Issued Under Office of Naval Research
Contract No. Nonr- 1858(16)

AGGREGATION OF PRODUCTION FUNCTIONS—
SOME EXPERIMENTAL RESULTS

John G. Cragg

Econometric Research Program
Research Paper No. 10
January 19, 1965

The research described in this paper was supported by the Office of Naval Research. Reproduction, translation, publication, use and disposal in whole or in part by or for the United States Government is permitted. The research described in this paper was performed while the author was a member of the Econometric Research Program, Princeton University, Princeton, N. J.

Princeton University
Econometric Research Program
92-A Nassau Street
Princeton, N. J.

ABSTRACT

Data available for use in estimating Cobb-Douglas production functions are frequently aggregated unsuitably. This paper reports the results of some distribution-sampling experiments conducted to investigate the effects of unsuitable aggregation. It is found that the usual least-squares estimates are biased. Misaggregation also seriously affects the performances of certain tests of hypotheses which were investigated.

Aggregation of Production Functions—
Some Experimental Results

1. Introduction

The production function which is most widely used in econometrics is undoubtedly the Cobb-Douglas (C-D) function:

$$(1) \qquad \log Y_t = \alpha_o + \alpha_1 \log X_{1t} + \alpha_2 \log X_{2t} + \dots + \alpha_n \log X_{nt} + u_t$$

where Y_t is an output (or production) at observation t, the X_{it} are inputs at the same observation and u_t is a random disturbance having the properties of disturbances in the usual regression model. The widespread use of the C-D function can probab ly be attributed to its simple form for estimation purposes, to the ease with which it can be manipulated mathematically, to its property of being homogeneous of degree $h = \sum_{i=1}^{n} \alpha_i$ and to its having constant elasticities of substitution between the input variables. But though it gives these convenient properties, a C-D specification for the production function (or for other economic functions) does entail the acceptance of rather strong economic assumptions about production relationships. It would seem to be desirable to test the hypothesis that production data are generated by a C-D function rather than by some other, especially since the C-D is not the only production function readily available for econometric use.

The alternative function we shall concentrate on in this paper is the "transcendental" production function of Halter, Carter and Hocking [1], referred to here as the H-C-H function:

$$\text{(2)} \qquad \log Y_t = \beta_o + \beta_1 \log X_{1t} + \ldots + \beta_n \log X_{nt} + \gamma_1 X_{1t} + \ldots + \gamma_n \log X_{nt} + \epsilon_t$$

where ϵ_t is defined as was u_t in (1). This function has none of the main properties of the C-D function unless all the $\gamma_i = 0$ — when it is the C-D function. In particular, it is not homogeneous and the elasticity of substitution is not a constant but depends on the values of the X_{it}. The H-C-H function does have the property that if $\beta_i > 1$, $\gamma_i \;\; 0$ ($i=1, \ldots, n$), it exhibits first increasing and then decreasing returns to scale. The H-C-H function is, of course, only one of any number of algebraic forms that the production function might take and may well have a great deal less appeal than the C-D function. However, the facts that the C-D function is only a special case of the H-C-H function and that the H-C-H function has, in general, quite different properties from the C-D make it of interest to test whether production data are explained better by the H-C-H or the C-D functions. The test suggested by the forms of equations (1) and (2) is the usual F-test of the hypothesis that $\gamma_1 = \gamma_2 = \ldots = \gamma_n = 0$. But in using the test for such a purpose, we would face the problem that aggregation of the data used in performing the test may make invalid the conclusions to be drawn.

The data available for estimating production functions are usually aggregated. Aggregation is performed both over inputs of similar kinds (e.g., over types of capital or of labor), over production processes (that is, the outputs and inputs of a number of separate production processes are aggregated to get establishment, firm or industry data) and over the same production processes being operated at different moments of time.

