AN ANALYSIS OF THE FACTORS RELATED TO THE VARIATION OF THE FOREIGN TRADE PROPORTION AMONG COUNTRIES

GLENN C. BLOMQUIST

THE FOREIGN TRADE proportion is considered to be the best summary measure of the extent of a nation's involvement in foreign trade.[1] Knowledge of the size of the foreign trade proportion is important because the proportion functions as a useful measure of the economic interaction between a country and the rest of the world. Despite the value of the foreign trade proportion, little has been accomplished toward explaining empirically its variation from country to country. The purpose of this paper is twofold: To examine unchallenged explanation offered by Kuznets and to quantitatively determine what factors are significant in the explanation of the variation in the size of the foreign trade proportion from country to country. In accomplishing the latter, the significance of the impact of institutional arrangements on the size of the foreign trade proportion is considered. Their influence on the foreign trade proportion has been suggested, but never before empirically supported.

I. BACKGROUND

Since the early twentieth century when Sombart proposed the "law of declining relative importance of foreign trade" the foreign trade proportion (FTP) has been thought to vary directly with the stage of economic development of a country in the early stages and inversely in the later stages. Sombart measured development by the degree of industrialization. Despite some sharp criticisms by Lipsey and others this FTP-Development hypothesis remained through the years and has been restated and supplemented by Kindleberger. What might be called the Sombart-Kindleberger hypothesis is: The FTP is low in primitive less-developed countries and rises with development until a country becomes a developed country. The FTP declines with further development of a developed country [13, pp. 368–76] [8, pp. 9–11, 18–32].

The Sombart-Kindleberger hypothesis was not formulated without support and criticism. The dependence of the FTP size on geographic size, population, per capita product, industrialization, etc., was hypothesized and argued in a priori fashion. Various empirical studies were made by Deutsch and Eckstein, Chenery, and Lipsey and others with the result being that the Sombart-Kindleberger hypothesis remained barely intact [3, pp. 267–77] [1, pp. 639, 634–45] [10, p. 37]. However, it was the only recognized explanation of the size of the FTP.

Recently the Sombart-Kindleberger hypothesis has been replaced by one proposed by Kuznets. He purports that the FTP size is determined primarily by the economic size of a country *and* its stage of economic development. Kuznets showed that the effect of development, if it could be isolated from other effects, was always to increase the FTP regardless of the stage of economic development. His cross-sectional study of sixty-two countries for the year 1958 supported his hypothesis [9, pp. 1–107]. Peter Lloyd's work with sixty countries for 1964 produced results consistent with Kuznets [11, pp. 23–28]. Up to this time there have been no challenges to the FTP work of either Kuznets or Lloyd. Their views could be summarized in a Kuznets-Lloyd hypothesis which would state: The FTP varies inversely with the economic size of a country and once this effect is considered the FTP varies directly with advancement in economic development regardless of whether or not the country is developed or less developed. This hypothesis is presently the best explanation of the FTP size.

II. OBJECTIVES

There are four specific objectives to be accomplished in this paper. The first is to check Kuznets' method of isolating the influence of economic development, i.e., method estimating the FTP based on economic size alone. Economic development is measured by per capita gross national product and economic size by gross national product. Kuznets produced foreign trade proportion estimates for each of the sixty-two countries. They were FTP size estimated on the basis of gross national product only. These estimates were then subtracted from each country's actual FTP. A FTP deviate unaffected by economic size was the remainder. This FTP was found by Kuznets to have significant direct relation with per capita product [9, pp. 15–25]. Thus it is extremely important that the FTP deviates have no influence of economic size reflected in them. The first objective is: (1) To check Kuznets' FTP deviates to see whether or not he successfully removed the influence of size. If his arithmetic method removed the effect, simple regression analysis will show that it did. The other objectives are: (2) To use the FTP deviates from an alternative method to provide a further test of the Kuznets-Lloyd hypothesis. A regression equation should provide FTP deviates equally as well as Kuznets' method. (3) To determine which economic variables among those suggested in the literature add significantly to the explanation of the FTP. These will be included in a prediction equation for the FTP and the explanatory power of the generalization about the size of the FTP will be assessed. Among the variables considered are two which reflect institutional trading arrangements. Variables of this nature have not been used quantitatively in previous empirical studies.

[1] The foreign trade proportion unless specifically defined differently is the ratio of the sum of exoprts plus imports to gross national product.

The Developing Economies 8 (June 1970): 180-197.

III. A CHECK ON KUZNETS' ARITHMETIC METHOD

The first objective is to check Kuznets' method of estimating the foreign trade proportion based on gross national product alone. To accomplish the check a simple regression of the FTP deviate as estimated by Kuznets as a function of Kuznets' data for 1958 gross national product is computed. The FTP deviate is defined as the true FTP for a country minus the FTP estimate based on size only. The results of this regression are an indication of the degree of Kuznets' success in eliminating the size factor from the FTP.

The following regression results were obtained:[2]

$$\text{(Kuznets' FTP deviate)} = \underset{(1.91)}{12.54025} - \underset{(1.00)}{0.11166}\,(\text{GNP}) \qquad \underset{(1.00,\ 500)}{R^2 = 0.016} \qquad (1)$$

where GNP is gross national product. The regression and regression coefficients are statistically insignificant at the 95 per cent confidence level ($\alpha = 0.05$), because the t values are less than the critical t value of 2.00 and the F value is less than the critical F value of 4.00. These results are evidence that support the contention that Kuznets' arithmetic method of accounting for the effect of economic size as measured by gross national product was successful.

