THE PROBLEM OF AGGREGATION IN PRODUCTION FUNCTIONS

By G. Sadasivan & S. C. Rai
I.A.R.S., New Delhi

A production function gives the relationship between output and inputs. There is a wide choice of algebraic forms which can be used to represent production functions. The Cobb-Douglas production function is probably the most popular one and it had a long and successful life without serious rivals. In its best known form with Y measuring outputs, K the quantities of capital and L the input of labour, we write

$$Y = AL^{\alpha} K^{\beta} \text{ where } \begin{array}{ll} Y > 0 & \alpha \geqslant 0 \\ K > 0 & \beta \geqslant 0 \\ L > 0 & A > 0 \end{array}$$

A is a constant and α and β are the elasticities of production with respect to labour and capital respectively.

Our problem arises in combining the production functions of different firms or different localities in order to get an average production function for an entire firm or entire locality. The systematic treatment of this problem was developed by Klein. He suggested that for obtaining the average production function and aggregate marginal productivity relations analogous to microfunctions we should take the weighted geometric means of the corresponding microvariables where the weights are proportional to the elasticities of each firm or farm. The elasticities of the macrofunction are the weighted average of the elasticities of microfunction with weights proportional to expenditure on the factor. The macrorevenue is the macroprice multiplied by the macroquantity. These are the Kleinian way of aggregation. Kleinian methods of aggregation leads to some troubles about wage rate, prices of output and capital. Hence other methods of aggregation have to be investigated.

The problems of linear aggregation was considered by Theil. The Cobb-Douglas function in its logarithmic form for the i-th firm is

$$(a)\ x_i=\alpha_i\, l_i+\beta_i\, k_i+a_i \qquad (i=1, 2,\dots n)$$

where $x_i=\log \times_i$, $l_i=\log L_i$ and $k_i=\log k_i$,.

From these a macroproduction function of the form

$$x=\alpha l+\beta k+a+t$$

for observation of the aggregate values for a time series can be set up. We first examine the regression of the labour input of the *i*th firm or aggregate labour input and aggregate capital input for the time series observation,

$$l_i=B_{li}\, l+C_{li}\, k+D_{ki}+U_{ki}$$

The coefficients *B* and *C* describe the systematic movements of microvariables as macro-quanities change. Substituting these values in equation (*a*) we get the macroproduction function for time series data. A survey of the aggregation problem in production function shows that we cannot approximate the basic requirements of sensible aggregation except for farms under same crops, firms under the same industry and other narrow sections of the economy.

Another way of obtaining the average production function is to take the geometric mean of all the microelasticities and put them in the aggregate function. In a sensible aggregation, the production function must be additivity separable (Nataf). In other words the output should be equal to a labour component plus a capital component. The Cobb-Douglas function does not satisfy the condition, but when trans formed logarithmically the function is additivity separable. In this paper we willdiscuss the different ways of obtaining the aggregate production function and present the variances of the elasticities obtained in each way.

Let there be two firms or farms and the Cobb-Douglas production function for these are

$$y_1=A_1\, x_1^{a_1} x_2^{a_2}\dots x_n^{a_n} \text{ and}$$

$$y_2=B_1\, x_1^{b_1} x_2^{b_2}\dots x_n^{b_n}$$

where $x_1, x_2,\dots x_n$ are the inputs and y_1 and y_2 are the outputs of the two firms or regions $a_1, a_2,\dots a_n$ and $b_1, b_2,\dots b_n$, are the elasticities. A_1 and B_1 are constants. Let us consider now the different methods of obtaining the average production function and estimate the variances of different elasticities. Combining the production function by taking the *AM* of the elasticities we have

$$y=A\, x_1^{\frac{a_1+b_1}{2}}\, x_2^{\frac{a_2+b_2}{2}} \dots x_n^{\frac{a_n+b_n}{2}} \qquad \dots(1)$$

If we take the geometric mean of the elasticities then the average production function becomes

$$y = A^* \, x_1^{\sqrt{a_1 b_1}} \cdot x_2^{\sqrt{a_2 b_2}} \ldots x_n^{\sqrt{a_n b_n}} \qquad \ldots(2)$$

Taking logarithms of (1) and (2) we have

$$\log y = \log A^* + \frac{a_1+b_1}{2} \log x_1 + \ldots + \frac{a_n+b_n}{2} \log x_n \qquad \ldots(3)$$

$$\log y = \log A + \sqrt{a_1 b_1} \log x_1 + \ldots + \sqrt{a_n b_n} \log x_n \qquad \ldots(4)$$

In general the elasticities of function (3) are greater than the corresponding elasticities of function (4).