The correct way to aggregate for the C-D function is geometrically— that is, the aggregate variables should be formed by summing the logarithms of the original variables.[1] For the H-C-H function there is no single correct way to aggregate the original variables. However, if for use in the aggregate function the variables to be used for the logarithmic terms of (2) are aggregated from the original variables geometrically while those to be used for the terms in the original inputs are aggregated arithmetically, the production function in these new variables has the same coefficients as the production function in the original, disaggregated variables. While these considerations point up the problems of correct aggregation, the economist is usually faced with data which have been aggregated arithmetically. The ability of the $\underline{F}$-statistic to provide a reliable

---

[1]See [2] and [3] for discussion of correct aggregation. By "correctness" is meant, essentially, that the coefficients of the production function in the aggregated variables are the same as in the original function while the disturbances have the same properties as the original disturbances.

test of the hypothesis that the C-D rather than the H-C-H function is the correct production function when used with inappropriately aggregated data may be doubtful. The effects of inappropriate aggregation on the bias, dispersion and standard errors of the estimates of the C-D function are also of interest. Since the problems seemed intractible to mathematical analysis, it was decided to explore the subject with some distribution-sampling experiments.

2. Design of the Experiments

Sampling experiments were conducted to investigate the effects of inappropriate aggregation on estimation of the C-D and H-C-H production functions and of the ability of the F-test to distinguish between them. The data for the experiments were generated as follows.

For each observation five "output" variables, Y_i, were generated from two "input" variables each, X_{1i} and X_{2i}, and an "error" term, u_i, by each of the functions:

$$\text{(1.a)} \quad \log \bar{Y}_i^{C} = \alpha_i \log X_{1i} + \alpha_{2i} \log X_{2i} + u_i \quad (i = 1, \ldots, 5)$$

$$\text{(2.a)} \quad \log \bar{Y}_i^{H} = \beta_1 \log X_{1i} + \beta_2 \log X_{2i} + \gamma_{1i} X_{1i} + \gamma_2 X_{2i} + u_i \quad (i = 1, \ldots, 5)$$

The "inputs," X_{ij}, were formed from independently and rectangularly distributed pseudo-random numbers. In most experiments the expected correlation between any two "inputs" of the same type was .87. However, in one experiment, referred to as experiment 0, the correlations were zero. The X_1 and the X_2 variables were distributed independently of each other. The "error" terms, $\underline{u}_i$, were normally and independently distributed pseudo-random deviates of mean zero and variance equal to ten per cent of the variance of the C-D output variable. The parameters of the C-D and the H-C-H functions used were held constant so that the problem of inappropriate forms of aggregation should not be beclouded by other aggregation problems. (Cf. [4].)

Aggregations were performed (arithmetically and geometrically) to group the five "outputs" of each type and the five $\underline{X}_1$ and $\underline{X}_2$ "inputs" to obtain an aggregated "observation" for use in estimating an "aggregated production function." Four different "observations" were obtained by aggregating the C-D and the H-C-H data "correctly" and by aggregating all the variables for use in estimation arithmetically.[2] The different types of data can be taken to correspond to the data economists would like to have and those which they usually are able to get.

Least-squares estimates of the coefficients of the functions and tests of various hypotheses were made from fifty such aggregated observations. These observations were independent of each other. In each experiment the process of generating observations and making estimates was repeated one hundred times to obtain samples of estimates and of the performances of the test statistics.

Four combinations of parameters were used in different

---

[2]That is, the arguments of the C-D function were of the form $\log \Sigma \underline{Q}_i$ rather than $\Sigma \log \underline{Q}_i$. The "incorrect" arguments for the H-C-H function were of the form $\log \sum_i \underline{Q}_i$ and $\sum_i \underline{Q}_i$ rather than $\sum_i \log Q_i$ and $\sum_i Q_i$.

experiments. These parameters and the experiments in which they were used are shown in Table 1.