IV. AN ALTERNATIVE TO KUZNETS' METHOD

Kuznets claims there is a positive, direct correlation between the foreign trade proportion and the stage of economic development as measured by per capita gross national product. His claim can be further tested by using an alternative method of obtaining the deviations from the "normal" FTP, where "normal" FTP is based upon economic size (gross national product) alone. The FTP deviates can be obtained through the application of simple regression analysis using Kuznets' 1958 data for sixty-two countries. The FTP is regressed as a function of only gross national product. The residuals from this regression are used as the FTP deviates, i.e., the estimates of the FTP considering only economic size.

Care was taken in order to obtain the best linear estimates of the FTP using only gross national product. The first equation attempted was the FTP regressed as a function of gross national product in levels, i.e., percentages and millions of dollars.

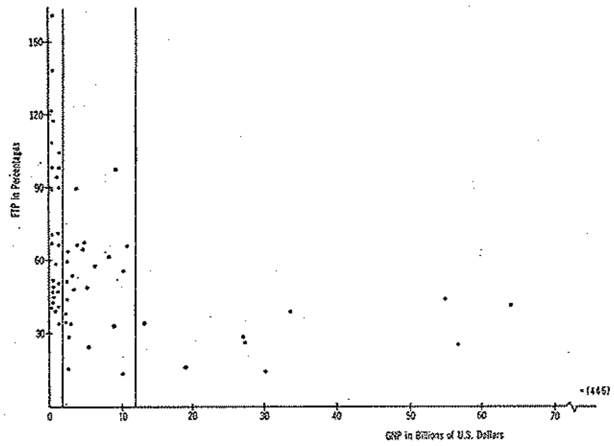
$$\text{FTP} = \underset{(14.61)}{59.06508} - \underset{(2.14)}{0.14697}\,(\text{GNP}) \qquad \underset{(4.58,\ 54.1)}{R^2 = 0.071} \qquad (2)$$

This regression was barely significant at the 95 per cent confidence level, as indicated by a calculated F statistic being only slightly greater than the critical F value, 4.00. Moreover, the calculated t statistic for the regression coefficient was also only slightly greater than the critical t value of 2.00. Since unsatisfactory results were obtained, the data plot of FTP and the independent variable, gross national product, was studied further in an attempt to find convenient groups into which the data could be divided. Dividing the data was considered for two reasons which are related to the assumptions of the linear regression model. Since Woytinsky, and Deutsch, Bliss, and Eckstein claimed that the FTP increased, reached a peak, and then decreased, a curvilinear relationship between the FTP and the stage of development had to be considered [2, p. 355] [17, pp. 65–67]. Application of linear regression analysis requires the assumption that the FTP-Development relationship is linear in the parameters (see [4, pp. 6, 81–85]). Data splitting would hopefully isolate the linear sections of the alleged curve and provide better linear estimates of the FTP than can be obtained from a straight estimating line through the entire group of data if this were in fact the case. This is the first reason.

In addition, it was obvious from the data plot that the variance of the FTP was not constant over the entire range of economic sizes. Application of regression analysis requires the assumption that the variance is constant at all observations in the data, i.e., homoscedasticity (see [7, pp. 207–11]). Data splitting would hopefully provide ranges which had constant variances, even though the variances were not necessarily the same. Segregation of ranges with constant variances would yield FTP estimates with a smaller error of estimate in comparison with

2 The values in parentheses beneath the regression coefficients are the respective computed t values. The number on the left in parentheses beneath the squared correlation coefficient is the computed F value. The value on the right is the standard error of the regression as a percentage of the dependent variable. All tests for significance are made at the 95 per cent confidence level with the appropriate degrees of freedom. These conventions are adhered to throughout this paper.

Fig. 1. FTP-Size Relationship for Kuznets' 1958 Data for Sixty-two Countries



Source: [4, Appendix Tables 1 and 3].

estimates from the entire data set. This is the second reason. No clear natural dividing points were observed upon plotting the data.

It did appear promising to separate the countries into three groups consisting of: ten countries with gross national products less than 2.20 billion 1958 dollars —Group A; twenty-one countries with gross national products between 2.20 and 12 billion—Group B; and thirty-one countries with gross national products greater than 12.00 billion—Group C. The divisions made it possible to regress the FTP as a function of gross national product for the three groups. In each of the three regressions the calculated values of the F and t statistics were below those required to be considered significant at the 95 per cent confidence level.

Group A:

$$\text{FTP} = \underset{(7.53)}{31.63181} - \underset{(1.55)}{0.04441}\,(\text{GNP}) \qquad \underset{(2.39,\ 40.0)}{R^2 = 0.230} \tag{3}$$

where 2.306 and 5.32 are the critical t and F values, respectively, and GNP is gross national product.

Group B:

$$\text{FTP} = \underset{(4.00)}{42.43913} + \underset{(0.74)}{1.26470}\,(\text{GNP}) \qquad \underset{(0.55,\ 45.2)}{R^2 = 0.028} \tag{4}$$

where 2.09 and 4.38 are the critical t and F values, respectively.

Group C:

$$\text{FTP} = \underset{(8.63)}{87.07827} - \underset{(1.93)}{18.17643}\,(\text{GNP}) \qquad \underset{(3.73,\ 45.7)}{R^2 = 0.114} \tag{5}$$

where 2.05 and 4.18 are the critical t and F values, respectively. Dividing the data was unrewarding. Yet, it appeared that the data might not have a constant variance. Given that there is heteroscedasticity it has been shown in econometric literature that more efficient estimating equations are achieved if the variance at different values of gross national product is estimated. Thus one could cope with the heteroscedasticity by assuming or estimating the variances of gross national products. Since the variance appears to get smaller as the gross national product increases one might consider this relationship to be the case: $E(U_i^2) = \sigma^2\, 1/(\text{GNP})^2$ where U_i is the residual, GNP is gross national product and σ^2 is a constant. Then the regression equation would take the form:

$$\text{FTP}_i(\text{GNP}_i) = \beta_1(\text{GNP}_i) + \beta_2(\text{GNP})_i^2 \tag{6}$$

(See J. Johnston for further explanation of this technique [7, pp. 210–11].) Indeed this type of equation would most likely produce better estimates of the FTP based only on size because the estimates would have a smaller variance. However, it was judged that the FTP variance was not statistically different at various ranges of gross national product. It would not alter to a significant degree the results obtained from a linear regression where the residual variances were considered constant, i.e., where there was indeed homoscedasticity.

