AGENCY FOR INTERNATIONAL DEVELOPMENT
WASHINGTON, D. C 20523
BIBLIOGRAPHIC INPUT SHEET

FOR AID USE ONLY
Batch 78

1. SUBJECT CLASSIFICATION

A. PRIMARY
Development and economics — DM00-0000-0000

B. SECONDARY
Industries and industrialization

2. TITLE AND SUBTITLE
The empirical relationship between nominal and effective rates of protection

3. AUTHOR(S)
Guisinger,S.E.; Schydlowsky,D.M.

4. DOCUMENT DATE	5. NUMBER OF PAGES	6. ARC NUMBER
1970	~~36p.~~ 39p.	ARC

7. REFERENCE ORGANIZATION NAME AND ADDRESS
Harvard

8. SUPPLEMENTARY NOTES *(Sponsoring Organization, Publishers, Availability)*
(prepared for Conf.on Effective Protection,Geneva,1970)

9. ABSTRACT

10. CONTROL NUMBER
PN-AAF-283

11. PRICE OF DOCUMENT

12. DESCRIPTORS
Economic factors
Tariffs
Industries
Protection
Rates

13. PROJECT NUMBER

14. CONTRACT NUMBER
CSD-1543 Res.

15. TYPE OF DOCUMENT

AID 590-1 (4-74)

CSD-1543 Res
Harvard
PN-AAF-283

THE EMPIRICAL RELATIONSHIP BETWEEN NOMINAL AND EFFECTIVE RATES OF PROTECTION

Prepared for the Conference
on Effective Protection, Geneva
December 18-20, 1970

S.E.
Stephen Guisinger
Daniel M. Schydlowsky

December 1, 1970

Appendix:
- Tables 1-3
- Appendices A,B and C
- Footnotes
- References

PROJECT FOR QUANTITATIVE RESEARCH
IN ECONOMIC DEVELOPMENT
Center for International Affairs,
Harvard University,
Cambridge, Massachusetts.

Portions of this research were supported by the Development Advisory Service, Harvard University through funds provided by the Agency for International Development under Contract CSD-1543. The views expressed in this paper do not, however, necessarily reflect the views of A.I.D.

I. Introduction

The concept of the effective rate of protection has been developed with the intention of providing a better measure of the interindustry resource allocative effects of the tariff system than nominal tariff rates, which have traditionally been used for this purpose. The effective rate, it is argued, is a better indicator of the incentives to factor movements among industries than the ad valorem tariff rates on an industry's output because it measures the increased factor rewards in the industry due to the tariff system and not just the increased price of that industry's product. Because the effective tariff is determined jointly by tariffs on inputs and outputs and by the share of value added in total output, there is no theoretical justification for assuming that the rank order of industries by their effective rate of protection is the same as the rank order by nominal rates of protection. In theory, at least, industries with high nominal tariffs on their outputs can have low effective rates of protection and low nominal tariff industries can have high effective rates.

One practical reason why the levels of nominal and effective rates should be expected to move together in industrial countries is the existence of cascaded tariff structures in those countries. With a cascaded tariff structure, the nominal rates on consumer goods generally exceed nominal rates on intermediate goods and tar-

iffs on intermediate goods exceed those on primary products. Given this pattern of tariffs, it is easy to see why effective rates for these major sectors will be cascaded in a similar fashion. Consumer goods' industries, afforded a high level of protection on their outputs, use the lesser protected intermediate goods as inputs. The intermediate goods' industries, enjoying a moderate protection on their output, use inputs from primary goods' producers which receive low levels of protection. Thus, in a simple three-sector model with cascaded tariffs, one would expect a positive correlation between the two measures of protection. In a more disaggregated model this previous reasoning does not necessarily hold. How close an association exists between nominal and effective protection at the two and three digit industry level becomes a question for empirical research.

One recent empirical study of the relation of nominal to effective rates of protection performed by Benjamin Cohen [2] suggests that for a large number of developed and less developed countries a high correlation between these two tariff measures does exist. Not only is the rank order of industries in each country quite similar using the two measures, but in most cases there is a strong linear relation between the nominal tariff for an industry and its effective tariff. On the basis of these findings, Cohen concludes: "If we want rather broad notions about either the relative average height of a country's tariffs or the ex post resource allocation effects of the tariff system, my results suggest one learns a lot--

but not everything--by looking at the ranking[of industries] by nominal tariffs." 1/

The purpose of this study is to examine closely what one can and cannot learn about the level of an industry's effective rate of protection from only the information contained in the nominal tariff rate on the industry's output. One important reason for studying the relationship between nominal and effective tariffs is the potential savings available to economists if a country's entire system of effective protection could be predicted from just the levels of nominal tariffs on outputs. The use of nominal rates on outputs to estimate effective rates would eliminate the need to compute weighted averages of input tariffs and to calculate the hypothetical "free trade" value added. If a reliable estimating equation could be determined, then empirical research in effective rates of protection would stand to gain a large windfall. The findings of Cohen's empirical research suggest that such an equation is certainly within the realm of possibility.

