

A Translog Cost Function Analysis of U.S. Agriculture:

A Dynamic Specification

By

Charles B. Moss,

Kenneth W. Erickson,

V. Eldon Ball,

and

Ashok K. Mishra^{*}

November 13, 2003

^{*} Charles Moss is a professor in the Department of Food and Resource Economics at the University of Florida. Kenneth Erickson, Eldon Ball and Ashok Mishra are economists with the United States Department of Agriculture, Economic Research Service. The views presented in this manuscript are those of the authors and do not necessarily represent those of the USDA.

Abstract

A Translog Cost Function Analysis of U.S. Agriculture:

A Dynamic Specification

Estimation of dual cost and profit functions has become a standard approach in applied economics since their inception in the early 1970s. However, growing concerns about cointegration have raised questions about the applicability of these dual specifications using time series data. This study applies a dual specification based on the error-correction model that nests the possibility of cointegration within the classical cost function framework. The estimated support the error-correction mechanism, but fall short of non-stationarity. In addition, the error-correction model yields consistent estimates with respect to the concavity of the cost function.

Keywords: dual cost function, error-correction, cointegration.

JEL Numbers: Q11, C32

A Translog Cost Function Analysis of U.S. Agriculture –

A Dynamic and Regional Specification

The advent of the dual (or indirect) approach to production economics has proven a windfall for applied economic analysis in general and to agricultural economics in particular. Using dual cost or profit functions, researchers analyzed changes in agricultural productivity (e.g., Capalbo and Denny 1986, Lambert and Shonkwiler 1995, Lim and Shumway 1997, Huffman et al. 2002) and changes in a multitude of agricultural and trade policies. However, most of the applications of duality have relied on the static formulation of the indirect objective function.¹ This emphasis on static econometric formulations is inconsistent with the growing econometric literature on cointegration (Engle and Granger 1987). Cointegration literature grew out of Granger and Newbold's (1974) finding that regressions between non-stationary time-series could yield spurious regression results. Thus, any dual specification including non-stationary time-series could be tainted by spurious regression problems. In addition to the difficulties raised by non-stationarity, the empirical specification of dual cost and profit functions typically included quasi-fixed inputs (Chambers and Vasavada 1983). The inclusion of quasi-fixed inputs allowed researchers to test long-run economic concepts [e.g., asset fixity in the case of Chambers and Vasavada (1983)] using annual or short-run data. This study examines the implications of the short-run specification of the standard Translog cost specification along with the possible implications of non-stationarity by estimating a dynamic Translog cost specification complete with dynamic share equations for U.S. agriculture using an empirical approach developed by Urga and Walters (2003).

¹ Exceptions to this emphasis include Lambert and Shonkwiler who analyze changes in factor bias over time.

The Translog specification of the dual cost function was initially justified as a second-order Taylor-series approximation to an unknown dual cost function (Berndt and Christensen, 1973). Mathematically, letting $C_t(y_t, w_t)$ denote the minimum cost of producing a given vector of outputs, y_t , given a vector of input prices, w_t , in period t a second order Taylor-Series expansion of the natural logarithm of this unspecified function becomes

$$\begin{aligned} \ln(C(y_t, w_t)) = & \ln(C(\bar{y}_t, \bar{w}_t)) + \nabla_{\bar{y}, \bar{w}} \ln(C(\bar{y}_t, \bar{w}_t)) \begin{pmatrix} \ln(y_t) - \ln(\bar{y}_t) \\ \ln(w_t) - \ln(\bar{w}_t) \end{pmatrix} \\ & + \frac{1}{2} \begin{pmatrix} \ln(y_t) - \ln(\bar{y}_t) \\ \ln(w_t) - \ln(\bar{w}_t) \end{pmatrix}' \nabla_{(\bar{y}, \bar{w}) \times (\bar{y}, \bar{w})}^2 \ln(C(\bar{y}_t, \bar{w}_t)) \begin{pmatrix} \ln(y_t) - \ln(\bar{y}_t) \\ \ln(w_t) - \ln(\bar{w}_t) \end{pmatrix} \end{aligned} \quad (1)$$

