The Philippine Statistician
Vol. xxiii, Nos. 3 - 4, 1974

A STATISTICAL TEST OF TWO-COUNTRY PRODUCTIVITY DIFFERENCES*

by Florian A. Alburo**

I

This paper is the result of an attempt to test the empirical validity of standard neoclassical assumptions explaining international differences in economic efficiency. The results in brief, created more problems although in the course of the experiment, a few statistical and theoretical observations showed themselves worthy of further experimentation.

Specifically, we shall (a) provide a brief framework of the existing theory, (b) define the course in which we shall conduct the test, (c) present our results and their interpretations, and (d) point out the inherent problems and their suggested direction for further work.

II

International trade theory explaining international productivity differences generally and implicitly build a model around two crucial assumptions[1]: (a) all firms within the country and across national boundaries employ the same neoclassical production function, and (b) all factors are homogeneous and ferfectly mobile. In a way, these assumptions are too restrictive and attempts have been made to relax these assumptions especially with respect to trade in specific manu-

---

* Professors Robert NcNown and Fred Glahe read an earlier draft of this paper while Herminio Blanco (now at Chicago) helped in programming the data. They are not responsible however for errors or inconsistencies here.

** Visiting Research Associate, U.P. School of Economics

[1] See any standard text on international trade theory. For example, B. Ohlin, *Interregional and International Trade* (Cambridge, 1933) or G. Haberler, *The Theory of International Trade* (London, 1936).

factured goods.[2] Yet in the more formal models the above assumptions are still employed.[3]

Given the assumptions and (assuming) perfect competition in all markets, the industry then becomes the representative firm, with factor prices and factor returns the same in all markets in a given country. The differences in productivity between countries are then due to differences in factor proportions employed. All countries lie on the same given common production function though they may be at different points at different times. It then becomes a short step to conclude the comparative advantage basis of trade.

In the context of economic development, this theory has been critically examined and in many ways has either been abandoned or replaced by a surrogate.[4]

III

The classic approach to the theory is the 2-country, 2-factor, 2-commodity model. In this paper we shall essentially employ the same model, the 2-countries being the United States and the Philippines.[5] We postulate a linear homogenous production function common to both countries, the arguments of which are labor and capital.[6] And the industry we shall consider is the 2-digit manufacturing industry.

---

[2] See for example, C.P. Kindleberger, *International Economics* 3rd edition (Homewiod, Ill.: R.D. Irwin, 1963), Chapter 7, *Idem, Foreign Trade and the National Economy* (New Haven: Yale University Press, 1962), Chapter 4, P.T. Ellsworth, *The International Economy* 3rd edition (New York: Wiley, 1964), Chapter 9, and S. Linder, *An Essay on Trade and Transformation* (New York: Wiley, 1961), pp. 90-92.

[3] R.R. Nelson says that even though "several papers admit the possibility of total productivity differences across nations . . . this is brought in an empirical fact of life, not as something intrinsic to the basic model." R.R. Nelson, "A 'Diffusion' Model of International Productivity Differences in Manufacturing Industry." *American Economic Review* 58 (December 1968), p. 1220 fn. 2.

[4] See C.P. Kindleberger, *Economic Development* 2nd edition (Tokyo: Kogakusha, 1965), Chapter 16.

[5] Our choice of countries is not constrained by the fact that the U.S. and Philippine economy are at both extremes in many magnitudes but for two reasons: (1) Philippine foreign trade with the U.S. is more than 30 per cent of total trade though not vice versa, and (2) if the above theory provides the basis of trade it would hold for any n, m country.

[6] There are of course other arguments, discussed below.

Our measure of productivity is value added per employee.[7] This is obtained by dividing the total value added of all firms within an industry by its total number of employees. There are of course statistical differences in the measurement of value added.[8]

Capital as used here is the book value of fixed assets for the year in which the census was conducted. And total employment are all working owners and unpaid family workers, production and related workers, including all other employees on paid holidays and paid vacations during the period for which data were gathered.