Since little advantage was gained by splitting the data, the data as a whole was considered again. Four regressions were considered for estimators of the FTP deviates. The four equations plus equation (2) formed the group from which

the best estimating equation was selected. The first of the four equations tried was the FTP regressed as a function of the logarithm of gross national product. The second was the logarithm of the FTP as a function of gross national product. The third was the logarithm of the FTP regressed as a function of the logarithm of gross national product. The fourth of the equations was the FTP regressed as a function of gross national product, gross national product squared, and gross national product cubed. The regression tried previously, equation (2), was the FTP as a function of gross national product. Out of this group of five prediction equations the one measuring both FTP and gross national product in logarithms was chosen to estimate the FTP based on size (gross national product) alone. Statistically this question was the most significant of the five equations on the basis of F and t statistic values.

$$\log \text{FTP} = \underset{(60.94)}{4.06640} - \underset{(5.91)}{0.19991}\,(\log \text{GNP}) \qquad \underset{(34.93,\ 12.1)}{R^2 = 0.368} \tag{7}$$

where GNP is gross national product. In addition, the calculated F and t values are well above the minimum critical values of 4.00 and 2.00 for the F and t statistics respectively. The regression of equation (7) was computed and predicted values of the FTP were obtained for each of the sixty-two countries. FTP deviates similar to Kuznets FTP deviates were calculated by subtracting the logarithm of the FTP predicted by the logarithm of gross national product, from the logarithm of the actual FTP for each country. These FTP deviates were then regressed as a function of the logarithm of per capita gross national product:

$$\log \text{FTP deviate} = -\underset{(3.76)}{1.29924} + \underset{(3.80)}{0.21910}\,(\log \text{PCP}) \qquad \underset{(14.46,\ \text{undefined})^{a}}{R^2 = 0.194} \tag{8}$$

where PCP is per capita gross national product.

The results of this regression method can be compared to the results of Kuznets' arithmetic method. The significance and sign of the regression coefficient of per capita gross national product are of primary interest. Equation (8) does produce a significant regression, since the calculated F value, 14.46, is greater than the critical value of 4.00. The positive regression coefficient of per capita gross national product is significant, because the calculated t, 3.80, is greater than the critical value of 2.00. The results of this alternative method strengthen Kuznets' claim that a significant, positive relationship between a typical country's FTP and its stage of economic development does exist.

V. EMPIRICAL ANALYSIS PERTINENT TO THE EXPLANATION OF THE SIZE OF THE FOREIGN TRADE PROPORTION: THE WORK OF LLOYD

After the exploration of the relationship between the foreign trade proportion and the stage of economic development which focused upon the work of Kuznets, it

[a] The error of estimate as a percentage of the mean of the dependent is undefined because the mean of FTP deviates is 0.0000.

was decided to examine the FTP's relationship with numerous other economic variables besides gross national product and per capita gross national product. The objective was to explain as much of the variation in the size of the FTP as possible. For Kuznets' data, the economic size of the country as measured by gross national product and the stage of economic development as measured by per capita product explain only part of the variation in the size of the FTP.

$$\log FTP = \underset{(7.06)}{2.55323} - \underset{(7.85)}{0.25718} (\log GNP) + \underset{(4.24)}{0.26387} (\log PCP) \qquad \begin{matrix} R^2 = 0.516 \\ (31.42,\ 10.6) \end{matrix} \qquad (9)$$

In order to make the squared multiple correlation coefficient in equation (9) comparable to one expressed in levels so that it can be more readily interpreted, the FTP is regressed as a function of the antilogarithm of the predicted FTP of equation (9).

$$FTP = \underset{(1.21)}{9.01325} + \underset{(7.01)}{0.90027} (\widehat{FTP}) \qquad \begin{matrix} R^2 = 0.450 \\ (49.16,\ 45.0) \end{matrix} \qquad (10)$$

where $\widehat{FTP}$ is the antilogarithm of the predicted FTP of equation (9). On the average a bit less than one-half of the FTP size can be explained by economic size and stage of economic development as they are here defined.

Suggestions of important, relevant variables for explaining the FTP were found in the literature. Six independent economic variables were selected for multiple regression analysis. Economic size as measured by gross national product or gross domestic product and population was considered because Kuznets, Haberler, and others hypothesized that the larger countries have smaller FTP's. Geographic size as measured by land area was included since Lewis and others claimed that countries with little geographic area supposedly have larger FTP's since their resources are likely to be skewed. Per capita product (either per capita gross national product or per capita gross domestic product) was chosen, as Kuznets' and Lloyd's work showed the FTP increased as the stage of development was advanced. The percentage of total aggregate produce produced by industry was included because Sombart, and Deutsch and Eckstein maintain that as industrialization occurs the FTP rises at first and then declines. Fixed capital formation as a percentage of gross aggregate product was included because Ellsworth and Kindleberger maintain that a relatively high rate of fixed capital formation leads to a high FTP because allegedly much of the capital is imported [8, pp. 183–84] [5, p. 854].