To work out the variances of these elasticities we use the following lemma. *Lemma.* If T is asymptotically normally distributed about θ, the function $F(T)$ of T is asymptotically normally distributed about $F(\theta)$ with variance $\left(\frac{dF}{d\theta}\right)^2 \psi(\theta)$ where $\psi(\theta)$ is the variance of T, provided that $\frac{dF}{dT}$ is continuous in the neighbourhood of θ. Applying the above lemma the estimates of variances of the elasticities of production function (2) can be worked out under the assumption that the partial regression coefficients b_i are asymptotically normally distributed with means as Bi.

Put $$A = \sqrt{a_1 b_1}$$

so that

$$\log A = \tfrac{1}{2} \log a_1 + \tfrac{1}{2} \log b_1.$$

We will now obtain the variances of $\log a_1$ and $\log b_1$. Put

$$F = \log a_1$$

$$\frac{dF}{da_1} = \frac{1}{a_1}$$

$$\therefore \quad \text{Var}(\log a_1) = \frac{1}{a_1^2} \text{Var}\, a_1.$$

Similarly

$$\text{Var}(\log b_1) = \frac{1}{b_1^2} \text{Var}\, b_1.$$

Assuming the elasticities a_1 and b_1 to be independent we have

$$\text{Var}(\log A) = \frac{1}{4a_1^2} \text{Var}(a_1) + \frac{1}{4b_1^2} \text{Var}(b_1).$$

Now put

$$F=\log A$$

$$\frac{dF}{dA}=\frac{1}{A}$$

$$\text{Var}(\log A)=\frac{1}{A^2}\text{Var}\, A$$

or

$$\text{Var}\, A=A^2\,\text{Var}(\log A)$$

i.e.

$$\text{Var}(\sqrt{a_1b_1})=\frac{a_1b_1}{4}\left[\frac{1}{a_1^2}\text{Var}(a_1)+\frac{1}{b_1^2}\text{Var}(b_1)\right]$$

Similarly

$$\text{Var}(\sqrt{a_2b_2})=\frac{a_2b_2}{4}\left[\frac{1}{a_2^2}\text{Var}(a_2)+\frac{1}{b_2^2}\text{Var}(b_2)\right]$$

$$\ldots \qquad \ldots \qquad \ldots$$

$$\text{Var}(\sqrt{a_nb_n})=\frac{a_nb_n}{4}\left[\frac{1}{a_n^2}\text{Var}(a_n)+\frac{1}{b_n^2}\text{Var}(b_n)\right].$$

Variances of the elasticities of the production function (1) are given below.

$$\text{Var}\left(\frac{a_1+b_1}{2}\right)=\frac{1}{4}\left[\text{Var}(a_1)+\text{Var}(b_1)\right]$$

$$\text{Var}\left(\frac{a_2+b_2}{2}\right)=\frac{1}{4}\left[\text{Var}(a_2)+\text{Var}(b_2)\right]$$

$$\ldots \qquad \ldots \qquad \ldots$$

$$\text{Var}\left(\frac{a_n+b_n}{2}\right)=\frac{1}{4}\left[\text{Var}(a_n)+\text{Var}(b_n)\right].$$

Comparing the variances of different elasticities of the two production functions, we observe that

$$\text{Var}\left(\frac{a_1+b_1}{2}\right) < \text{Var}\sqrt{a_1b_1} \text{ if}$$

$$\text{Var}(a_1)+\text{Var}(b_1) < \frac{b_1}{a_1}\text{Var}(a_1)+\frac{a_1}{b_1}\text{Var}(b_1) \qquad \ldots(5)$$

But

$$\frac{\text{Var}(a_1)}{a_1} < \frac{\text{Var}(b_1)}{b_1} \text{ if } a_1 > b_1$$

and

$$\frac{\text{Var}(b_1)}{b_1} < \frac{\text{Var}(a_1)}{a_1} \text{ if } b_1 > a_1 \text{ [from (5)].}$$

In other words if the elasticity of a particular input of one firm is greater than the corresponding elasticity of another firm and if the ratio of the variance of the elasticity of one firm is less than the corresponding ratio of the other firm, then the aggregate production function obtained by taking the mean elasticities of that input is more efficient than the production function obtained by taking the G.M. of the elasticities.