where $\nabla_x f(x)$ denotes the gradient (or vector of first derivatives) of function $f(x)$ with respect to the vector x and $\nabla_{xx}^2 f(x)$ denotes the Hessian matrix (or matrix of second derivatives) of function $f(x)$ with respect to the vector x . Assuming that the gradient vector and Hessian matrix are constants and collapsing the geometric constants into an intercept term, this specification becomes

$$\begin{aligned} \ln(C(y_t, w_t)) = & \alpha_0 + \alpha' \ln(y_t) + \frac{1}{2} \ln(y_t)' A \ln(y_t) + \\ & \beta' \ln(w_t) + \frac{1}{2} \ln(w_t)' B \ln(w_t) + \ln(y_t)' \Gamma \ln(w_t) \end{aligned} \quad (2)$$

where α_0 , α , A , β , B , and Γ represent estimated parameters (constants).

Applying Sheppard's lemma to the dual cost function specification in equation (2) yields a vector of cost-share equations

$$\frac{\partial \ln(C(y_t, w_t))}{\partial \ln(w_t)} = \frac{\partial C(y_t, w_t)}{\partial w_t} \frac{w_t}{C(y_t, w_t)} = \frac{x^*(y_t, w_t) w_t}{C(y_t, w_t)} = S(y_t, w_t) \quad (3)$$

where $s_i(y_t, w_t)$ denotes the share of cost expended on input i in the cost-minimizing solution.

Based on the Translog cost function in equation 2, these share equations can be specified as

$$S(y_i, w_i) = \alpha + A \ln(w_t) + \Gamma \ln(y_t) \quad (4)$$

Finally, the dual function specification implies theoretical restrictions on the parameters in the empirical models. Commonly imposed restrictions include the homogeneity restriction which guarantees that the individual demand equations sum to total cost and the symmetry restriction which guarantees that the estimated function obeys Young's theorem. In the Translog specification homogeneity implies that

$$\sum_{i=1}^M \alpha_i = 1, \text{ and } \sum_{i=1}^M A_{ij} = 0 \forall j = 1, \dots, M \quad (5)$$

where M is the number of inputs. The symmetry condition implies that

$$A_{ij} = A_{ji} \forall i, j = 1, \dots, M \text{ and } B_{ij} = B_{ji} \forall i, j = 1, \dots, N \quad (6)$$

where N is the number of outputs. Both homogeneity and symmetry can be imposed through linear restrictions. Curvature restrictions are typically nonlinear and, hence, less frequently imposed (Diewert and Wales 1987). Taken together, these restrictions give the second-order flexible functional form theoretical content. However, cynical practitioners could ask: If economic theory suggests these restrictions for empirical cost functions, why are researchers frequently forced to impose them using linear restrictions (in the case of homogeneity) and nonlinear restrictions (in the case of curvature restrictions).

In addition to questions arising from the theoretical restrictions on these cost functions, the possibility of non-stationarity in input prices and output levels raises additional difficulties for the Translog cost function. Non-stationary time series are best described as time series that

follow a random walk, or possess a unit root. Empirically, a non-stationary time series can be defined by the equation

$$x_t = \rho x_{t-1} + \varepsilon_t \quad (7)$$

where $\rho \rightarrow 1$. If the time-difference of a variable is stationary, then the series is integrated of order 1, $I(1)$. Among the important properties of non-stationary time-series, from an econometric perspective, is the fact that non-stationary time-series have an infinite variance. This infinite variance has been used to derive econometric methods for estimating econometric models of non-stationary time series. Specifically, two time-series are cointegrated if the residual of a linear relationship between the two non-stationary time series has a finite variance (or is stationary). Within the Translog cost function this implies that for the estimated cost function to be econometrically valid, all variables (costs, input prices, and output levels) must be integrated of the same order. Several studies (e.g., Lambert and Shonkwiler 1995 and Lim and Shumway 1997) have explicitly taken non-stationarity into account in the estimation of the dual.