Data used for the multiple regression was based on that of Lloyd for sixty countries for the year, 1964 [11, pp. 128–30]. A small part of the data is for 1963. The Lloyd data is similar to Kuznets' data. There are thirteen countries included in Kuznets' sixty-two countries that are omitted from Lloyd's list. They are India, Brazil, Sweden, Switzerland, New Zealand, Nigeria, Cuba, Algeria, Rhodesia, Nyasaland, Ghana, Iceland, and British Guiana. There are eleven countries included in Lloyd's sixty which are omitted from Kuznets' list: Sudan, Togo, Tunisia, Bolivia, Paraguay, Uruguay, Cambodia, Iran, Iraq, Jordan, and South Viet-Nam. The differences in the samples of countries do not appear to be important. Gross domestic product is used as the measure of aggregate economic activity in the Lloyd data. Kuznets used gross national product. The difference between the two products is that the former includes the net flow of factor incomes across boundaries while the latter does not. This will not influence the results to any significant degree because the difference between gross national and gross domestic product is a small percentage of either one of the products. By the same reasoning results will be acceptable if per capita gross domestic product is used in place of per capita gross national product. Percentage of gross domestic product produced by industry uses industry as "total industry" as defined by the United Nations [16]. Population is for the middle of 1964. Land area in hectares is the measure used for geographic area.[4]

The results that support the Kuznets-Lloyd hypothesis that are achieved through multiple regression analysis of the 1964 Lloyd data are:

$$\begin{aligned} FTP = {} & 0.295147 - \underset{(2.22)}{0.000008} (GNP) + \underset{(2.64)}{0.000251} (PCP) + \underset{(0.01)}{0.000042} (TIP) \\ & + \underset{(0.96)}{0.000185} (POP) - \underset{(1.59)}{0.000035} (AREA) + \underset{(1.47)}{0.010431} (FKF) \end{aligned} \qquad (11)$$

(2.22) (2.66) (2.64) (0.01)

$R^2 = 0.267$

where GNP is gross national product, TIP is percentage of gross domestic product produced by total industry, POP is population, AREA is land area, and FKF is percentage of gross domestic product expended on capital formation [11, pp. 23–28]. The significant, negative regression coefficient for gross domestic product and significant, positive regression coefficient for per capita product are indications of support for expecting a smaller FTP the larger the economic size of a country, while expecting a larger FTP, the more developed is a country once the size effect is considered.

For some unknown reason one does not get the same regression results when Lloyd's data is used, the same variables are used, the variables are regressed in the same order, and the variables are measured in levels. Equation (12) should be the same as equation (11).

$$\begin{aligned} FTP = {} & \underset{(2.25)}{25.23242} + \underset{(3.83)}{0.00041} (GDP) - \underset{(0.95)}{0.00817} (PCP) + \underset{(0.54)}{0.17350} (TIP) \\ & - \underset{(5.26)}{1.25947} (POP) - \underset{(1.78)}{0.03266} (AREA) + \underset{(3.70)}{2.41362} (FKF) \end{aligned} \qquad (12)$$

$R^2 = 0.480$
(8.15, 40.3)

These results disaffirm the Kuznets-Lloyd hypothesis in that no significant regression coefficient for per capita product is obtained and there is a significant positive coefficient for gross domestic product. The only support for the Kuznets-Lloyd hypothesis is that a significant, negative regression coefficient for population is found. These results cast doubt on the Kuznets-Lloyd hypothesis.

[4] One hectare is equal to 2.47 acres, approximately [16, Table 3, pp. 457–65].

Because the differences between Lloyd's results and equation (12) were significantly different, Lloyd's data was checked for errors or misprints. Reference to the *Production Yearbook for 1965* [6], *United Nations Monthly Bulletin of Statistics* [16], and *United Nations Yearbook of National Accounts Statistics 1965* [17] showed that there were no major discrepancies in Lloyd's data and that it was in fact essentially correct. The data was reproduced. Two observations in Lloyd's data were doubted. The percentage gross domestic product produced by industry for Uruguay should have been reported as 24, not 46. The fixed capital formation as a percentage of gross domestic product for the United States should have been reported as 17, not 15.[5] Thus, the differences between equation (11) and equation (12) are due to differences in the regression analysis. It was noted that the mean of GDP is 20, 109, not 11, 176. The OSU/ECON Regression Program written by Jon Cunnyngham was used to produce the results appearing in equation (12).[6]

Before an investigation of the effect of institutional arrangements on the Kuznets-Lloyd hypothesis an attempt was made to improve the explanation of the foreign trade proportion size by checking the sensitivity of the Lloyd-type equation (12) to mathematical form. Kuznets' success in the use of logarithms in his analysis of variance suggested that expressing the variables in logarithms for multiple regression analysis would be a helpful statistical technique. By allowing the FTP and six independent variables to be expressed in logarithms the explanatory power of the regression equation can be improved. This alteration alone provides a better explanation of the FTP than does the regression analysis which measured the variables in levels, i.e., hectares, dollars, rather than rates of change. When equation (12) was expressed in logarithms the following resulted:

$$\begin{aligned} \log \text{FTP} = {} & 2.85641 - 0.19759 (\log \text{GDP}) + 0.09561 (\log \text{PCP}) \\ & (6.81) \qquad (4.29) \qquad\qquad (1.16) \\ & + 0.13184 (\log \text{TIP}) - 0.06022 (\log \text{AREA}) + 0.60617 (\log \text{FKF}) \\ & \quad (1.10) \qquad\qquad (2.08) \qquad\qquad (3.37) \\ R^2 = {} & 0.625 \\ & (17.99,\ 8.7) \end{aligned} \tag{13}$$

