Thailand - Agriculture

1979

MITIGATING THE EFFECTS OF MULTICOLLINEARITY USING EXACT AND STOCHASTIC RESTRICTIONS: THE CASE OF AN AGGREGATE AGRICULTURAL PRODUCTION FUNCTION FOR THAILAND

by

Ron C. Mittelhammer, Douglas L. Young,
Damrongsak Tasanasanta and John T. Donnelly

Contributed paper submitted to Annual Meeting of the
American Agricultural Economics Association
Pullman, Washington, July 29-August 1, 1979

ABSTRACT

Mixed estimation and OLS were each used to estimate an aggregate agricultural production function for Thailand for which data were highly multicollinear. The final mixed model obtained using minimax regret in the pretest appeared to significantly outperform OLS in terms of precision and overall reasonableness, mitigating a serious multicollinearity problem.

MITIGATING THE EFFECTS OF MULTICOLLINEARITY USING EXACT AND STOCHASTIC RESTRICTIONS: THE CASE OF AN AGGREGATE AGRICULTURAL PRODUCTION FUNCTION FOR THAILAND

In recent years econometricians have made substantial progress in developing a number of alternatives to ordinary least-squares (OLS) regression techniques. These alternatives which include ridge regression [Hoerl and Kennard], principal components regression [Hill, Fomby, and Johnson], exact linear restrictions [Goldberger], Bayesian econometrics [Zellner], and mixed estimation [Theil and Goldberger], have the potential for mitigating the effects of multicollinearity. While these estimating techniques may produce biased estimates, it has been demonstrated that they can achieve substantial reductions in mean square error (MSE) under appropriate circumstances. The theoretical progress in development of biased alternatives to OLS regression procedures has been accompanied by development of the statistical properties of pre-test estimators in regression analysis [Bock, Yancey, and Judge; Brook].

Notwithstanding the immediate relevance of these theoretical developments to applied economists, there has been relatively little consideration of either biased regression techniques or pre-test implications in applied econometric research.

The purpose of this paper is twofold: (1) to evaluate in an applied study the use of regression with exact and stochastic linear restrictions to mitigate the effects of serious multicollinearity and, (2) to discuss the implications of pre-test considerations. The techniques are applied to the estimation of an aggregate agricultural production function for Thailand using time series data from 1950 to 1976.[1]

Aggregate Production Function Estimation and Agricultural Growth Measurement

Serious multicollinearity frequently reduces the precision of parameter estimates for aggregate production functions to the point where confidence in their signs and magnitudes is very low. Compounding the multicollinearity problem have been efforts to measure technological progress in the aggregate production process by using some function of time as a proxy. Aggregate measures of land, labor, and capital inputs tend to be highly correlated with each other and with time.

The specific substantive objectives of the present application, whose successful achievement was made possible by the use of the stochastic linear restrictions (mixed estimation) regression technique, were (1) to obtain precise estimates of production elasticities for land, labor, capital, and technology in Thai agriculture in order to decompose 1950-1976 agricultural growth between neutral technological progress and increased use of conventional inputs, and (2) to test the hypothesis that three government sponsored economic planning periods between 1961 and 1976 had no discernible impact on the rate of any neutral technological progress in the agricultural sector.

As the largest sector of the Thai economy, agriculture currently provides one-third of the GDP, 70 percent of export earnings, and employment for 78 percent of the total labor force [Ministry of Agriculture and Cooperatives; Sriplung]. The critical role of agriculture in generating foreign exchange was accentuated by the marked reduction of capital inflows associated with the withdrawal of American troops during the early 1970's and by continuing deficits in the balance of trade [Ministry of Agriculture and Cooperatives].

In spite of Thailand's generally accepted promising agricultural potential, there is a widespread feeling that actual performance has fallen short of that potential. In large part, this feeling is based on the observed relatively slow rate of adoption of the new high yielding varieties (HYVs) of rice, and the related minimal technological impact of the "Green Revolution," in Thailand as compared to certain other Asian countries [Griffin, pp. 41-43; Myint, Ch. 2]. This phenomenon has been attributed to both environmental and administrative constraints. First, as is well known, a necessary condition for realizing the yield potential of the HYVs of rice is water control facilities or exceptionally well distributed rainfall [B. Johnston]. In Thailand, the lack of flood control facilities has constrained the adoption of the short-stemmed HYVs in the deepflooding river basins, while inadequate irrigation development has limited adoption in the Central Plains [Griffin]. Secondly, and more generally, in Thailand a scarcity of effective agricultural research and extension resources, together with high input costs, incentive weakening commodity price policies, and infrastructural and administrative deficiencies may have seriously limited the realization of productivity gain potentials in agriculture [Silcock].