Exchange rate used for the conversion of Philippine data values into dollars was the free rate quoted by the International Monetary Fund.[9]

Consider a production function:

$$(1) \qquad Q_{ij} = f(K_{ij}, L_{ij}) \qquad f_K > 0 \quad f_{KK} < 0 \qquad f_L > 0 \quad f_{LL} < 0$$

where Q_{ij} is output of country i in industry j, K_{ij} is the value of fixed assets of country i in industry j and L_{ij} is total employment of country i in industry j. The assumption of diminishing marginal products is also posited.

If we follow the neoclassical assumptions and postulate (1) to be linear homogeneous, we can rewrite (1) as:

$$(2) \qquad \left[\frac{Q}{L}\right]_{ij} = f\left\{\left[\frac{K}{L}\right]_{ij}\right\} \quad f' > 0 \quad f'' < 0$$

[7] A better measure would perhaps be value added per man-hour but data on Philippine labor man-hour were not available during the conduct of this test.

[8] Value added in the U.S. is an adjusted figure.

[9] *International Finance Statistics* (Washington D.C.: International Monetary Fund, July 1968), p. 262.

where the symbols are similarly defined as in (1) above. This expresses output per worker as now a function of the factor proportion (i.e. the capital-labor ratio) employed.

Equations (1) and (2) however are general equations. We need to further specify the function. Two choices are available: The Cobb-Douglas or the more general CES production function. Since (a) the CES has been used in another study[10] and (b) it has been shown that the results of the use of the CES do not significantly differ from the results of the use of the simpler (more restrictive Cobb-Douglas, we employ the Cobb-Douglas production function.[11]

With a Cobb-Douglas function of the form

$$(3) \qquad Q_{ij} = \alpha K_{ij}^{B_1} L_{ij}^{B_2}$$

where α is an index of total productivity growth and the B's are the elasticity coefficients, and the assumption of linear homogeneity and that the elasticities add up to unity, (3) can be expressed as:

$$(4) \qquad \left[\frac{Q}{L}\right]_{ij} = \alpha \left[\frac{L}{K}\right]_{ij}^{B_1} \qquad B_2 = 1 - B_1$$

Transforming (4) into logs yields:

$$(5) \qquad \log \left[\frac{Q}{L}\right]_{ij} = \log \alpha + B_1 \log \left[\frac{K}{L}\right]_{ij}$$

Equation (4) is now expressed as a linear equation in logs to which the methods of least squares estimation can be used to estimate the parameters. Knowing B_1 from the coefficient in (5) and the assumption $B_1 + B_2 = 1$ we can specify all the

[10] K.J. Arrow, H.B. Chenery, B.S. Minhas, and R.M. Solow, "Capital-Labor Substitution and Economic Efficiency," *Review of Economics and Statistics* 43 (August 1961), pp. 225-250.

[11] R.R. Nelson, "The CES Production Function and Economic Growth Projections," mimeo., 1965.

values of the parameters in (3). One can of course determine the parameter values by using (3) in logs and thereby not restricting the values of B's but in keeping with the general assumption of linear homogeneity preference is made to (4) and the estimating equation (5).

IV

Our least squares estimation yielded the following result:

(6)

$$\log \left[\frac{Q}{L}\right]_{ij} = \log 1.17 + \underset{(8.74)}{.722} \log \left[\frac{K}{L}\right]_{ij} + U_{ij} \qquad \begin{matrix}(R^2 = .67)\\(\text{s.e.} = .24)\end{matrix}$$

where U_{ij} is the stochastic residual term with mean assumed be zero and constant variance. R^2 is the coefficient of determination and the bracketed term under the coefficient of log $\frac{K}{L}$ is the t-value; s.e. measures the standard error of the estimate. The regression coefficient of equation (6) is significant at both the one and five per cent levels. The F-value indicating the explanatory power of the variable is also significant.