The variable for population was colinear with that for gross domestic product to the degree that the former's diagonal elements dropped below the tolerance level for matrix computation. Population was statistically unimportant in equation (13) and was deleted from the regression. In order to compare the two squared multiple correlation coefficients and standard errors of equations (12) and (13) the antilogarithms of the estimated FTP were found and the FTP regressed as a function of the antilogarithms as was done in equation (10). This procedure provided statistics for (13) comparable to those for (12) in levels.

$$\text{FTP} = 5.26494 + 0.93768\,(\widehat{\text{FTP}}') \qquad R^2 = 0.682 \tag{14}$$
$$(1.10) \qquad (11.15) \qquad\qquad (124.25,\ 31.3)$$

where $\widehat{\text{FTP}}'$ is the antilogarithm of the predicted FTP values from the equation expressing all values in logarithms, (13). The outcome was that the regression in logarithms possessed larger t and F statistic values (F, 124.25—greater than 8.15), a larger R^2 (0.682—greater than 0.480), and with a smaller standard error relative to the mean of the FTP (31.3—less than 40.3). Thus the search for further explanation of the FTP beyond that provided by gross aggregate product and per capita product and the investigation of the effect of institutional arrangements upon the FTP-Development relationship was performed using the logarithmic mathematical form of a multiple regression equation.

In order to be certain that approximately the same amount of FTP variation can be explained by gross aggregate product and per capita product in Lloyd's data as with Kuznets' data, equation (13) was again regressed using only the logarithms of these two as independent variables. When these results were made comparable to a regression in levels, it became apparent from equations (10) and (15) that in both data sets approximately 50 per cent of the variation in the FTP is explained by economic size and stage of economic development. The explanation had to be roughly the same in order to substitute the Lloyd data for the Kuznets data for the analysis of the FTP where more than these two variables were considered. A bit more than half was explained in Lloyd's data for 1964:

$$\text{FTP} = 1.13802 + 1.04037\,(\widehat{\text{FTP}}'') \qquad R^2 = 0.517 \tag{15}$$
$$(0.16) \qquad (7.87) \qquad\qquad (62.01,\ 39.4)$$

where $\widehat{\text{FTP}}''$ is the antilogarithm of the predicted FTP values of the regression of Lloyd's FTP's as a function of gross domestic product and per capita product in logarithms.

Thus, the Kuznets-Lloyd hypothesis explained only one-half of the various FTP sizes as evidenced by the squared multiple correlation coefficients of equations (10) and (15).

VI. ACCOMMODATING INSTITUTIONAL ARRANGEMENTS IN A QUANTITATIVE EXPLANATION OF THE FOREIGN TRADE PROPORTION

Two provisions for institutional arrangements were made. A dummy variable for economic integration was created. The reason for establishing it was that economic integration produced higher FTP's in the participant countries than in the countries which were non-participants according to Lloyd. A dummy variable for satellite countries was established since Kuznets and Woytinsky suggest that satellites have higher FTP's than countries which are sovereign nations with no historic political foreign trade ties [9, p. 31] [17, p. 65].

The dummy variable for economic integration was defined in the following

[5] [6, Table 1, pp. 3–8] for AREA; [15, Table 1, pp. 1–4] for POP; [16, Table 2, pp. 448–55] for FTP and FKF; [16, Table 3, pp. 457–65] for TIP; and [16, Table 9B, pp. 449–502] for GDP and PCP.

[6] Correspondence with Lloyd revealed that his regression results were very likely incorrect. Since he is in Australia and his material is in Michigan, he could only speculate that some observations in his data were mispunched.

manner. If a country was a member of either the European Economic Community or the European Free Trade Association in 1964, the country was considered to be significantly influenced by the effects of economic integration. Any other country was not deemed to be actively and effectively participating in economic integration. As of 1964 the Latin American Free Trade Association and the Central American Common Market were not effective enough to be considered. Ten of the sixty countries were in one of the two effective organizations and were designated as active participants in economic integration. Portugal, Norway, Denmark, Austria, and the United Kingdom were participants in the European Free Trade Association. The Netherlands, Belgium, Luxemburg, West Germany, Italy, and France were participants in the European Economic Community.

The dummy variable for satellite countries was defined with the aid of *The Statesman's Yearbook 1964–65* [14] and the *Rand-McNally Commercial Atlas and Marketing Guide, 1968* [12]. Each of the sixty country's political status was investigated along with its primary foreign trade partners and its geographical and economic size. Each country must have a strong enough historical, political relationship with a center country that, say over 50 per cent of a satellite's foreign trade is with it or that satellite trades at least as much with that center country as any other country. The satellite must have been a dependent territory in 1964, have received its independence within the ten years prior to 1964, or have some other peculiarly strong, relationship with a center country. In addition the satellite must have been small in area, say less than fifteen thousand hectares and have had a small population of, say less than 10 million people in 1964. These are indeed small countries in that the mean area for the sixty countries was eighty-five thousand hectares and the mean population, 17 million people. Ten countries qualified as satellite countries. Mauritius, Barbados, Jamaica, Trinidad and Tobago, Cyprus, Ireland, and Malta qualified because of their relationship with the United Kingdom. Puerto Rico and Panama were included because they were associated with the United States. Togo qualified because of its association with France.