TABLE 1

Parameters Used in the Experiments

Experiment	α_0	α_1	α_2	β_0	β_1	β_2	γ_1	γ_2
0 and 1	0	.65	.35	0	7.5	4.5	-.035	-.025
2	0	.75	.45	0	4.5	3.5	.035	.025
3	0	.6	.3	0	.95	.85	-.005	-.005
4	0	.95	.05	0	.95	4.5	.035	-.025

3. Aggregation in the C-D Function

Incorrectly aggregating the variables biased the estimates of the coefficients of the C-D function. Table 2 shows the averages and the standard deviations of the estimates of the C-D function made from the correctly and incorrectly aggregated C-D data. In all cases incorrect aggregation biased the estimates of the constant downwards and those of the other coefficients upwards. This tendency was a good deal more pronounced in experiment 0, where the inputs to be aggregated were uncorrelated with each other, than in the other experiments. It is also interesting that in experiment 4, where one of the inputs played a very minor role in generating output, the bias was slight and for the coefficient of the second input was not significant even at the .25 level. All other biases were significant beyond the .005 level. It is also worth noting that the dispersions of the estimates , as measured by their standard deviations, were larger when the data

TABLE 2

Estimates of Coefficients of the C-D Function; Means and Standard Deviations

Experiment	Coefficient	True Value	Correctly Aggregated		Incorrectly Aggregated	
			Means	Standard deviations	Means	Standard deviations
0	α_o	0	-.009	.107	-.931	.409
	α_1	.65	.651	.005	.723	.044
	α_2	.35	.350	.005	.429	.046
1	α_o	0	-.010	.115	-.229	.148
	α_1	.65	.652	.011	.670	.050
	α_2	.35	.350	.012	.368	.054
2	α_o	0	-.013	.138	-.589	.181
	α_1	.75	.752	.014	.772	.061
	α_2	.45	.449	.014	.472	.066
3	α_1	0	-.009	.104	-.047	.132
	α_2	.6	.601	.010	.619	.044
	α_3	.3	.299	.010	.316	.048
4	α_1	0	-.010	.119	-.047	.049
	α_2	.95	.952	.012	.956	.016
	α_3	.05	.049	.012	.052	.017

were incorrectly aggregated. This tendency was most pronounced for experiment 0 and least in experiment 4. (In experiment 4, indeed, the standard deviation of the estimates of the constant was smaller when inappropriately aggregated data were used than when the correctly aggregated set was employed.)

One of the more interesting tests that can be made with the C-D production function is the t-test for constant returns to scale. (That is, of the hypothesis that $\alpha_1 + \alpha_2 = 1$.) In experiments 0, 1, and 4 the function used to generate the data exhibited constant returns to scale. Table 3 records the frequencies with which the hypothesis of constant returns to scale was rejected at the .05 level on a two-tailed test using the correctly and incorrectly aggregated data. When the data were aggregated incorrectly, the hypothesis was rejected more frequently than when correct aggregation was performed. This increase in Type I error was most pronounced in experiment 0. It was least evident in experiment 4. In experiments 2 and 3 misaggregation only slightly

TABLE 3

Frequence of Rejection of Hypothesis of Constant Returns to Scale

Experiment	Correct Aggregation	Incorrect Aggregation
0	6	47
1	5	24
4	5	11

reduced the power of the test. The differences between the performances of the test on the two types of data were not sufficient to allow the conclusion that the power of the test was reduced. This may be because the C-D functions in these experiments had returns to scale substantially different from constant. In all experiments misaggregation led to the t-test rejecting more frequently the hypotheses that the values of α_1 and α_2 were indeed the ones used to generate the data. The differences were similar to those shown in Table 3 for the test of constant returns to scale. The results of experiments 2 and 3 were very similar to those for experiment 1.

4. Comparisons with the H-C-H function

The H-C-H function was estimated from correctly aggregated data and from inappropriately aggregated data. The estimates of some of the coefficients of the H-C-H function were biased significantly when unsuitably aggregated data were used. This was not, however, true for all the coefficients. Table 4 records the means and standard deviations of the estimates made from the incorrectly aggregated data and indicates which of the biases were significant. The most pronounced aspects of the table are the large downward biases of the estimates of γ_1 and γ_2 and the lack of significant biases in the estimates of β_1 and β_2. The standard deviations of the estimates were apt to be a great deal larger when the data were inappropriately aggregated than when suitable aggregation was performed.