Dynamic Specification of the Dual Translog Cost Function

Direct estimation of the Translog cost system as specified in equations 2 and 4 implicitly assumes that all costs are variable, or stated differently, the time-series observations represented a long-run equilibrium. Some studies (e.g., Chambers and Vasavada 1983) have modeled the deviation from the long-run equilibrium by appending quasi-fixed variables, or variables representing inputs that cannot be instantaneously varied within a single time period. However, this study estimates the long-run Translog cost function using the error-correction approach developed by Urga and Walters (2003) based on the general error correction form introduced by Anderson and Blundell (1982).

Urga and Walters (2003) begin by specifying an autoregressive distributed lag process for the share equations:

$$S_t = D_1 S^*(y_t, w_t) + D_2 S^*(y_{t-1}, w_{t-1}) + D_3 S_{t-1} + \eta_t \quad (8)$$

where S_t is the observed shares for each input, $S^*(y_t, w_t)$ is the optimal (long-run) input shares defined by equation 4, and D_1 , D_2 , and D_3 are $M \times M$. Based on Anderson and Blundell (1982), equation 8 can be used to derive a generalized error correction mechanism for the share equations

$$\Delta S_t = G \Delta S^*(y_t, w_t) + K(S^*(y_{t-1}, w_{t-1}) - S_t) + \eta_t \quad (9)$$

where Δ is defined as the difference operator ($\Delta x_t = x_t - x_{t-1}$), $G = D_1$, and $K = D_1 + D_2 + D_3$. Within the error-correction context, as $G \rightarrow I$ (where I is the identity matrix) and $K \rightarrow 0$ the error correction model yields the standard long-run Translog cost function specification. Adding further structure to this framework, Urga and Walters let

$$G = mI \quad (10)$$

where m becomes a constant rate of adjustment for all share equations. The error correction model for the share equations then becomes

$$\Delta S_t = m \Delta S^*(y_t, w_t) + K(S^*(y_{t-1}, w_{t-1}) - S_t) + \eta_t. \quad (11)$$

Based on the error-correction model for the share equations, the cost function containing both equilibrium and disequilibrium terms becomes

$$\begin{aligned} \ln(C_t) = & m \ln(C^*(y_t, w_t)) + (1-m) \ln(C^*(y_{t-1}, w_{t-1})) + \\ & (1-m) \left(\sum_{i=1}^M S_{i,t-1} \ln(w_{i,t}) - \sum_{i=1}^M S_i^*(y_{it}, w_{i,t}) \ln(w_{i,t-1}) \right) + \\ & \sum_{i=1}^M \sum_{j=1}^M b_{ij} (S_j^*(y_{t-1}, w_{t-1}) - S_{j,t-1}) \ln(w_{it}) \end{aligned} \quad (12)$$

where $\ln(C_t)$ is the observed cost and $\ln(C^*(y_t, w_t))$ is the minimum cost from equation 2. The error correction model for the share equations [depicted in equation 11] and the cost function specification in equation 12 can be estimated simultaneously by defining $K = mI + B$ in equation 11. Further, since the share equations sum to one by definition, one of the rows in the K matrix is redundant.

To test the implications of the dynamic Translog Specification presented in equations 11 and 12 we estimate the parameters of the Translog cost function for U.S. agriculture as specified in equations 2 and 4 using a long-run approach and the error correction specification. For discussion purposes, we will focus primarily on the violation of curvature restrictions of the dual cost specification.

Data

The data used in our analysis are from the United States Department of Agriculture, Economic Research Service (USDA/ERS). They include U.S.-level estimates of purchased inputs, labor, and capital inputs for 1948 through 1999 as well as two indices representing the quantity of crops and livestock output produced by the sector.

Our measures of output include two quantity indices, one for crop outputs and the second for livestock outputs. In each case, disaggregated crop and livestock outputs are aggregated into a single product using producer prices. The prices corresponding to each disaggregated output reflect the value of that output to the producer (i.e., subsidies are added to the market price received for each output while indirect taxes are deducted from the market value). These data were compiled by the USDA/ERS and exclude the production (both crop and livestock) that is used as farm inputs.