Since our data are cross-section values of manufacturing industries, we did not test for autocorrelation. However a test was conducted to determine if the residuals of the least squares estimation had a constant variance under the assumption of the random error term U_{ij}. We partitioned our observations into two and conducted separate regression equations on both.[12] A ratio of the sums of squares of the least squares residuals was computed and an F-test conducted. This ratio is 5.11 which indicate under one and five per cent level of significance a rejection of the hypothesis that the residual term has a constant variance.

The significance of this result is crucial. This means that we had omitted an (or some) important explanatory variable(s)

---

[12] H. Theil, *Principles of Econometrics* (New York: Wiley, 1971), pp. 196-198, 214-216.

in our specification of the production function which is exerting an influence in the residual term. Or that we have pooled together non-homogeneous data.

Since our 2-country, 2-factor model is obviously between a developed and an underdeveloped economy, the next best thing to do was to add a dummy variable to our model, indicating a presumed level of development.

$$\text{Define} \quad Z^* = \begin{cases} 1 & \text{if U.S.} \\ 0 & \text{if Philippines} \end{cases}$$

Equation (2) now becomes

$$(7) \qquad \left[\frac{Q}{L}\right]_{ij} = f\left\{\left[\frac{K}{L}\right]_{ij}, Z^*_{ij}\right\}$$

and retaining the Cobb-Douglas function, our estimating equation (5) becomes

$$(8) \qquad \log\left[\frac{Q}{L}\right]_{ij} = \log \alpha + B_1 \log\left[\frac{K}{L}\right]_{ij} + B_2 Z^*_{ij}$$

Z^* here is not expressed in logs since Z^* takes on values only of either one or zero. In any case letting Z^* assume a value of 10 (the log of which is 1) for the U.S. and zero for the Philippines will not change the significance of the coefficient value.[13] The difficulty of interpreting the third term of equation (8) aside, the crucial question is whether differentiating countries by level of development improves the fit and whether the variable Z^* is significant.

Running a regression on equation (8) yielded the results:

$$(9) \qquad \log\left[\frac{Q}{L}\right]_{ij} = \log 2.03 + \underset{(4.65)}{.405} \log\left[\frac{K}{L}\right]_{ij}$$

$$+ \underset{(.30)}{.428}Z^*_{ij} + V_{ij} \qquad \underset{(\text{s.e.}=.18)}{(R^2=.81)}$$

[13] Transforming (8) to (4) would seem to be where a problem would lie.

where the definitions are similar to equation (6), V_{ij} is the new stochastic residual term, and Z* is the dummy variable indicating the level of development (or more appropriately as shown below, the degree of technology). More generally however we can consider Z* as a shift parameter (apart from lag α).[14] Notice that the coefficient of determination now increases to .81 from .67 in equation (6). All the regression coefficients are significant at the one and five per cent levels of significance. The F-value indicates that all variables have significant explanatory power. Conducting a t-test[15] whether there is any significant difference between the coefficient of $\frac{K}{L}$ in (9) and (6), we found that the coefficients are significantly different.

It is thus clear from (9) and (6) that the additional notion that productivity differences are also functions of levels of development is significant. Equation (7) therefore becomes a better model than (1) or (2). However resolving the issue by the addition of Z*, finding that this shift parameter is significant, and concluding that the two countries are not on the same production function but may be on different points on different production functions does not end our problem. It may just be starting. For the inclusion of the dummy variable does not eliminate the fact that as we found earlier in equation (6) our stochastic residual error term is heteroscedastic.[16] Thus while our fit improved, our main problem still remains. To seek ways of removing or reducing this heteroscedasticity involves another study. But rather than ending this paper with the knowledge that we know no better than what we know before the start, we end by briefly extending our statistical analysis in the direction that provides a link with some theory.

V

Nelson,[17] has suggested that one way of looking at international productivity differences is as a technological diffusion

[14] Recall α is a measure of total productivity growth index.

[15] Since the sample exceeded 30 we also conducted the usual normal test and the same results hold.

[16] Johnston, *Econometric Methods* (New York: McGraw-Hill, 1963), pp. 221-228.