In view of these widespread predominantly qualitative speculations on agricultural productivity trends in Thailand, a quantitative examination of the sources and rates of agricultural growth, and the impact of government planning periods on that growth, appears to be justified.

OLS Model and Results

A variant of Tinbergen's specification of the Cobb-Douglas production function which permits a constant rate of neutral technological progress was selected.[2] Upon modification to permit differential technological progress among government planning periods, the specification became:

$$(1)\quad Q = Ae^{t(\alpha_0 + \alpha_1 D_1 + \alpha_2 D_2 + \alpha_3 D_3)} L^{\beta} N^{\gamma} K^{\delta} e^{u}$$

where

Q = aggregate agricultural output in billion baht at constant 1956 prices, including crops, livestock, fishery, and forestry output

A = constant

t = time, 1, 2, ..., 27

L = land input, harvested acreage in million rai

N = agricultural labor force in million workers

K = agricultural capital in billion baht at constant 1956 prices

D_1 = 1 in 1961-1966, 0 elsewhere; represents first government planning period

D_2 = 1 in 1967-1971, 0 elsewhere; represents second government planning period

D_3 = 1 in 1972-1976, 0 elsewhere; represents third government planning period

u = error term

Given a value of t, the production function is conditionally homogeneous of degree $(\beta + \gamma + \delta)$ in land, labor, and capital. The time variable $e^{t(\alpha_0 + \alpha_1 D_1 + \alpha_2 D_2 + \alpha_3 D_3)}$, is introduced as a proxy for technology. The rate of technological change is $(dQ/dt)/Q = \alpha_0 + \alpha_1 D_1 + \alpha_2 D_2 + \alpha_3 D_3)$.

Taking the natural logarithm of both sides of equation (1), the following estimating equation for the production function was obtained:

(2) $\ln Q = \ln A + \alpha_0 t + \alpha_1 D_1 t + \alpha_2 D_2 t + \alpha_3 D_3 t + \beta \ln L + \delta \ln N + \gamma \ln K + u$

Examination of the simple and multiple correlations involving regressors in (2) indicated that a very serious multicollinearity problem existed.

Pure OLS estimation of the highly multicollinear model in (2) (see model 1, Table 1) resulted in all but two of the parameter estimates being insignificant at the .05 level of type I error ($t_{.05,\ 19df} = 2.093$). The extremely high estimated elasticity on the labor variable and the negative growth rate attributed to technology seem particularly unreasonable.

A simultaneous test of hypothesis that all three coefficients associated with the government planning period dummy variables were zero was accomplished using an F test.[3] The obvious information provided by the test was a statistical decision regarding the significance of the government planning periods in affecting Thai agricultural output. A second motivation for the test was to ascertain whether the three exact linear restrictions on the parameters of the model in the form of deletion

Table 1. Production Function Results Using Alternative Estimation Techniques

Model No.a/	Constant	Estimated Coefficient of:									
	ln A	ln L	ln N	ln K	t	D_1t	D_2t	D_3t	D-W	R^2	F
1	-4.5108	0.4937	3.2300	0.1668	-0.0553	-0.0036	-0.00222	0.0018	1.683	.988	
		(2.74)* b/	(2.86)*	(0.38)	(-1.86)	(-0.88)	(-0.33)	(0.25)			
2	-1.7410	0.4907	1.3244	0.1776	-0.0115					.985	1.78
		(2.98)*	(1.84)	(0.87)	(-0.50)						
3	-.0273	.3384	.5949	.1578	.0099	.0009	-.0002	.0014		.985	
		(4.025)*	(4.002)*	(2.141)*	(1.614)	(.302)	(-.05)	(.368)			
4	-.0620	.3535	.5775	.1562	.0110					.984	1.91
		(4.401)*	(3.954)*	(2.296)*	(1.971)*						

a/ Model 1 is pure ordinary least squares (OLS), model 2 is OLS with exact linear restrictions, model 3 is mixed estimation, model 4 is mixed estimation with exact linear restrictions.

b/ Values in paretheses are T-values.