Intermediate inputs include feed, seed, and livestock purchases. Intermediate goods produced within the farm sector are included in intermediate input only if they also have been included in output. Another component is agricultural chemicals. To account properly for changes in input characteristics or quality, we construct price indexes of fertilizers and pesticides using the hedonic regression technique. The basic premise underlying this approach is that price differences across goods are due mainly to quality differences that can be measured in terms of common attributes. The final components of Intermediate Inputs are energy (composed of petroleum fuels, natural gas, and electricity) and other purchased inputs.

The indexes of labor input incorporate data from both establishment and household surveys. Estimates of employment, hours worked, and labor compensation are controlled to industry totals based on establishment surveys that underlie the U.S. national income and product accounts. These totals are allocated among categories of the work force cross-classified by the characteristics of individual workers on the basis of household surveys. The resulting estimates of hours worked and average compensation per hour are used to construct the indexes of labor input.

Estimates of the capital stock were constructed for each asset type. For depreciable assets, we employ the perpetual inventory method to estimate capital stocks from data on investment. Estimates of the stocks of land and inventories are implicit quantities based on balance sheet data. We constructed estimates of rental prices for each type of asset. Implicit rental prices are derived based on the correspondence between the purchase price of an asset and the discounted value of future service flows derived from that asset.

Depreciable capital assets include nonresidential structures, motor vehicles, farm tractors and other equipment. Data on investment are obtained from the U.S. Department of

Commerce's Bureau of Economic Analysis's (BEA) *Fixed Reproducible Tangible Wealth in the United States*.

Results

Both the long run (static) and dynamic Translog specifications were estimated using a concentrated maximum likelihood approach assuming a multivariate normal distribution function.² The estimated parameters for each specification are presented in Table 1. The impact of the dynamic specification on the Translog is apparent in the estimated value of m . The m coefficient determines the overall autoregressive structure of the model (i.e., parameters the constant autoregressive effect in each estimated equation). Its estimated value of 0.36 is statistically different than zero at any conventional level of confidence; indicating that the dynamic structure of the factor demands is important. Further, assuming normality the

² Using the dynamic Translog as an example, we specify the empirical model as

$$\left\{ \begin{array}{l} \eta_{1t} = \ln(C_t) - \left[m \ln(C^*(y_t, w_t)) + (1-m) \ln(C^*(y_{t-1}, w_{t-1})) + \right. \\ \quad (1-m) \left(\sum_{i=1}^M S_{i,t-1} \ln(w_{i,t}) - \sum_{i=1}^M S_i^*(y_{it}, w_{i,t}) \ln(w_{i,t-1}) \right) + \\ \quad \left. \sum_{i=1}^M \sum_{j=1}^M b_{ij} (S_j^*(y_{t-1}, w_{t-1}) - S_{j,t-1}) \ln(w_{it}) \right] \\ \eta_{2t} = \Delta S_t - [m \Delta S_t^* + K(S_{t-1}^* - S_{t-1})] \end{array} \right\} = \eta_t$$

where η_{1t} is the scalar residual of the dynamic cost function and η_{2t} is the $M-1$ vector of residuals from the dynamic share equations. The parameters (α_0 , α , A , β , B , Γ , m , and B) are then estimated by maximizing the concentrated likelihood function

$$L = -\frac{T}{2} \ln \left[\det \left(\frac{\eta' \eta}{T} \right) \right]$$

where T is the number of time-series observations ($t = 1, \dots, T$)

hypothesis that $m = 1$ can also be rejected at any conventional confidence level.³ The rejection of the hypothesis that $m = 1$ implies that non-stationarity is not present in the specification.

Given that the autoregressive structure is economically significant, we examine the effect of the dynamic specification on the economic implications of the estimates. In our estimations, homogeneity and symmetry are imposed on both cost systems. In the case of the long-run Translog cost function, we impose homogeneity by normalizing each price by one of the prices (in this case the price of capital inputs). In the dynamic Translog specification, differencing the share equations imposes homogeneity. In both specifications, symmetry is imposed by linear restrictions on the estimated parameters. In this application, however, the most interesting impact of the dynamic Translog specification involves the curvature restrictions.