The empirical analysis described in this study, however, leads to a somewhat different conclusion. On the basis of our results, we would paraphrase Cohen by saying that one learns <u>something</u> but by no <u>everything</u> about the levels of effective rates of protection among industries by studying the levels of nominal tariffs on their outputs. The reason why Cohen was able to make such an optimistic statement of what could be learned about the levels of effective rates in a country

from the levels of nominal tariffs is that he unnecessarily restricts the uses to which effective rates can be put. He correctly states that, "The measurement of tariff levels [both effective and nominal] can be of at least two practical uses: as a means of comparing the level of protection among various countries (either on the 'average' or for a particular industry) and as a way of studying the effect of the tariff structure on resource allocation within a country." 2/ Cohen then proceeds to evaluate the predictions made by nominal rates in the light of these two uses, and primarily the use in measuring the resource allocative effects of the tariff system. But in this evaluation, he omits an important third use of effective rates: the measurement of the static welfare costs of tariff protection. 3/

This omission is significant because more information about the effective rate of protection is needed to study the static welfare costs of a tariff system than its resource allocative effects. The resource allocative effects can be determined knowing only the <u>relative</u> position of industries according to their effective rate whereas the welfare costs can be calculated only by knowing the <u>absolute</u> (i.e. cardinal) level of effective protection.

In this study we will demonstrate that in general the nominal tariff rate predicts much better the rank order of industries by their effective rate than it does the industry's absolute level of effective protection. Since it is possible to infer rank order from absolute measures but not the reverse, the predictions of effec-

tive rates made from nominal rates are therefore only helpful in one of the three uses of effective rates listed above. We will also demonstrate that any observed accuracy of nominal rates in predicting the absolute levels of effective rates is largely due to the type of definition used for the rate of effective protection,[4] and to the level of aggregation in the classification of industries. Finally, we will produce empirical evidence suggesting that regression equations between nominal and effective rates of protection are different between countries so that even if one country shows a very significant relationship between nominal and effective rates, the parameters of that relationship could not be "borrowed" by another country wishing to estimate absolute levels of effective protection from the levels of its own nominal tariff rates. Thus, even if a strong linear relationship did exist between nominal and effective rates for a country, it would still have to calculate at least a small sample of effective rates in order to determine the parameters of the estimating equation peculiar to its own situation.

II. The Pattern of Nominal-Effective Relationships Across Countries

This last conclusion is particularly important because a single algebraic relationship between nominal and effective rates common to all countries would make it unnecessary for any country to calculate an effective rate; it could simply estimate effective rates using the relationship observed in other countries. A visual comparison of the slopes and intercepts of the linear regressions between nominal

and effective rates in Cohen's sample (see Appendix A) suggests strongly that there is no common pattern. The slopes of the regressions vary from .32 to 2.29 and the intercepts from 68.04 to -15.01. Given that most of these coefficients are highly significant, it is doubtful that the observed range of variation in these coefficients could be due to random sampling error alone. In fact a statistical test for the homogeneity of regression parameters using a sub-sample drawn from Cohen's sample of countries bears this point out.

We performed the test for similarity among regression parameters with data from four of the countries in Cohen's sample--France, Netherlands, Belgium and Italy. [5/] The data on nominal and effective rates for these four countries have the distinct advantage that they were compiled using an identical industrial classification scheme and a common procedure for calculating effective rates. This consistency in the source of data eliminates any bias due to differing levels of aggregation and computational techniques. An added advantage is that by comparison to the coverage in Cohen's sample, these four countries are quite similar with respect to tariff levels and economic structure. If regression parameters are not similar for these countries, it is unlikely that they are similar for the rest of Cohen's sample.