TABLE 4

Estimates of the H-C-H Function Made from Incorrectly Aggregated Data: Means and Standard Deviations

Experiment	Quantity	β_0	β_1	β_2	γ_1	γ_2
0	True Value	0	7.5	4.5	-.035	-.025
	Mean	-47.26*	6.201*	3.913	-.0104*	-.0046*
	Standard deviation	28.44	4.615	4.615	.0180	.0181
1	True Value	0	7.5	4.5	-.035	-.025
	Mean	-15.82*	7.207	4.504	-.0082*	-.0043*
	Standard deviation	12.77	3.406	3.876	.0131	.0149
2	True Value	0	4.5	3.5	.035	.025
	Mean	-14.57*	5.050	3.826	.0033*	.0039*
	Standard deviation	12.88	3.249	3.843	.0124	.0147
3	True Value	0	.95	.85	-.005	-.005
	Mean	-18.05*	.927	.965	-.0008*	-.0011*
	Standard deviation	18.42	.527	.603	..0021	.0024
4	True Value	0	.95	4.5	.035	-.025
	Mean	-5.933*	.830	4.376	.0074*	-.0048*
	Standard deviation	5.339	1.. 370	1.750	.0054	.0068

* Significant at the .005 level. Unstarred entries are also not significant at the .05 level.

The hypothesis that the data were generated by a C-D function rather than by an H-C-H function was tested using the data generated by each type of function. In Table 5 are recorded the frequencies with which the F-test rejected the hypothesis at the .05 significance level for the correctly and incorrectly aggregated data of each type. Two things stand out from the table. First, incorrect aggregation led to some increase in the occurrence of Type I error. Second, the test incorrectly accepted the hypothesis that H-C-H data were generated by a C-D function much more frequently when mis-aggregation occurred than when correct aggregation was performed. Both these findings suggest that drawing conclusions from incorrectly aggregated data is a very precarious undertaking.

TABLE 5

Frequency of Rejection of Hypothesis that Production Function Was C - D

Experiment	C-D Generation of Data		H-C-H Generation of Data	
	Correct Aggregation	Incorrect Aggregation	Correct Aggregation	Incorrect Aggregation
0	1	9	100	14
1	3	16	100	40
2	3	15	100	16
3	3	16	100	36
4	3	8	100	33

5. Conclusion

The experiments conducted are, of course, subject to strong limitations. The method of generating the data far from reflects fully the complexities of the processes from which data might come for use in estimating production functions. Nevertheless, the results of the experiments do suggest some conclusions about the effects of incorrect aggregation of the data for use in estimating C-D production functions. Although the incorrect aggregation of the data did significantly bias the estimates, the sizes of the biases are probably not so large as to render estimates of the coefficients of C-D production functions useless. On the other hand, the poor results of the F-test for distinguishing the C-D function from the H-C-H and of the t-test for constant returns to scale indicate that inappropriate aggregation can destroy the usefulness of standard tests of hypotheses. The experiments indicate that it is precarious to ignore the dangers of incorrect aggregation when double-log regressions are used.

REFERENCES

1. Halter, A. N., H. O. Carter and J. G. Hocking, " A Note on the Transcendental Production Function," Journal of Farm Economics, Vol. 39 (1957), pp. 966-974.

2. Nataf, André, "Sur la Possibilité de Construction de Certains Macro-Modèles, "Econometrica, Vol. 16 (1948), pp. 232-244.

3. Klein, Lawrence R., "Remarks on the Theory of Aggregation," <u>Econometrica,</u> Vol. 14 (1946), pp. 303-312.

4. Theil, H., <u>Linear Aggregation of Economic Relationships,</u> Amsterdam: North-Holland, 1954.