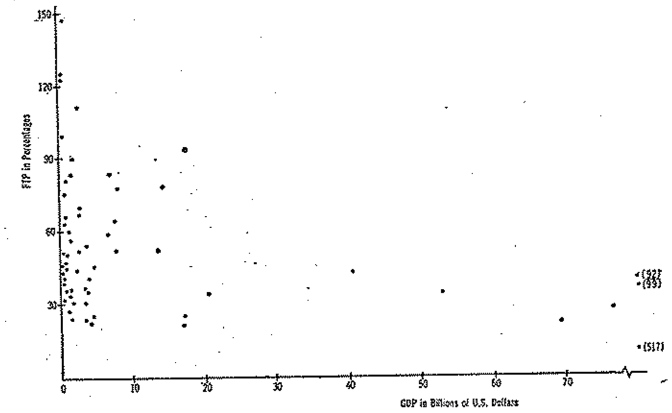
Before the multiple regression analysis was performed the FTP was plotted as a function of gross aggregate product for data similar to Kuznets' data, the data gathered by Lloyd.

The countries were grouped according to where they were bunched and the economic characteristics of each country within the group were compared. For example, it was noted that the five highest FTP's belonged to countries which were satellites of larger countries and their land areas and populations were extremely small relative to the other countries. These and other comparisons were fruitful enough to warrant further study using the multiple regression analysis.

VII. THE EXPLANATION OF THE FOREIGN TRADE PROPORTION PROVIDED BY THE CONSIDERATION OF INSTITUTIONAL ARRANGEMENTS AND THE EFFECT OF THIS CONSIDERATION ON THE KUZNETS-LLOYD HYPOTHESIS

The third objective of this paper is to choose the best statistical prediction equation which explains the most FTP variation and investigate the effect of the consideration of institutional arrangements on the Kuznets-Lloyd hypothesis.

Fig. 2. FTP-Size Relationship for Lloyd's 1964 Data for Sixty Countries



Source: [11, Appendix 1].

In order to determine how much more explanation could be rendered by considering economic integration and satellite countries, the dummy variables for these intercountry bonds were incorporated into a regression equation expressed in logarithms which included gross domestic product, per capita product, fixed capital formation percentage, geographic area, and total industrial product percentage. This did improve the explanation only a small amount over the regression without the dummy variables, equation (13):

$$\begin{aligned} \log FTP = {} & 3.09440 - 0.22208\,(\log GDP) + 0.67689\,(\log FKF) \\ & \quad (7.62) \qquad (4.73) \qquad\qquad (4.02) \\ & + 0.46288\,(EI) + 0.32157\,(SATEL) + 0.08053\,(\log TIP) \qquad (16) \\ & \quad (3.16) \qquad\qquad (1.99) \qquad\qquad (0.72) \\ & + 0.03319\,(\log PCP) - 0.00566\,(\log AREA) \qquad R^2 = 0.690 \\ & \quad (0.41) \qquad\qquad (0.18) \qquad\qquad (16.52,\ 8.0) \end{aligned}$$

where GDP is gross domestic product, FKF is fixed capital formation expenditure as a percentage of GDP, EI is economic integration, SATEL is satellite country, TIP is total industrial product as a percentage of GDP, PCP is per capita gross domestic product and AREA is land area. The standard error as a percentage of the logarithm of the FTP was slightly smaller (8.0 is less than 8.7) and the squared multiple correlation coefficient was slightly greater (0.690 is greater than 0.625), but the F value was a bit smaller (16.52 is less than 17.99). Superficially it appears as though the dummy variables add little to explanation of the FTP.

Intercorrelation and insignificance among the independent variables must be considered.

Before the regression is transformed to obtain results comparable to the regression expressed in levels a better prediction equation is chosen according to statistical criteria than that which contains all the independent variables. The total industrial percentage, per capita product, and geographical area variables were deleted because their respective values of 0.72, 0.41, and 0.18 were far below the value of 2.00 required for significance. Population was not included in the prediction equation because it was highly colinear with gross domestic product and its diagonal element dropped below the tolerance level in computation. This meant that the matrix was nearly singular and that population and gross domestic product could not be considered independent economic variables. This was not surprising in that both are measures of economic size. Gross domestic product was chosen over population as the better measure of economic size because alternative equations which utilize gross domestic product as the measure for economic size, were more significant, did not result in large negative indirect effects for any variable, and possessed a larger squared multiple correlation coefficient.[7] Per capita product displayed a net negative indirect effect when it appeared in a regression where population was used as the measure of economic size. The net effect was −1.18 and the indirect to direct effect ratio, −1.07. Different multiple regression equations containing gross domestic product and population were compared because whenever these two variables appeared in the same equation one was insignificant due to the large amount of colinearity between them. Their simple correlation coefficient was 0.833. Gross aggregate product is generally claimed to be the better measure of economic size. Thus the comparison of each "best" equation containing population to one containing gross domestic product was justified. When the equations containing gross domestic product were found superior to those containing population and the former concentrated upon, the large negative indirect effect of per capita product disappeared. However, the regression coefficient for per capita product was insignificant.

Eliminating those variables which did not meet the statistical standards left the following prediction equation:

$$\log FTP = \underset{(8.32)}{3.16697} - \underset{(7.06)}{0.20379}\,(\log GDP) + \underset{(5.21)}{0.74575}\,(\log FKF) + \underset{(4.22)}{0.51289}\,(EI) + \underset{(3.20)}{0.38891}\,(SATEL) \qquad \underset{(29.72,\ 7.8)}{R^2 = 0.684} \tag{17}$$

Although the squared multiple correlation coefficient for equation (17) is slightly lower than for equation (16) (0.684 is less than 0.690) the t and F values are greater (F, 29.72 is greater than 16.52) and the relative standard error is smaller (7.8 is less than 8.0). A small amount of explanatory power is foregone for an equation which has less multicolinearity, is more significant, and more efficient.

[7] Jon Cunnyngham, course in applied regression analysis, Ohio State University, Columbus, Ohio, April–May, 1968. See Appendix for a discussion of direct and indirect effects.