The description of the tests for homogeneity and the results are contained in Appendix B. The results can be briefly summarized

as follows. The null hypothesis of overall homegeneity—i.e. both identical slopes and identical intercepts in the regression equations for the four countries—can be rejected at a better than 5% significance level. However, neither the null hypothesis for identical slopes, the intercept effect being held constant, nor the null hypothesis for identical intercepts, the slope effect being held constant, can be rejected at even a 10% level of significance. Therefore, by elimination the lack of overall homogeneity can be traced to the joint effect of both intercepts and slopes. Thus, in spite of each of these country's strong association between nominal and effective rates, the patterns of association are quite different.

It is possible, of course, that the parameters in the relationship between nominal and effective rates vary in some systematic way across countries and that the appropriate slope and intercept term for each country could be determined once the specific causes of this variation are identified. It is possible, for example, that the values of the regression parameters in a country are influenced by such factors as the importance of manufacturing in total product, the level of per capita income, export or import-competing orientation, or the average height of nominal protection.

To test for this possibility, the slope coefficients, intercepts and R^2's reported in Cohen's table of results (Appendix A of this study) were used as dependent variables in regression equations. As independent variables, the following were used: the share of the

manufacturing sector in GDP (an indicator of the level of industrial development); the level of per capita income (a measure of overall economic development); the number of industries covered in each country study (an indicator of the fineness of disaggregation); and the average level of the nominal tariff. The data and tests used in this cross-section analysis are described in Appendix C; the results can be summarized as follows.

The level of industrial development exerts a significant influence on all three parameters. The variable reflecting the relative size of the manufacturing sector is negatively associated with the intercept term and positively associated with the slope coefficient and the coefficient of determination. The other independent variables occasionally make contributions to the corrected R^2 but in only one case--the average nominal tariff in the equation with the slope parameter as the dependent variable--is one of these variables statistically significant. In no case can these independent variables explain more than 70 percent of the variance in the dependent variables, suggesting that predictions of one country's regression parameters based on these cross-section estimates would be subject to considerable error. This would be particularly true for countries with relatively small manufacturing sectors because the "goodness of fit"--i.e. the R^2's--tend to be quite low for these countries.

III. The Level of Aggregation and the Accuracy of Prediction

In his analysis Cohen noted that there appeared to be a negative association between the level of the coefficient of determination

for each country's regression and the number of industries defined for the country; he found a Spearman rank order correlation of -.42 between these two variables for the 26 countries in his sample. The regression analysis just described, however, found no significant association between these two variables: the t statistic for the coefficient of the "observations" variable in the equation with R^2 as the dependent variable was less than 1. 6/ Nevertheless, this lack of strong linear relation between these two variables should not obscure one very important characteristic of the country studies in Cohen's sample. None of the countries with highly disaggregated industrial classifications have high R^2's. The three studies with the largest number of observations--the U.S. (281 industries); Korea (218 industries); and the Philippines (89 industries)--all have coefficients of determination below the mean R^2 for Cohen's sample. 7/ The absence of a strong negative correlation is due to the fact that many studies with a small number of observations have low R^2's as well.

Since the number of observations in each study is determined by differing levels of disaggregation, the association of highly disaggregated studies with low R^2's suggests that there is some type of aggregation bias. The possibility of an aggregation bias in the coefficient of determination would pose a dilemma to the country wishing to estimate its structure of effective rates from the pattern of nominal rates. A country could make relatively good predictions at very broad levels of sectoral classification, but at more disaggregated and per-

haps more useful classifications, the accuracy of predictions would diminish.

The presence of an upward bias does seem plausible on the basis of the limited empirical evidence available. In the seven country study of effective protection directed by Bela Balassa [1], the nominal and effective rates for each country are presented at two levels of aggregation: the first at a very broad industrial classification with 10 divisions and the second at a finer breakdown containing 56 industrial classifications. In Table 1 below we show the results of regression analysis for these 14 sets of data. In all cases, the coefficients of determination are lower for the disaggregated equations than for the aggregated ones; in some cases the differences are considerable.