* Denotes significance at the level of 0.05 or higher.

of the dummy variables should be imposed to obtain estimators of the parameters in the model that were superior to unrestricted least square estimators. The assessment of superiority utilized the quadratic risk function,

(3) $E(X\hat{\beta} - X\beta)' (X\hat{\beta} - X\beta) = E(\hat{\beta} - \beta)'X'X(\hat{\beta} - \beta)$.

where X is the N x K matrix of explanatory variable values and $\hat{\beta}$ is an estimate of β (for additional details see Wallace or Brook). While it would have been preferable to base decisions on the loss function associated with $\hat{\beta}$, $E(\hat{\beta}-\beta)' (\hat{\beta}-\beta)$, rather than on the loss associated with the linear combination $X\hat{\beta}$, the currently existing literature on pretesting (see Wallace, and Brook) does not suffice for ready application of the former criterion.

The minimax regret approach of Brook was used in choosing a critical value for the F test. The minimax regret criterion results in a final (pretest) estimator of the model whose quadratic risk function exhibits the smallest maximum deviation from the minimum of the risks of the OLS and restricted least squares estimators throughout the range of possible biases in the restrictions (see Brook). It is noteworthy that this minimax regret critical value for 3 restrictions and 19 degrees of freedom of

1.93 is substantially below a more conventional critical value of 3.13 which would represent a test of the hypothesis at a .05 level of type I error. It is, therefore, recognized that the choice of the restricted estimator based on minimax regret implies acceptance of the hypothesis that the three government planning dummy variables have coefficients not significantly different from zero at the .05 level of significance.

The calculated F value of 1.78 warrants both the conclusion that the restricted estimator (model 2 in Table 1) should be chosen based on minimax regret, and the acceptance of the hypothesis that the government planning period dummy variable coefficients are equal to zero. Examination of the reduced OLS model reveals that although standard errors of the parameter estimates were generally reduced, the parameter estimates associated with the capital and time variables remained relatively imprecise. The estimated output elasticity for labor of 1.3244, although reduced, still seems unreasonably high on theoretical grounds. In response to the unsatisfactory nature of the preceding results, and since it was felt that prior stochastic bounds on some of the output elasticities could be constructed, the technique of mixed estimation was examined.

Mixed Estimation

Often researchers have prior notions concerning the signs and general ranges of magnitudes of at least some of the parameters in their models. This prior information emanates from previous research, economic theory, the institutions surrounding the system, and/or simple introspection by the researcher. The mixed estimation technique [Theil and Goldberger] is a method of combining a pure model (based on sample information only) with prior information in the form of stochastic linear constraints on the parameters of the model. The most important property of the mixed estimator from the standpoint of the multicollinearity problem is that the potential exists for obtaining estimates that are superior to (pure) OLS based on mean square error considerations.

Reflecting on similar research in other Asian countries [Ho; Akino and Hayami; Tang, 1963; Khan; and Tang, n.d.] and on direct observation of the agricultural sector in Thailand, the following interval estimates for output elasticities with respect to land, labor, and capital were decided upon by the authors:

(4) $E_{Q,LAND} \varepsilon [.10, .50]$ with probability .95

$E_{Q,LABOR} \varepsilon [.25, .85]$ with probability .95

$E_{Q,CAPITAL} \varepsilon [0., .30]$ with probability .95

The prior point estimates of the elasticities were designated as the midpoints of the intervals in (4), and estimates were assumed to be generated from normal priors with variances .01, .0255, and .005625, respectively, and with zero covariances. (See Theil and Goldberger for the linear representation of prior constraints.)

The production function (1) was estimated subject to the stochastic constraints on the magnitudes of the output elasticities with respect to land, labor, and capital using mixed estimation (model 3 of Table 1). Judging by the t-values associated with the parameter estimates, the coefficients on land, labor, capital, and time appear to have been estimated relatively precisely, while the dummy variable coefficient estimates appear to be much less precise.

In this case, there are six constraints that must be examined in choosing the final model for the Thai aggregate production function. The first set of three restrictions, as before, refer to the deletion of the government planning period dummy variables from the model and are exact linear restrictions. The second set of three constraints refer to the stochastic restrictions on the output elasticities with respect to land, labor, and capital (4). The final estimator involved a choice between the OLS estimator, and the restricted estimator incorporating both the exact and stochastic linear restrictions (a reduced mixed model).

An F statistic corresponding to the six restrictions was calculated [Judge, Yancey, Bock, p. 36] and its value of 1.91 was compared to the minimax regret critical value of 1.96 for six and nineteen degrees of freedom. Based on this test, the restricted estimator (mixed estimator with the government planning period dummy variables deleted) was chosen as the final model. It should be noted that acceptance of the hypothesis that the coefficients on the government planning period dummy variables are not significantly different from zero as well as acceptability of the hypothesis of compatibility between sample and prior information [Theil] is implied in the pre-test above.