Another method of aggregation is to find the weighted arithmetic mean of the micro-elasticities of an input with weights inversely proportional to their variances. The aggregate production function will then become

$$S=A_1x_1^{C_1}x_2^{C_2}\ldots x_n^{C_n}$$

where A_1 is a constant and c_1, c_2, c_n are the macro-elasticities given by

$$c_1=\frac{\dfrac{a_1}{V(a_1)}+\dfrac{b_1}{V(b_1)}}{\dfrac{1}{V(a_1)}+\dfrac{1}{V(b_1)}}=\frac{a_1V(b_1)+b_1V(a_1)}{V(a_1)+V(b_1)}$$

Similarly

$$c_2=\frac{a_2V(b_2)+b_2V(a_2)}{V(a_2)+V(b_2)}$$

$$\ldots \qquad \ldots \qquad \ldots$$

$$c_n=\frac{a_nV(b_n)+b_nV(a_n)}{V(a_n)+V(b_n)}.$$

Taking the logarithms of the macro-function we have

$$\log S=\log A_1+c_1\log x_1+c_2\log x_2+\ldots+c_n\log x_n$$

The variances for ($c_i i=1, 2,\ldots n$) can be worked out as follows :—

$$\text{Var }(c_1)=\left[\frac{V(b_1)}{V(a_1)+V(b_1)}\right]^2V(a_1)+\left[\frac{V(a_1)}{V(a_1)+V(b_1)}\right]^2V(b_1)$$

$$=\frac{[V(b_1)]^2V(a_1)}{[V(a_1)+V(b_1)]^2}+\frac{[V(a_1)]^2V(b_1)}{[V(a_1)+V(b_1)]^2}$$

$$=\frac{V(a_1)V(b_1)[V(a_1)+V(b_1)]}{[V(a_1)+V(b_1)]^2}$$

$$=\frac{V(a_1)V(b_1)}{V(a_1)+V(b_1)}.$$

Similarly

$$\text{Var}\,(c_2)=\frac{V(a_2)V(b_2)}{V(a_2)+V(b_2)}$$

$$\ldots \qquad \ldots \qquad \ldots$$

$$\text{Var}\,(c_n)=\frac{V(a_n)V(b_n)}{V(a_n)+V(b_n)}.$$

Comparing the variances of $c_1, c_2, \ldots c_n$ with those of the macro elasticities

$$\frac{a_1+b_1}{2}, \frac{a_2+b_2}{2}, \ldots \frac{a_n+b_n}{2}$$

obtained by taking the arithmetic mean of micro-elasticities we observe that

$$\text{Var}\,(c_1) \text{ is always less than } V\left(\frac{a_1+b_1}{2}\right)$$

as $$\frac{V(a_1)V(b_1)}{V(a_1)+V(b_1)} < \frac{V(a_1)+V(b_1)}{4}$$

or $$2V(a_1)V(b_1) < [V(a_1)]^2+[V(b_1)]^2$$

or $$[V(a_1)-V(b_1)]^2 > 0$$

which is true.

Thus aggregation by this procedure is more efficient than aggregation by averaging.

Another way of developing an aggregation production function is by weighting the micro-elasticities by the number of farms or firms on which each production function is based and taking the arithmetic mean.

Let the two functions be given by

$$y_1=A_1x_1^{a_1}\ x_2^{a_2}\ldots x_n^{a_n}$$

$$y_2=A_2x_1^{b_1}\ x_2^{b_2}\ldots x_n^{b_n}$$

We shall take the aggregate production function as

$$y_3=A_3x_1^{c_1}\ x_2^{c_2}\ldots x_n^{c_n}$$

where

$$c_1=\frac{n_1a_1+n_2b_1}{n_1+n_2}$$

$$\ldots \qquad \ldots \qquad \ldots$$

$$c_n=\frac{n_1a_n+n_2b_n}{n_1+n_2}$$

We will obtain the expression for the variance of $c_1, c_2, \ldots c_n$

$$\text{Var } c_1 = \text{Var}\left(\frac{n_1 a_1 + n_2 b_1}{n_1 + n_2}\right)$$

$$= \frac{1}{(n_1+n_2)^2}\left[n_1^2 \text{ Var }(a_1) + n_2^2 \text{ Var }(b_1)\right]$$

$$\text{Var } c_2 = \frac{1}{(n_1+n_2)^2}\left[n_1^2 \text{ Var }(a_2) + n_2^2 \text{ Var }(b_2)\right]$$

$$\ldots \qquad \ldots \qquad \ldots$$

$$\text{Var } c_n = \frac{1}{(n_1+n_2)^2}\left[n_1^2 \text{ Var }(a_n) + n_2^2 \text{ Var }(b_n)\right]$$