The theoretical relationship between the aggregated and disaggregated R^2's have been analyzed by Grunfeld and Griliches [4] and their findings offer some basis for expecting upward bias in the correlation of nominal and effective rates as the level of aggregation is increased. They derived the following relationship between two coefficients of determination: R_D^2, the coefficient of determination based on a basic, disaggregated data set and R_A^2, the coefficient of determination based on a regression involving the same data, but estimated from the means of groups within that data.

$$1. \quad \frac{R_A^2}{R_D^2} = \frac{b^2 + \dfrac{\sigma_u^2}{\sigma_t^2}}{b^2 + \dfrac{\sigma_u^2\,[1 + r_u\,(K-1)]}{\sigma_t^2\,[1 + r_t\,(K-1)]}}$$

where: r_t, r_u are the intraclass coefficients defined as:

$$r_t = \frac{K}{K-1}\,\frac{\sigma_{t*}^2}{\sigma_t^2} - \frac{1}{K-1}$$

$$r_u = \frac{K}{K-1}\,\frac{\sigma_{u*}^2}{\sigma_u^2} - \frac{1}{K-1}$$

b is the slope parameter of the two regression equations. (b_A is assumed to equal b_D)

σ_t^2, σ_{t*}^2 are the variances of the independent variable in the disaggregated and aggregated form, respectively

σ_u^2, σ_{u*}^2 are the variances of the error term in the disaggregated and aggregated equations

K is the number of observations grouped into each category of the aggregated data set. K is assumed equal across all categories

Grunfeld and Griliches show that $R_A^2 \gtreqless R_D^2$ as $r_t \gtreqless r_u$, thus we can expect an upward bias in the R^2 whenever the intraclass coefficient for the independent variable is relatively larger than the intraclass coefficient for the error term.

When K is relatively large, r_t is approximately equal to $\frac{\sigma^2_{t*}}{\sigma^2_t}$ and the numerator of this last fraction will be larger, the more similar are the observations grouped in each category. Thus, if dissimilar nominal tariffs were grouped together in each category so that the group means showed little variation, σ^2_{t*} would be small and r_t would be low. However, the aggregated classification schemes actually used in tariff analysis tend to preserve the cascaded effect of tariff systems. Therefore, we should expect to observe a relatively high intraclass coefficient for nominal tariffs. This implies a large r_t and hence the expectation of an upward bias in the coefficient of determination, R^2_A. Such an expectation is supported by the results in Table 1 where r'_t and r'_u are approximations of r_t and r_u obtained by computing the ratios $\frac{\sigma^2_{t*}}{\sigma^2_t}$ and $\frac{\sigma^2_{u*}}{\sigma^2_u}$. 8/

IV. The Definition of the Effective Rate and the Accuracy of Predictions

The degree of upward bias in the coefficients of determination does not seem particularly severe until one considers the definition for the effective rate of protection used by Cohen in his regression analysis. The two most common concepts of the effective rate are usually symbolized by the letters U and Z. U measures the subsidy given an industry as a percentage of the industry's value added at domestic prices while Z measures the subsidy as a percentage of the industry's value added at world prices. Since U is a non-linear transformation of Z,9/ one would

expect differences in the goodness of fit between equations using U and Z as dependent variables. To test this proposition empirically, we fit regressions to data on effective and nominal rates of protection for both U and Z forms and at two levels of aggregation from data collected from the Korean study [8]. The results of these regressions are shown in Table 2.

Note that at the aggregated level (34 observations), there is little difference in the fit of the two equations as measured in the R^2's. At the more disaggregated level (207 observations), the two forms of effective protection show quite radically different fits. For both the U and Z forms, the overall coefficient of determination declines with higher levels of disaggregation, but with U as the dependent variable, the R^2 declines to .57 while with Z, it falls to .02.

The low correlation for t and Z in the Korean case can be explained in part by one of the characteristics of least squares estimators, namely the tendency for very large values of the dependent variable to dominate a regression line. In the Korean data, the highest Z in the aggregate sample is only 93 while in the disaggregated set the highest is 56,385 (several other observations range between 4,000 and 7,000). As the median Z for the disaggregated set is only 61, these few observations at the upper end of the scale exert a strong pull on the regression line. Because industries with extremely high Z's were only small contributions to total industrial output, the extreme values disappear in the aggregated data and do not therefore influence the regression line to the same extent as in the disaggregated case.

With the U form of effective protection, "extreme" rates of effective protection disappear for another reason. An industry with a Z of 56,000 has a U of approximately 100. For any country the range of effective rates that will be observed when U is used runs from approximately -100 to +100,[10/] compared with the virtually unlimited range when Z is used. The range over which U is defined is very similar to the range over which nominal tariffs for most countries are observed so that the problem of an extreme value for the dependent variable which dominates a regression equation is eliminated by the use of U.