The mixed estimates (model 4) of the output elasticities with respect to land, labor, and capital are each lower in value than are the corresponding final least square estimates. The most notable differences among estimates are the estimate of the labor elasticity, which was reduced in magnitude by over half, and the effect of the time-technology proxy which changed from having a negative to having a positive effect. The estimate of the homogeneity of the production function conditional on time was reduced from 1.9927 estimated via exactly restricted least squares to 1.0872 estimated from the reduced mixed model. The coefficients estimated by the latter procedure are considered more reasonable on theoretical grounds than those estimated by the pure OLS or the exact linear restrictions models. Examination of the t-values associated with these mixed coefficient estimates suggest that the precision in estimation has been substantially improved over the least squares approach, mitigating the effects of a rather severe multicollinearity problem.

Contribution of Increased Input Use and Technological Progress to Agricultural Output Growth

To ascertain the respective contributions of conventional inputs and neutral technological progress on Thai agricultural output growth, the increase in agricultural output between 1950 and 1976 was decomposed into: (1) a movement along the 1950 production surface due to increased use of conventional inputs, and (2) an upward shift in the production surface due to neutral technological progress (see Table 2).

The exactly restricted least squares estimates would indicate that there was technological decay rather than growth. Given the imprecision with which the

Table 2. Shares of Conventional Input Expansion and Technological Progress in Increased Agricultural Output in Thailand, and Growth Rate due to Technological Advance, 1950-1976

Estimating Technique	Input Share	Technology Share	Technology Induced[1] Output Growth Rate
Restricted Least Squares	1.349	-.349	-.0115 (-.50)
Mixed Estimation	.628	.372	.0110 (1.971)

[1] The growth rate is given as $(dQ/dt)/Q = \alpha o$, that is, the coefficient of t in the models. The value in parenthesis is the T-value associated with the estimated growth rate.

coefficient on the technology variable was estimated, together with the suspect size of the output elasticity with respect to labor, it seems advisable to concentrate on the results provided in the final mixed model.

These results indicate that the increase in conventional input levels from 1950 to 1976 would account for roughly 63% of the agricultural output growth, while technological growth would account for the remainder.

Summary and Conclusions

Mixed estimation markedly improved the statistical precision and theoretical reasonableness of estimated parameters on land, labor, capital, and time (a technology proxy) for an aggregate agricultural production function for Thailand. Using the final mixed model, the rate of Thai output growth attributed to neutral technological progress was about 1.1 percent per year between 1950 and 1976, whereas using OLS a negative rate was estimated. The mixed model indicated the degree of homogeneity of the utilized Cobb-Douglas production function to be about 1.1, compared with an OLS estimate of nearly 2.0. The results of mixed estimation attributed approximately one third of total Thai output growth during 1950-76 to neutral technological progress and the remaining two-thirds to expand use of land, labor, and capital. Neither OLS nor mixed estimation detected a statistically significant differential impact of economic planning periods on the rate of neutral technological progress in Thai agriculture. It should be recalled that in this application the authors' prior information concerning land, labor, and capital production elasticities was shown to be statistically compatible with the sample information.

These quantitative results, the authors believe, allow the qualitative conclusion that advances in technology materially contributed to Thai agricultural growth during the past three decades.[4] Possible explanations for the apparent realization of neutral technological progress in Thai agriculture, even in the absence of any Green Revolution there, as well as the apparent absence of any variable impact of the Thai government's different economic planning periods, have been explored elsewhere [Tasanasanta].

This application demonstrated the potential usefulness of mixed estimation in increasing precision of estimated parameters in the presence of extreme multicollinearity and formally addressed implied pretest considerations. The precise estimation of this particular aggregate agricultural production function, and the subsequent analysis of agricultural output growth in Thailand, could not have been accomplished had the mixed estimation alternative to OLS not been available.

Footnotes

Ron C. Mittelhammer and Douglas L. Young are Assistant Agricultural Economists and Assistant Professors in the Department of Agricultural Economics at Washington State University. Damrongsak Tasanasanta was a graduate student and John T. Donnelly is an Associate Professor in the Department of Economics at Washington State University.