Comparing these variances with the variance of the macro-elasticities obtained by the unweighted arithmetic mean we find that the present method of aggregation will be more efficient provided

$$\frac{n_1^2}{(n_1+n_2)^2} \text{ Var }(a_1) + \frac{n_2^2}{(n_1+n_2)^2} \text{ Var } b_1 \leqslant \frac{\text{Var }(a_1)}{4} + \frac{\text{Var }(b_1)}{4}$$

or $$\frac{4n_1^2}{(n_1+n_2)^2}\frac{V(a_1)}{V(b_1)} + \frac{4n_2^2}{(n_1+n_2)^2} \leqslant \frac{V(a_1)}{V(b_2)} + 1$$

or $$\frac{V(a_1)}{V(b_1)}\left[\frac{4n_1^2}{(n_1+n_2)^2} - 1\right] \leqslant 1 - \frac{4n_2^2}{(n_1+n_2)^2}$$

i.e. $$\frac{V(a_1)}{V(b_1)}(3n_1^2 - n_2^2 - 2n_1n_2) \leqslant n_1^2 - 3n_2^2 + 2n_1n_2$$

i.e. $$\frac{V(a_1)}{V(b_1)}(3n_1 + n_2)(n_1 - n_2) \leqslant (n_1 - n_2)(n_1 + 3n_2).$$

Thus if $n_1 > n_2$

$$\frac{V(a_1)}{V(b_1)} \text{ should be } \leqslant \frac{n_1+3n_2}{3n_1+n_2}$$

and if $n_2 > n_1$

$$\frac{V(a_1)}{V(b_1)} \text{ should be } \geqslant \frac{n_1+3n_2}{3n_1+n_2}$$

in order that the present method of aggregation is more efficient than simple averaging.

Yet another method of developing an aggregate production function is by weighting the elasticities with output. The aggregate elasticity in this case works out to be

$$\frac{a_1 y_1 + b_2 y_2}{y_1 + y_2}$$

and the variance works out to be

$$\left(\frac{y_1}{y_1+y_2}\right)^2 V(a_1)+\left(\frac{y_2}{y_1+y_2}\right)^2 V(b_1)$$

where y_1 and y_2 are the outputs of the two places or farms or firms. Money value of output can also be used in this case as weights.

The aggretation of A values in the case of the simple a.m. is obtained by taking the geometric mean. In other cases since A is the value of the output for unit values of the inputs $x_1, x_2, \ldots x_n$, the aggregate value can be taken once the aggregate elasticities are known.

The aggregation problem for the SMAC production function has also been investigated. The basic change introduced by SMAC is to allow the elasticity of substitution, σ, to be a constant at a value other than unity (Cobb-Dougals) or zero (input-output). The function is

$$X=\gamma\left[\delta K^{-\rho}+(1-\delta)L^{-\rho}\right]^{-\frac{1}{\rho}}.$$

This function is linear and homogeneous *i.e.* there are constant returns to scale. The efficiency parameter is γ which changes output for given quantities of input ; the distribution parameter is

$$\delta(0\leqslant\delta\leqslant 1)$$

which determines the division of factor income. The substitution parameter ρ is a simple function of the elasticity of substitution and thus $\sigma=1/(1+\rho)$. The marginal product of capital is

$$\delta\gamma^{-\rho}(X/K)^{1+\rho}.$$

The limits to the value of ρ are derived from σ. The function is additivity separable when written in the form,

$$\gamma^{\rho} X^{-\rho}=\delta L^{-\rho}+(1-\delta)K^{-\rho}.$$

Since the variances of the statistics cannot be estimated, the aggregation method is not discussed at present.