Yet, the increase in R^2's obtained by using U rather than Z in regressions involving effective and nominal rates does not necessarily improve predictions because a small error in a predicted U corresponds to a large error in the prediction of Z. A confidence interval of $\pm$ 10 for an estimated U of 80 turns out to be a much wider confidence interval for the corresponding Z value, which is 400, since values of U of 70 and 90 translate into Z values of 233 and 900. If only small changes in Z are important in analyzing the welfare costs of a tariff system, then only fractional errors in the estimates of U can be tolerated. In this light, the R^2's of equations predicting U with nominal tariff rates substantially overstate their actual ability to predict levels of effective protection measured by Z.

V. The Rank Order of Industries by Nominal and Effective Rates

In the previous sections we have demonstrated that nominal rates are not good predictors of the absolute levels of effective protection,

especially when the Z form of the effective rate is used and the number of separate industrial categories defined is quite large. For predicting the relative position of an industry's effective rate of protection, nominal tariff rates are apparently quite good and are insensitive to the types of biases just discussed. Table 3 presents rank order correlation coefficients for the aggregated and the disaggregated data in the seven country study by Balassa [1] and the Korean study [8]. Although there is a decline in all but one of the Spearman correlation coefficients in going from the aggregated to the disaggregated data, it is probably not significant, which suggests that there is no "aggregation bias" comparable to the bias observed when the absolute levels of protection are being predicted. Since U is a continuous, monotonic transformation of Z (with negative value added industries being excluded), the type of definition used for effective protection will not affect the rank order of industries. Thus, as far as the rank order of industries is concerned, one does learn a lot by looking at the levels of the nominal tariff rates, though one may still wish to calculate effective rates in order to eliminate even the small error that would occur if they were simply estimated from the rank order of nominal rates.

VI. Summary and Conclusions

In this note, we have examined the empirical correlation between nominal and effective tariffs and found that within each country, regression equations relating effective protection (measured by either Z or U) to nominal protection at relatively high levels of aggregation have reasonably high coefficients of determination. With a greater dis-

aggregation in the number of industrial classifications defined, however, the quality of this association deteriorates, especially for correlations involving the Z-form of the effective rate.[11/] Thus, what a country can learn about its structure of effective protection from the schedule of nominal rates depends critically on the desired level of aggregation. Our guess is that most countries will choose to determine the rates of effective protection for their industries at such a detailed level of disaggregation that the direct calculation of effective rates will be necessary.

If a country wishes to know the absolute levels of protection for its industries at even a high level of aggregation, it will still have to calculate at least a small sample of effective rates because, while the nominal and effective rates in each country may be highly associated, the values of the regression equation vary from country to country in a way that cannot be easily predicted. A country may choose to estimate the levels of effective rates, but will still need a small sample of effective rates from which the predictive equation specific to its own situation can be estimated. Analysis of covariance has shown that in the case of four European countries, the hypothesis of homogeneity in the slopes and intercepts of regression equations relating effective to nominal rates can be rejected, suggesting that countries cannot usefully borrow estimating equations from one another.

If only the rank order of industries by their effective rate of protection is desired, then Cohen is quite correct in stating that

one learns a lot about the levels of effective protection by studying the levels of nominal tariffs. We feel, however, that the uses of the effective rate of protection extend beyond just the measurement of the direction of the resource allocative effects of the tariff system. These other uses require a knowledge of the absolute level of effective rates at low levels of disaggregation. Only the direct calculation of effective rates of protection can accurately provide this type of information.

APPENDIX TO:

THE EMPIRICAL RELATIONSHIP BETWEEN NOMINAL AND EFFECTIVE RATES OF PROTECTION

Stephen Guisinger
Daniel M. Schydlowsky

Includes:
- Tables 1-3
- Appendices A,B and C
- Footnotes
- References

TABLE 1

	Aggregated Data						Disaggregated Data							
	(1) Obs.	(2) R^2	(3) a	(4) b	(5) σ_t^2	(6) $\hat{\sigma}_u^2$	(7) Obs.	(8) R^2	(9) a	(10) b	(11) $\hat{\sigma}_t^2$	(12) $\hat{\sigma}_u^2$	(13) r'_t (5÷11)	(14) r'_u (6÷12)
Brazil	10	.95	-24.04 (-2.54)	2.02 (12.01)	1419.12	362.82	56	.79	-23.27 (-1.92)	2.19 (14.36)	3011.45	7835.10	.47	.094
Chile	10	.81	14.88 (1.56)	1.95 (6.00)	828.32	784.97	56	.56	22.22 (1.73)	2.19 (8.28)	2034.84	629.22	.41	.10
Malaysia	9	.84	.87 (-.46)	1.58 (6.06)	40.69	22.01	53	.40	2.32 (.63)	1.16 (5.07)	310.81	215.19	.13	.035
Mexico	10	.95	-1.68 (-.68)	1.77 (12.91)	222.84	37.53	54	.85	-2.73 (-.11)	1.89 (16.92)	326.36	1351.70	.68	.17
Pakistan	9	.95	10.06 (.88)	2.33 (11.08)	2141.50	758.56	40	.89	13.99 (1.94)	2.15 (17.60)	2327.32	521.79	.92	.56
Philippines	10	.80	-7.92 (-1.14)	2.31 (5.61)	122.04	185.62	56	.73	4.95 (1.41)	1.87 (12.32)	410.44	162.54	.28	.36
Norway	10	.93	-.25 (-.12)	1.94 (10.27)	83.96	26.95	56	.63	1.03 (.53)	1.47 (9.65)	127.32	162.54	.66	.17

