)

Note: Estimates are maximum likelihood; standard errors appear in parentheses below each estimated coefficient. The subscripts are defined as follows: 0 represents the intercept term, q denotes output, m is cottonseed input, k is capital input, and w is labor input.

"The nonhomothetic form of the cost function appropriately represents the industry structure."

sensitive; it changes value considerably when homogeneity is imposed on the system. Because some of the estimated parameters of each model are not highly significant statistically, it is not apparent which of the three models is preferable and acceptable for describing the industry.

A decision concerning acceptability is possible by using the χ^2 test described previously. Table 3 presents values of the test statistic $-2 \ln \lambda$. In each case, model A is taken as the unconstrained form to compute $|\Omega_u|$, and $|\Omega_c|$ is computed with models B and C. Because the computed χ^2 statistics are greater than the critical level with 99-percent confidence, the homothetic and homogeneous forms can be rejected for both data sets. The nonhomothetic form of the cost function appropriately represents the industry structure.[8]

Table 3—Test statistics for homotheticity and homogeneity

Test	Hypothesis	
	Homotheticity	Homogeneity
Restrictions	2	3
Critical χ^2 (1%)	9.21	11.35
χ^2 Data set I	38.28	40.30
χ^2 Data set II	15.795	16.068

Second-order conditions require that the cost function be monotonic and convex in input-price space. These conditions are satisfied if the fitted $s_{ij} > 0$ $i = m,k,w$, $j = 1, \ldots, n$, and the relevant bordered Hessian is negative definite. These conditions are satisfied for all model A observations in both data sets.

Mean estimates of scale economies are presented in table 4 for five groups of firms ranked by output from smallest to largest. For each estimate, t statistics appear in parentheses for the null hypothesis that mean $\sigma = 0$. For both data sets, significant scale economies (t statistics greater than 2, positive values for σ) are

[8] Additional testing of model A against nonhomothetic, and homogeneous Cobb-Douglas cost functions indicated the acceptance of this model.

Table 4—Estimated firm scale economies

Size class	Data set I				Data set II			
	$\bar{\sigma}_a$	$\bar{\sigma}_b$	$\bar{\sigma}_c$	$\bar{q}$	$\bar{\sigma}_a$	$\bar{\sigma}_b$	$\bar{\sigma}_c$	$\bar{q}$
1	0.073 (20.23)	0.094 (23.34)	0.033 (1.57)	8.3	0.028 (11.50)	0.026 (16.54)	0.016 (.55)	5.5
2	.055 (26.46)	.073 (29.07)	.033 (1.57)	12.0	.020 (14.15)	.020 (56.6)	.016 (.55)	11.8
3	.042 (42.30)	.062 (47.25)	.033 (1.57)	16.1	.018 (14.99)	.018 (119.94)	.016 (.55)	16.0
4	.031 (25.71)	.048 (17.34)	.033 (1.57)	20.0	.017 (11.97)	.016 (112.49)	.016 (.55)	20.1
5	−.012 (1.26)	−.003 (.39)	.033 (1.57)	60.4	.008 (2.97)	.010 (5.41)	.016 (.55)	65.2

Note: Oil output is in million pounds, and t-statistics are in parentheses. Subscripts refer to models, and bars denote mean values.

". . . at least partially, the cause of declining firm numbers in the cottonseed processing industry is due to the realization of firm economies within the industry."

indicated for size classes 1 through 4, when the cost function is nonhomogeneous. When homogeneity is imposed, however, no significant scale economies are indicated. Thus, model specification is crucial in determining whether scale economies exist.

Based on group estimates of σ, the largest scale economies exist for the smallest size class of firms for both data sets. For example, using the nonhomothetic form, the value of σ is 0.073 for the smallest size class of data set I. For data set II, the corresponding value is 0.028. As size classes get larger, σ values decline, which indicates less opportunity for reducing average costs by expanding firm size. This behavior conforms to prior expectations. On an intertemporal basis, opportunities for scale economies decreased from 1960-62 through 1972-74, even for smaller sized firms. Even so, significant unexploited scale economies existed in 1972-74.

Given the estimates of the γ_{ij} in table 2, derived demand price elasticities for inputs were obtained using the relation:

$$\eta_i = (\gamma_{ii} + s_i\,(s_i - 1))/s_i^2 \qquad i = m,k,w, \qquad (14)$$

where η_i is the derived demand price elasticity for the *i*th input (*5*). Derived demand price elasticities were computed for cottonseed, capital, and labor used in cottonseed processing (table 5).