One of the main problems of aggregation is that the data are usually published in the form of arithmetic averages or totals, whereas our system of aggregation requires geometric averages. If the deviations from the mean value are relatively small we can show that the geometric mean G is approximately related to the arithmetic mean by the formula

$$G=\bar{X}\left(1-\frac{1}{2}\frac{\sigma^2}{\bar{X}^2}\right)$$

In this case the a.m. gives an upward relative bias which is approximately half the relative variance.

The biases in the estimates of elasticities will be different if we adopt different methods of fitting the production function. The behaviour of these biases under aggregation is an interesting study which is under investigation. In this paper aggregation methods are developed for the Cobb-Dougals production function under the assumption that the functions have been fitted already by the least square technique. The behaviour of aggregation under the covariance matrix method, factors-shares method and instrumental variables method are under investigation.

The investigations in this paper are illustrated with the help of data from studies in the economics of farm management in West Godavari District. In this study 104 farms have been taken out of which 67 belong to paddy zone and 37 to tobacco zone. The variables included in the study are y=value of all crops raised on each farm (in Rs.), $x1$ (land in standard acres), $x2$, (human labour in man-days), $x3$, (Total capital invested in agriculture excluding the value of land) and $x4$ (Production expenses excluding the value of human labour). The production functions derived for the different zones are the following

Paddy zone

$$y=45{\cdot}8575\ x_1^{0{\cdot}49001}\ x_2^{0{\cdot}22676}\ x_3^{0{\cdot}00006}\ x_4^{0\ 27228}$$

Tabacco zone

$$y=6{\cdot}60753 x_1^{\cdot 17096} x_2^{0{\cdot}34883}\ x_3^{0{\cdot}01603}\ x_4^{0{\cdot}50631}.$$

The elasticities along with the variance for the different outputs worked out by the different procedures are given in the table. Pooling the elasticities by using the reciprocal of the variances in seen to be perhaps the best method here after adding all inputs and outputs.

SUMMARY

The problem of aggregation of the Cobb-Dougals production functions and also the other types are briefly discussed in this paper. Six methods of aggregation are presented of which aggregating by using the inverse of the variances as weights is seen to be the most efficient method after the method of adding all outputs and inputs. The behaviour of aggregation under different systems of fitting is under investigation. How the SMAC behaves under aggregation is also being investigated.

REFERENCES

1. A.A. Walters (1963). Production and Cost Functions Econometrica (Vol. 31), Number 1-2, pp. 1-66.

2. J. Kmenta and M.E. Joseph (1963). A Monte-carlo Study of Alternative Estimates of Production Functions. Economictrica Vol. 31, No. 3.

3. K.J. Arrow, H.B. Chenery, B.S. Minhas and R.M. Slow (1961). Capital Labour Substitution and Economic efficiency. Review of Economics and Statistics, Vol. 43, pp. 225-250.

4. B.S. Rao ('57-'58). Studies in Economics of Farm Management in West Godavari District, Andhra Pradesh, Published by the Directorate of Economics and Statistics, Ministry of Food and Agriculture, Government of India.

5. C.R. Rao (1952). Advanced Statistical Methods in Biometric Research. J. Wiley and Sons, N. York.

6. Klein, L.R. (1946). Remarks on the Theory of Aggregation. Econometrica, Vol. 14, pp. 303-312.

TABLE

Combined Production Function

$$Y = A\, x_1^{b_1}\; x_2^{b_2}\; x_3^{b_3}\; x_4^{b_4}$$

Method	A	x_1		x_2		x_3		x_4	
		b_1	$v(b_1)$	b_2	$v(b_2)$	b_3	$v(b_3)$	b_4	$v(b_4)$
Adding all inputs and outputs.	26·0823	0·4236	0·004948	0·23996	0·006874	0·C0071	0·000745	0·34660	0·002864
A.M. of Elasticities.	9·461	0·33048	0·006856	0·28780	0·013026	0·00804	0·001003	0.38930	0·006106
G.M. of Elasticities.	11·82	0·28950	0 013587	0·28130	0·011389	0·00099	0·155159	0·37140	0·005108
Elasticities pooled with weights as inverse variances.	15·27	0·37790	0·006251	0·25759	0·009835	0·00931	0·000978	0·32517	0·004273
Elasticities pooled with weights as number of farms.	12·21	0·37650	0·006251	0·27019	0·010391	0·00574	0·001178	0·35554	0·004684
Elasticities pooled with weights as outputs.	8·56	0·31244	0·007405	0·29470	0·014652	0·00895	0·000980	0·40253	0·006942