[1] A complete description of the data and data sources can be found in Tasanasanta. Tasanasanta's study analyzed Thai output growth using a similar approach but a slightly different production function specification.

[2] The many limitations of the simple Cobb-Douglas form and the assumption of neutral technological progress have been well documented in the literature, e.g., see Murray Brown. However, given the limited adoption of "Green Revolution" technolgies in Thai agriculture, the assumption of neutral technological progress seemed reasonable in this initial investigation. This manageable specification also facilitated comparisons of the estimation techniques whose evaluation was a principle objective of this research.

[3] The test is calculated as:

$$\frac{(C\,\hat{\beta})'\,[C(X'X)^{-1}C']^{-1}\,(C\,\hat{\beta})}{gS^2} \sim F(n-k)$$

where g = number of independent constraints to be tested (in this case g equals the number of variables deleted, i.e., the number whose parameters have been constrained to zero), n = number of observations, k = number of parameters estimated, S^2 is the square of the standard error of the equation, and C is appropriately defined.

[4] This conclusion was further strengthened by remarkably similar quantitative results from the use of principle components regression on this data; these results were deleted from this paper due to space limitations.

References

Akino, Masakutsu and Yujiro Hayami, "Sources of Agricultural Growth in Japan, 1880-1965," Quarterly Journal of Economics, 88 (August 1974): 454-470.

Bock, M.E., T.A. Yancey, and G. Judge, "The Statistical Consequences of the Preliminary Test Estimators in Regression," J. Amer. Statist. Ass., 68 (1973):109-116.

Brook, Richard J., "On the Use of a Regret Function to Set Significance Points in Prior Tests of Estimation," J. Amer. Statist. Ass., 71(1976):126-131.

Brown, Murray, On the Theory and Measurement of Technological Change, Cambridge University Press, London, 1966.

Goldberger, A.S., Econometric Theory, John Wiley and Sons, New York, 1964.

Griffen, Keith, The Political Economy of Agrarian Change: An Essay on the Green Revolution, Harvard University Press, Cambridge, Mass., 1974.

Hill, R.C., Thomas B. Fomby and S.R. Johnson, "Component Selection Norms for Principal Components Regression," Commun. Statist., A6(1977):309-334.

Ho, Yhi-min, Agricultural Development of Taiwan, 1903-1960, Vanderbilt University Press, 1966, p. 63.

Hoerl, A.E. and R.W. Kennard, "Ridge Regression: Biased Estimation for Nonorthogonal Problems," Technometrics, 12(Feb. 1970):55-67.

Johnston, Bruce F., "Agriculture and Structural Transformation in Developing Countries: A Survey of Research," Journal of Economic Literature, 8 (1970):369-404.

Judge, G.G., T.A. Yancey, and M.E. Bock, "Properties of Estimators After Preliminary Tests of Significance When Stochastic Restrictions Are Used In Regression," J. Econometrics, 1(1973):29-48.

Khan, N.A., Problems of Growth of an Underdeveloped Economy," Asia Publishing Bombay, 1961.

Ministry of Agriculture and Cooperatives, Agricultural Statistics of Thailand, Bangkok, 1975-76.

Myint, H., Southeast Asia's Economy: Development Policies in the 1970's, Praeger Publishers, New York, 1972.

Silcock, T.H., The Economic Development of Thai Agriculture, Cornell University Press, 1970.

Sriplung, S., Poverty of Farmers and Solutions, Rusm Mitra Thai Publisher, Bangkok, 1976, (Thai version).

Tang, A.M., "Research and Education in Japanese Agricultural Development, 1880-1938," The Economic Studies Quarterly, (May 1963): 93.

------, Input-Output Relations in Agricultural Output of Communist China, 1952-65, Nashville, Vanderbilt University, n.d., (mimeographed).

Tasanasanta, Damrongsak, Sources of Agricultural Growth in Thailand, 1950-1976, unpublished Ph.D. dissertation, Washington State University, 1978.

Theil, H., "On the Use of Incomplete Prior Information in Regression Analysis," J. Amer. Statist. Ass., 58(1963):401-414.

------ and A.S. Goldberger, "On Pure and Mixed Statistical Estimation in Economics," International Economic Review, 2(1961), p. 65-78.

Yancey, T.A., G.G. Judge, and M.E. Bock, "A Mean Square Error Test When Stochastic Restrictions Are Used in Regression," Comm. In Statist., 3(1974):755-768.

Zellner, A., An Introduction to Bayesian Inference in Econometrics, John Wiley and Sons, New York, 1971.