The reported derived demand elasticities for all three inputs by size class and data set are all inelastic. This finding is consistent with prior expectations—in highly specialized plants, inputs are demanded with little sensitivity to price. In this respect, the demand for cottonseed is more inelastic than the demand for capital or labor. Firm size has only a limited effect on the level of derived demand elasticities, across size classes, except perhaps for the derived demand for labor in 1972-74. From 1960-62 to 1972-74, derived demands for capital and labor become more inelastic while the converse was true for cottonseed. This behavior is consistent with increased plant specialization in capital and labor use, and greater efficiency in production scheduling in cottonseed use.

Table 5—Derived demand elasticities

Size class	1960-62			1972-74		
	$\bar{\eta}_m$	$\bar{\eta}_k$	$\bar{\eta}_w$	$\bar{\eta}_m$	$\bar{\eta}_k$	$\bar{\eta}_w$
1	-0.09	-0.59	-0.73	-0.15	-0.50	-0.48
2	-.06	-.57	-.73	-.12	-.50	-.45
3	-.07	-.59	-.73	-.11	-.48	-.38
4	-.06	-.57	-.72	-.13	-.51	-.38
5	-.07	-.58	-.71	-.13	-.50	-.37

CONCLUSION

With the potential of increasing cotton production in the years ahead, cottonseed processing will become more important. Over the last 20 years, the number of firms and plants in the industry declined considerably. This has raised concern among many economists that market power might be increasing in the industry.

Based on the acceptance of a nonhomothetic cost function, results indicate that, at least partially, the cause of declining firm numbers in the cottonseed processing industry is due to the realization of firm economies within the industry. This process is most likely a consequence of effective competition, although it is not clear why unexploited scale economies have persisted over time.

Further, the cost structure of the industry is nonhomothetic, input demands are inelastic, and they vary little by firm size. Because of the similarity in the level of measures of scale economies for different size classes, the "flat" region of the average cost curve for firms in the industry is lengthy. Thus, the cottonseed processing industry exhibits a considerable dispersion in firm size.

REFERENCES

(*1*) Berndt, E., and L. R. Christensen. "The Translog Function and the Substitution of Equipment, Structures, and Labor in U.S. Manufacturing 1929-68." *J. Econometrics.* Vol. 1, 1973, pp. 81-113.

(*2*) Binswanger, H. P. "A Cost Function Approach to the Measurement of Elasticities of Factor Demand and Elasticities of Substitution." *Am. J. Agr. Econ.* Vol. 56, 1974, pp. 377-386.

(*3*) ________ "The Measurement of Technical Change Biases with Many Factors of Production." *Am. Econ. Rev.* Vol. 65, 1974, pp. 964-976.

(*4*) Christensen, L. R., D. Jorgenson, and L. J. Lau. "Conjugate Duality and the Transcendental Logarithmic Production Function." *Econometrica.* Vol. 39, 1971, pp. 255-256.

(*5*) Christensen, L. R., and W. H. Green. "Economies of Scale in U.S. Electric Power Generation." *J. Polit. Econ.* Vol. 84, 1976, pp. 655-676.

(*6*) Griffin, J. M. "The Econometrics of Joint Production: Another Approach." *Rev. Econ. and Statis.* Vol. 59, 1977, pp. 389-397.

(*7*) ________ "Joint Production Technology: The Case of Petrochemicals." *Econometrica.* Vol. 46, 1978, pp. 379-397.

(*8*) National Cottonseed Products Association. *The International Green Book.* The Cotton Gin and Oil Mill Press, Dallas, TX, issued annually.

(*9*) Stuart, C., and W. R. Morrison. "Economies of Size in Soybean Processing Plants," Agr. Expt. Sta., Univ. Ark., Bull. 743, Fayetteville, 1969.

(*10*) U.S. Department of Commerce. *1972 Census of Manufactures.* U.S. Govt. Print. Off., Wash., D.C., 1974.

In Earlier Issues

The rapid growth of the population of the United States during the last decade, together with the wider use of statistics relating to the future, have brought about an increase of interest in population projections . . . No official attempt has been made to publish periodically a systematic set of projections for the major geographic subdivisions of the United States . . . Fertility and mortality occur with more statistical regularity than does migration, and to the extent that they do they are more predictable. Because of this, it seems reasonable to conclude that projections of the total population of the United States can be made with more accuracy than projections for geographic subdivisions of the country. [Hagood and Siegel accepted census projections to 1975 of the U.S. population and projected the distribution among the nine census divisions. Their widest relative error was an underestimate by only 1.35 percent of the share that would be in the South Atlantic region in 1975. (ed.)]

Margaret Jarman Hagood
& Jacob S. Siegel
April 1951, Vol. 3, No. 2, pp. 41-52.