As discussed in Chambers (1984), a dual cost function is concave in input prices and convex in output levels. Thus, the curvature conditions for the Translog imply that

$$\nabla_{ww}^2 C^*(y_t, w_t) = \left[A + S^*(y_t, w_t) S^*(y_t, w_t)' \right] = A^* \quad (13)$$

If the Translog cost function is concave in input prices A^* is negative definite. In order for A^* to be negative definite, A must be negative definite (i.e., $S^*(y_t, w_t) S^*(y_t, w_t)'$ is positive semidefinite by construction). The empirical estimates of Translog cost function presented in Table 1 imply different concavities. The parameter estimates in Table 1 imply two estimates of the A matrix:

³ Given the formulation of the model, the distribution of m may follow the distribution of the Dickey-Fuller statistic. However, given that our estimated value of m is less than 0.75 the assumption of normality may be appropriate (i.e., the autoregressive parameter does not approach a unit-root).

$$\begin{aligned}\tilde{A} &= \begin{bmatrix} 0.2120 & -0.0883 \\ -0.0883 & 0.1054 \end{bmatrix} \text{ for the long-run Translog, and} \\ \hat{A} &= \begin{bmatrix} -0.1047 & 0.2052 \\ 0.2052 & -0.4060 \end{bmatrix} \text{ for the dynamic Translog}\end{aligned}\tag{14}$$

Based on these estimated values, it is obvious that the estimates of the long-run Translog cannot obey the concavity restrictions for dual cost functions (i.e., the diagonal elements of the matrix are positive). For symmetric real valued matrices, the definiteness of the matrix can be inferred from the matrix's eigenvalues. If all the eigenvalues of the matrix are positive, the matrix is positive definite. On the other hand, if all the eigenvalues of the matrix are negative, the matrix is negative definite. However, if the eigenvalues of the matrix have mixed signs, the matrix is indefinite. Table 1 also presents the eigenvalues for each A matrix. These results indicate that the estimated A matrix for the long-run Translog is actually positive definite which contradicts economic theory. However, the dynamic eigenvalues of the dynamic Translog are both negative implying that the estimated A matrix is negative definite, consistent with production theory.

In addition to the implications of the dynamic specification for the overall concavity, the significance of the dynamic specification for each individual share equation is demonstrated by the statistical significance of the B parameters. Of these parameters, only b_{12} fails to be statistically significant at the 0.01 level of confidence, and it is statistically significant at the 0.05 level of confidence.

Conclusion

This study has used an empirical approach developed by Urga and Walters (2003) to examine the implications of the short-run specification of the standard Translog cost specification along with the possible implications of non-stationarity. We have estimated a

dynamic Translog cost specification complete with dynamic share equations for U.S. agriculture and compared it to the static, long-run specification. Homogeneity and symmetry restrictions are imposed through linear restrictions. Also, the possibility of non-stationarity in input prices and output levels is examined.

We estimated both the long-run (static) and the dynamic Translog specifications using maximum likelihood estimation finding that the dynamic Translog specification yielded more significant parameter estimates, and yielded results that are consistent with economic theory. In particular, the coefficient m (the adjustment cost parameter) determines the overall autoregressive structure of the model. The fact that its estimated value (0.36) is statistically different from zero at any conventional level of confidence indicates that the dynamic structure of the model is important. This finding illustrates the superiority of the short-run, dynamic specification over the static, long-run model.

Since we found that the autoregressive structure is economically significant, we examined the economic implications of using the dynamic specification. A key finding of our study is that whereas the static, long-run model's estimates did not obey the concavity restrictions for dual cost functions, the dynamic Translog model's estimates are consistent with concavity restrictions, and thus with production theory.