Because the Kuznets-Lloyd hypothesis was found to be significant in previous explanation of the FTP variations above, PCP was reentered into equation (17) to check its importance. The regression coefficient for development (PCP) was insignificant with a t value of 0.72. The additional explanation was less than 1 per cent of the unexplained variation. The four variables already in the equation remained quite significant. When the prediction equation, (17), is made comparable to those regressed in levels, it can be seen that the incorporation of the dummy variables into the prediction equation has in fact increased the amount of FTP variation that can be explained from country to country. The transformation of equation (17) produces:

$$FTP = \underset{(0.63)}{2.89391} + \underset{(12.10)}{0.98000}\,(\widehat{FTP}''') \qquad \underset{(146.43,\ 29.4)}{R^2 = 0.716} \tag{18}$$

where $\widehat{FTP}'''$ is the antilogarithm of the FTP predicted by equation (17).

Considerably more explanation of the foreign trade proportion has been accomplished than was accomplished with regressions in levels without the economic integration and satellite dummy variables. If the prediction equation is compared to Lloyd's results, equation (11), then the unexplained variation in the FTP has been reduced by 61 per cent. If the prediction equation is compared to the regression using Lloyd's data, equation (12), then the unexplained variation is reduced 45 per cent, the relative standard error of regression diminishes (40.3 to 29.4) and the regression is more significant (F, 146.43 is greater than 8.15). On the average, over 70 per cent of the FTP size can be accounted for by economic size, relative size of fixed caiptal formation expenditure, economic integration and satellite relationships.

Examination of the residuals from equation (17) revealed the countries with the ten largest deviates. They were South Africa, Tanzania, Canada, Venezuela, Iraq, Malaya, and The Netherlands with actual FTP's greater than predicted FTP's and Togo, Paraguay, and Japan with actual FTP's less than predicted FTP's. It is interesting to note that South Africa, Venezuela, and Iraq are exporters of some very valuable commodity, gold and oil, respectively. Malaya almost qualified for the satellite dummy variable, but was a bit large in geographic area. This suggests why four of the underpredictions occurred. The overprediction of Togo occurred probably because it is not as economically dependent as the other satellite countries. There are no apparent reasons why Tanzania, Canada, Netherlands, and Japan were not predicted better.

The quantitative consideration of institutional arrangements modifies the Kuznets-Lloyd hypothesis in that the positive correlation of the FTP with the stage of econòmic development becomes insignificant. The inverse relationship between the FTP and economic size is revealed once again as an important determinant of the FTP size. This is consistent with the Kuznets-Lloyd hypothesis. However, the hypothesis must be valued with less worth in the sense of being a dominant relationship which *determines the FTP size. The size of fixed capital expenditures relative to gross aggregate product of a country and the country's involvement in economic integration or with a large center country overshadow* the FTP-Development relationship.

VIII. CONCLUSIONS

1. The Kuznets' arithmetic method of estimating a foreign trade proportion that has the effect of economic size removed was successful. Kuznets' FTP deviates showed no linear association with gross national product.
2. The regression method of estimating FTP deviates as an alternative to Kuznets' arithmetic method, further affirmed the Kuznets-Lloyd hypothesis. This hypothesis states that the FTP size depends primarily on economic size with which it varies inversely and the stage of economic development with which it varies directly.
3. The Kuznets-Lloyd hypothesis accounts for approximately one-half of the variation from country to country in the size of the FTP.
4. Over 70 per cent of the variation in FTP size can be explained by consideration of economic size (GNP), relative fixed capital formation expenditure (FKF), economic integration (EI), and satellite dependency (SATEL). A country's FTP size becomes smaller: the larger the economic size, the smaller the relative amount spent on fixed capital formation, and the less the country relies upon institutional trade arrangements.
5. The importance of the Kuznets-Lloyd hypothesis is overshadowed by that of the general prediction equation. The hypothesis is neither disproven nor disaffirmed by the variables included in the general prediction equation. Rather, it was shown that the FTP-Development relationship coupled with the FTP-Economic size relationship is not so good at accounting for variation in FTP size as was the general prediction equation with its four variables.
6. Consideration of participation in economic integration and satellite dependency does offer a significant amount of explanation of FTP size. Institutional arrangements are important, relevant determinants of the size of the foreign trade proportion.

APPENDIX

DIRECT AND INDIRECT EFFECTS

The following notation is used in this Appendix:

x_i is the i-th variable.

$\beta_{ij.z}$ is the true regression coefficient for X_j, given X_z in a three variable multiple regression, where X_i is the dependent variable.

"Beta weight" of a regression coefficient is the coefficient divided by its standard error.

r_{ij} is the simple correlation coefficient for variables, X_i and X_j.

r is the correlation matrix.

R^2 is the multiple correlation coefficient.

D_{ij} is the determinant of the correlation matrix for variables, X_i and X_j.

The use of direct and indirect effects is a method for dissecting the squared multiple correlation coefficient in a manner so as to render obvious the independent influence of each "independent" variable upon the dependent variable. It is a method in which the squared multiple correlation coefficient is decomposed into direct and indirect effects. The impetus for analysis of direct and indirect effects is the elimination of variables from a prediction equation which are very colinear with other "independent" variables and which claim their significance in the regression equation because of their colinearity with other truly significant variables. The direct effect of an independent variable upon the dependent variable is the effect upon explanation of the dependent variable not dependent upon any colinearity with another independent variable. The direct effect of variable X_2 in a three variable model where X_1 is the dependent variable and X_2 and X_3 are the independent variables, is the product of the beta weight of $\beta_{12.3}$ times the beta weight of $\beta_{12.3}$ times the simple correlation coefficient of X_2 with X_2. $\beta_{12.3}$ is the true regression coefficient of X_2 in a multiple regression with X_2 and X_3 as independent variables. The beta weight for $\beta_{12.3}$ is the quotient of the regression coefficient $\beta_{12.3}$ divided by the standard error of this regression coefficient. The indirect effect is the explanation of the dependent variable by the particular independent variable that is due only to this independent variable's multicolinearity with other independent variables. The indirect effect between variables X_2 and X_3 is the product of beta weight of $\beta_{12.3}$ times the beta weight of $\beta_{13.2}$ times the simple correlation of X_2 and X_3.