[17] R.R. Nelson, "A 'Diffusion' Model . . . ", *op. cit.*

process from "leaders" to "followers." Over time therefore as diffusion increases rapidly in the "lag" countries and rates of increases of technological advancement are less than diffusion rates, we can expect "lag" countries to catch up with the "lead" countries — and at some particular time they may then be on similar production functions.

There are evidences which lend support to this thesis. Hufbauer[18], for example, finds that trade in the synthetic materials industry does not follow the Heckscher-Ohlin factor proportions theorem but yields to a technological-lag theorem. That is, he finds nations with abundant supply of the basic resources for synthetic materials are the net importers of these goods while the technologically advanced nations not possessing them are the net exporters. So with Keesing[19] and Posner.[20]

It must be admitted that we have not done justice to our assertions in II and III. Surely there have been more addition of variables to production function estimates like education, some measure of technical change, the proportion of skilled labor to total work force, and many others.[21] But our purpose is simply to remove the basic estimation problems in part IV.

Taking into consideration this diffusion process idea, we conducted regressions on the same variables as equation (8) but for different years — 1957 for the U.S. and 1962 for the Philippines. Our regressions resulted in:

$$(10) \qquad \log \left[\frac{Q}{L}\right]_{ij} = \log 1.27 + \underset{(6.89)}{.685} \log \left[\frac{K}{L}\right]_{ij} + W_{ij} \qquad \begin{matrix}(R^2=.67)\\ (s.e. = .29)\end{matrix}$$

[19] D.B. Keesing, "The Impact of Research and Development on U.S. Trade," *Journal of Political Economy* 75 (February, 1967), pp. 38

[20] M.V. Posner, "International Trade and Technical Change," *Oxford Economic Papers* 13 (October 1961), pp. 323-75.

[21] M. Brown (ed.), *The Theory and Empirical Analysis of Production* (New York: Columbia University Press, 1967).

$$(11) \quad \log \left[\frac{Q}{L}\right]_{ij} = \log 1.99 + \underset{(4.45)}{.410} \log \left[\frac{K}{L}\right]_{ij} + \underset{(5.27)}{.448} Z^* + e_{ij} \qquad \begin{matrix} (R^2 = .75) \\ (\text{s.e.} = .22) \end{matrix}$$

where the definitions are the same as before and W_{ij} and e_{ij} are the stochastic residual terms. All the coefficients of (10) and (11) are significant and all variables in both equations have significant explanatory power by their F-values. The coefficient of determination has reduced by as much as 11 per cent comparing (10) and (6) and as low as 6 per cent comparing (9) and (11). We conducted tests to determine whether there are significant differences in the coefficients of equations (10) and (11) and (6) and (9). At both one and five per cent significance levels there are no differences in the coefficients despite the 5-year lag, except for the dummy variable coefficient. Testing for the constant variance assumption for the residual terms we found that the hypothesis of heteroscedasticity is now rejected at the five per cent level of significance though not at one per cent. The ratio decreased measurably from 5.11 to 2.86.

Although we have not added new variables but simply lagging our data in accordance with the diffusion idea, the level of heteroscedasticity of our residual error term decreases with no significant differences in the regression coefficients.

This is no proof of course that the technological lag theorem holds. For one, we would have expected the coefficient of the dummy variable shift parameter to decrease with a lag; but it in fact increased tending to show that production function differences would widen, presuming this is evidence of the hypothesis of different production functions. The interpretation of Z^* however remains a thorny problem. For another, the coefficient of determination decreased from .81 to .75 which appear to indicate a trade-off between reducing heteroscedasticity with a lag and lower fit and increasing heteroscedasticity with a higher fit. Finally, the standard error of estimate also increases in (11).

But the experiments seem to give a direction. Perhaps with additional lag and more variables associated with diffusion added (like the fraction of trade and licensing agreements or the proportion of international firm subsidiaries in the total industry) we will be more comfortable in abandoning the existing neoclassical formulations.

FLORIAN A. ALBURO

University of Colorado, U.S.A.
and
University of the Philippines