References

- Anderson, G. and R. Blundell. "Estimation and Hypothesis Testing in Dynamic Singular Equation Systems." *Econometrica* 50(1982): 1559–71.
- Capalbo, S.M. and M.G.S. Denny. "Testing Long-Run Productivity Models for the Canadian and U.S. Agricultural Sectors," *American Journal of Agricultural Economics* 68(1986): 615–25.
- Chambers, R.G. *Applied Production Analysis: A Dual Approach* New York: Cambridge University Press, 1984.
- Chambers, R.G. and U. Vasavada. "Testing Asset Fixity for U.S. Agriculture." *American Journal of Agricultural Economics* 76(1983): 761–769.
- Clark, J.S. and C.E. Youngblood. "Estimating Duality Models with Biased Technical Change: A Time Series Approach," *American Journal of Agricultural Economics* 74(August 1992): 583–91.
- Diewert, W.E., and T.J. Wales. "Flexible Functional Forms and Global Curvature Conditions." *Econometrica*, 55 (1987): 43–68.
- Engle, R.F. and C.W.J. Granger. "Cointegration and Error-Correction: Representation, Estimation, and Testing." *Econometrica* 55 (1987): 251-76.
- Granger, C.W.J. and P. Newbold. "Spurious Regressions in Econometrics," *Journal of Econometrics* 26(1974): 1045–66.
- Huffman, W.E., V.E. Ball, M. Gopinath, and A. Somwaru. "Public R&D and Infrastructure Policies: Effects on Cost of Midwestern Agriculture," In *Agricultural Productivity: Measurement and Sources of Growth*, V. E. Ball and G. W. Norton (eds) (Boston: Kluwer Academic Publisher, 2002): 167–184.

- Lambert, D.K. and J.S. Shonkwiler. "Factor Bias Under Stochastic Technical Change," *American Journal of Agricultural Economics* 77(1995): 578–90.
- Lim, H. and C.R. Shumway. "Technical Change and Model Specification: U.S. Agricultural Production," *American Journal of Agricultural Economics* 79(1997): 543–54.
- Urga, G. and C. Walters. "Dynamic Translog and Linear Logit Models: A Factor Demand Analysis of Interfuel Substitution in U.S. Industrial Energy Demand." *Energy Economics* 25(2003): 1–21

Table 1. Estimates of Long-Run and Dynamic Translog Parameters

	Long-Run Translog	Dynamic Translog
α_0	12.1985 (0.0715)	-3.0308 (39.4867)
α_1	0.3447 (0.0103)	-0.0328 (0.1381)
α_2	0.2794 (0.0123)	1.9533 (1.3523)
A_{11}	0.2120 (0.0198)	-0.1047 (0.0020)
A_{12}	-0.0883 (0.0107)	0.2052 (0.0016)
A_{22}	0.1054 (0.0144)	-0.4060 (0.0138)
β_1	4.0510 (0.8616)	-7.5641 (6.0028)
β_2	1.6134 (0.5028)	-0.3667 (8.5640)
B_{11}	16.7172 (4.9167)	-1.7301 (5.4119)
B_{12}	-5.4118 (3.4072)	-1.6759 (2.2364)
B_{22}	3.2963 (2.5615)	0.3041 (1.2302)
Γ_{11}	0.3736 (0.0562)	-0.2278 (0.0216)
Γ_{12}	0.0249 (0.0224)	-0.0168 (0.0020)
Γ_{21}	-0.3147 (0.0665)	0.9872 (0.0496)
Γ_{22}	-0.1151 (0.0384)	0.0371 (0.0640)
m		0.3633 (0.0018)

^aNumbers in parenthesis denote standard errors.

Table 1. Estimates of Long-Run and Dynamic Translog Parameters

	Long-Run Translog	Dynamic Translog
b_{11}		-0.3924 (0.0033)
b_{12}		-0.0047 (0.0025)
b_{21}		0.0167 (0.0002)
b_{22}		-0.2899 (0.0020)
<i>Eigenvalues for Input Price Coefficients</i>		
λ_1	0.0556	-0.5099
λ_2	0.2618	-0.0007