CLUSTER ANALYSIS AS A PREPROCESSOR FOR FITTING AGGREGATE PRODUCTION FUNCTIONS

Rafael Solis, Craig School of Business, California State University, Fresno, 5245 N. Backer Ave. Fresno, CA 93740. (559)278-2194, rafaels@csufresno.edu

K.C. Tseng,, Craig School of Business, California State University, Fresno, 5245 N. Backer Ave. Fresno, CA 93740. (559)278-4981, kctseng@csufresno.edu

ABSTRACT

In this paper we discuss the innovative approach of clustering algorithms to group countries based on a set of macroeconomic variables and their effect on the estimation of the parameters of a Cobb-Douglas type production function. We found that the clustering algorithms did produce some differences on the parameters estimation as compared to simple grouping of the countries based on per-capita income as is traditionally done by most economists. We suggest that the utilization of clustering methodologies could play a more important role when a better refined grouping of countries may be needed without adversely affecting the meaning of the parameters of the fitted Cobb-Douglas functions.

METHODOLOGY AND DATA SOURCES

Since we have not found any cross-sectional empirical study on aggregate production function at the national level in recent years, we have decided to apply the well known and tested Cobb-Douglas production function to the present study. In addition, this special function can be transformed into log-linear form with linearity in parameters so that standard statistical hypothesis testing can be conducted by using the ordinary least squares estimation method. As pointed out by Zellner, Kmenta, and Dreze [16], the least squares estimators are unbiased and consistent if random errors are normally distributed and independent across equations. The estimated parameters are the elasticities of aggregate output with respect to the inputs included in the equation. With properly defined input and output variables and under perfect competition and constant returns to scale, the exponent associated with each input indicates that factor share of national income of that particular factor of production. This is derived from the fact that under competitive markets, there exist no excess profit and all factors of production will exhaust national product. Furthermore, the summation of the exponents associated with all inputs included in the equation shows the returns to scale. For example, if the summation of exponents is greater (less) than one, there exist economies (diseconomies) of scale. The estimated parameter (exponent) is the output elasticity with respect to a particular input (factor). Finally, based on several time series empirical studies cited above and others, the Cobb-Douglas production function has performed well.

The data collected consisting of 70 countries and 9 variables (Table 1). Later two variables were dropped and some were transformed for the cluster analysis. Other variables were utilized for the fitting of the Cobb-Douglas production functions. The variables utilized for the cluster analysis were: GDP_PERC (GDP growth in percent), GDP_CAP (GDP per capita), ELEC_CAP (Electricity production per capita) and UNEMPLOY (unemployment rate). These variables were standardized due their different measurement scales.

Table 1
Macroeconomic Variables

Variable Name	Description
AREA_SQ	Country area (sq. km.)
ARABLE_L	% of land dedicated to agriculture (1996)
POPULATI	Population (1996 in million)
GDP	Gross Domestic Product (1996) in $ billion)
GDP_PERC	GDP growth in % (1995-1996)
GDP_CAP	GDP per capita (in 1996 $)
ELECT	Electricity Production (1996 billion of KwH)
LABOR_FO	Labor force (1996 in million)
UNEMPLOY	Unemployment rate (1996)

SUMMARY AND CONCLUSIONS

The study applies the well known and tested Cobb-Douglas production function to cross-sectional data from a reliable CIA data source. We have utilized multivariate clustering techniques to classify countries into similar clusters so that countries within the same cluster will be less heterogeneous and more suitable to fit into a production function. We have discussed the advantages and disadvantages of clustering analysis since this technique has never been applied to production function studies. Based on the multivariate clustering techniques all 70 countries are put into two groups. In addition, the traditional univariate approach used by economists based on variable such as GDP per capita is also used to separate the developed from the developing countries. The empirical findings from both approaches are somewhat different but quite similar.

The empirical findings from the Cobb-Douglas production function are consistent with theoretical expectations. First, Labor contributes much more than electricity production to GDP for both the developed and the developing countries. Second, technical differences among the developed countries are greater than those among the developing countries although the technical differences among the developed or the developing countries are also quite substantial. Third, the multivariate clustering techniques are a useful method for grouping countries into similar clusters/groups. More applications can be tested particularly in more complicated studies with large number of variables and observations. Finally, since there is a lack of cross-sectional studies on production functions, we hope this study may induce similar works in the future. Indeed, the implications from this type of studies are significant for international trade, factor movements and technical transmission. It is worthwhile continuing further researches in the future.

REFERENCES

[1] Anderberg, M.R., *Cluster Analysis for Applications.* Academic Press. 1973.

[2] Arrow, K.J., H.B. Chenery, B.S. Minhas, and R.M. Solow., "Capital-Labor Substitution and Economic Efficiency," *Review of Economics and Statistics* (August 1961), pp. 225-250, in A. Zellner(ed.), *Readings in Economic Statistics and Econometrics,* Boston:Little, Brown and Co. 1968, pp. 364-408.

[3] Basu, S. and J.G. Fernald, "Returns to Scale in U.S. Production: Estimates and Implications," *Journal of Political Economy*, 105 (1997), pp.249-283.

[4] Beckman, M.J. and R. Sato, "Aggregate Production Functions and Types of Technical Progress: A Statistical Analysis," *American Economic Review,* 59 (March 1969), pp. 88-101.

[5] Blackorby, C. and W. Schorm, "The Existence of Input and Output Aggregates in Aggregate Production Functions," *Econometrica*, 56 (1988), pp.613-643.

[6] Fisher, F.M., *Aggregation: Aggregate Production Functions and Related Topics.* Collected Papers of Franklin M. Fisher, edited by J. Monz, Cambridge: MIT Press, 1993.

[7] Fisher, W.D., *Clustering and Aggregation in Economics*. Johns Hopkins Press, Baltimore. 1969.

<most content and references were omitted in this 3-page version of the paper>