Consideration of the three variable equation illustrates the computation of the direct and indirect effects. Variable X_1 is the dependent variable to be explained. Variables X_2 and X_3 are the "independent" variables.

$$X_1 = \beta_0 + \beta_{12.3}X_2 + \beta_{13.2}X_3 + \varepsilon$$

The correlation matrix for these variables is:

$r =$		X_1	X_2	X_3
	X_1	1.00	0.90	0.80
	X_2	0.90	1.00	0.50
	X_3	0.80	0.50	1.00

The beta weights corresponding $\beta_{12.3}$ and $\beta_{13.2}$ are:

$$\frac{D_{12}}{D_{11}} = \frac{0.90 - (0.80 \times 0.50)}{1.00 - (0.50)(0.50)} = 0.67 \quad \text{and} \quad \frac{D_{13}}{D_{11}} = \frac{0.80 - (0.90)(0.45)}{1.00 - (0.50)(0.50)} = 0.47$$

respectively where D_{12}, D_{13}, D_{11} are the respective determinants of the correlation matrix.

The direct effects of X_2 and X_3 are:

(beta weight of $\beta_{13.2}$) (beta weight of $\beta_{13.2}$) $(r_{22}) = (0.47)^2(1.00) = 0.22$ and
(beta weights of $\beta_{12.3}$) (beta weight of $\beta_{12.3}$) $(r_{33}) = (0.67)^2(1.00) = 0.45$

respectively.

The indirect effect of X_2 and X_3 is:

(beta weight of $\beta_{12.3}$) (beta weight of $\beta_{13.2}$) $(r_{23}) = (0.67)(0.47)(0.50) = 0.16$.

The matrix of direct and indirect effect and multiple correlation coefficient are:

	X_2	X_3	
X_2	0.45	0.16	
X_3	0.16	0.22	R^2
Total	0.61	0.38	0.99

The diagonal elements are the direct effects. The off-diagonal elements are the indirect effects.

Knowledge of the direct and indirect effects of an independent variable is useful when combined with theoretical expectations. If economic analysis does not lead the economist to the conclusion that the independent variable in mind is one of the main relevant factors and the ratio of indirect to direct effect is negative and greater than one, then the variable should be removed from the multiple regression equation because of its colinearity with other variables. A negative indirect effect is a sign of the particular "independent" variable's dependency upon independent variables in the equation in order to be statistically significant. A multiple regression equation with more truly independent variables is achieved when large negative indirect effects are deleted from the equation.

REFERENCES

1. Chenery, H. B. "Patterns of Industrial Growth," *American Economic Review*, Vol. 50, No. 4 (September 1960), pp. 625–54.
2. Deutsch, K. W.; Bliss, C. I.; and Eckstein, A. "Population, Sovereignty, and the Share of Foreign Trade," *Economic Development and Cultural Change*, Vol. 10, No. 4 (July 1962), pp. 353–66.
3. Deutsch, K. W., and Eckstein, A. "National Industrialization and the Declining Share of the International Sector, 1890–1959," *World Politics*, Vol. 13, No. 2 (January 1961), pp. 267–99.
4. Draper, N. R., and Smith, H., Jr. *Applied Regression Analysis* (New York: John Wiley and Sons, Inc., 1967).
5. Ellsworth, P. T. *The International Economy, Its Structure and Operation* (New York, 1950).
6. Food and Agriculture Organization. *Production Yearbook 1965*, Vol. 19 (Rome: United Nations Statistical Division, 1966).
7. Johnston, J. *Econometric Methods* (New York: McGraw-Hill Book Co., Inc., 1963).
8. Kindleberger, C. P. *Foreign Trade and the National Economy* (New Haven, Conn.: Yale University Press, 1962).
9. Kuznets, S. "Quantitative Aspects of the Economic Growth of Nations, IX," *Economic Development and Cultural Change*, Vol. 13, No. 1 (October 1964), pp. 1–109.
10. Lipsey, R. E. *Price and Quantity Trends in the Foreign Trade of the United States* (Princeton, N. J.: Princeton University Press, 1963).
11. Lloyd, P. J. *International Trade Problems of Small Nations* (Durham, N. C.: Duke University Press, 1968).
12. Rand-McNally and Co. *Commercial Atlas and Marketing Guide*, 99th ed. (Chicago, Ill.: Rand-McNally and Co., 1968).
13. Sombart, W. *Die Deutsche Volkswirschaft in Neunzehnten Jahrhundert*, 3rd ed. (Berlin: Bondi, 1913).
14. Steinberg, S. J., ed. *The Statesman's Yearbook 1964–1965* (New York: St. Martin's Press, 1964).

15. United Nations. *Monthly Bulletin of Statistics* (New York: United Nations Publishing Service, July 1966).
16. ———. *Yearbook of National Accounts Statistics, 1965* (New York: United Nations Publishing Service, 1966).
17. Woytinsky, W. S., and Woytinsky, E. S. *World Commerce and Governments: Trends and Outlook* (New York: The Twentieth Century Fund, 1955).