NOTES TO TABLE 1

1. All data was taken from Balassa [1] chapter 4. For several countries, the regression equations were estimated from a smaller sample than the 10 industries in the aggregated data set and the 56 industries in the disaggregated data set compiled by the Balassa study. This reduction was made necessary by the presence of industries with "negative values added" at world prices, whose effective rates of protection cannot be measured on the same interval scale with the effective rates of industries having positive values added at world prices. For a discussion of the measurement problem posed by "negative value added" industries, see Guisinger [5].

2. The regression equation estimated in each case is: $Z = a + b \cdot t + \varepsilon$. Values of the t-statistic for coefficients are shown in parentheses.

TABLE 2

	Aggregated Data						Disaggregated Data							
	(1) No. of Obs.	(2) r^2	(3) a	(4) b	(5) σ^2_t	(6) $\hat{\sigma}^2_u$	(7) No. of Obs.	(8) r^2	(9) a	(10) b	(11) σ^2_t	(12) $\hat{\sigma}^2_u$	(13) r'_t (5÷11)	(14) r'_u (6÷12)
Korea (Z)	34	.66	2.37 (.40)	.72 (7.86)	814.58	226.00	207	.023	-122.00 (-.30)	12.54 (2.19)	2304.00	1555.	0[4]	.35	.00001
Korea (U)	34	.65	6.42 (2.05)	.379 (7.7[illegible])	814.58	64.00	207	.572	5.53 (2.21)	.585 (16.56)	2304.00	592.	0	.35	.11

NOTES TO TABLE 2:

1. As in Table 1, all "negative value added" industries were excluded from our regression analysis.
2. The regression equations estimated was Z (or U) $= a + b \cdot t + \varepsilon$.
3. All data was taken from Korean Development Association [8].
4. t-values are shown in parentheses.

TABLE 3

Country	Aggregated Data		Disaggregated Data	
	No. of Industries	Spearman R	No. of Industries	Spearman R
1. Norway	10	.96	56	.82
2. Philippines	10	.93	56	.83
3. Mexico	10	.95	53	.82
4. Pakistan	9	.93	54	.90
5. Chile	10	.90	40	.88
6. Malaysia	9	.79	56	.88
7. Brazil	10	.96	56	.94
8. Korea	34	.89	218	.92

Source: Balassa [1] (see Note to Table 1)

APPENDIX A

Country and Year	Number of Industries	Spearman Rank Correlation	Tariffs U = a + b N: a (t ratio)	b (t ratio)	R^2	F	
	(1)	(2)	(3)	(4)	(5)	(6)	
Argentina, 1953	29	.89	-15.01 (-1.44)	1.15 (9.72)	.78	94.4	*
Belgium, 1959	29	.83	-6.98 (-2.42)	1.50 (7.75)	.69	60.1	*
Brazil, 1966	21	.96	21.0 (5.68)	.32 (11.0)	.86	120.4	*
Brazil, 1967	21	.95	2.78 (.56)	.72 (8.89)	.81	79.0	
EEC, 1959	29	.94	-9.75 (-6.85)	1.50 (17.2)	.92	296.1	
EEC, 1962	36	.85	-4.60 (-1.81)	1.61 (8.68)	.69	75.3	
France, 1959	29	.92	-13.88 (-4.37)	1.22 (12.80)	.86	163.9	*
Israel, 1961	30	.80	-3.38 (-.37)	.70 (5.09)	.48	25.9	*
Italy, 1959	29	.96	-12.95 (-7.20)	1.56 (17.4)	.92	303.3	*
Japan, 1962	36	.71	3.09 (.97)	1.11 (6.65)	.57	44.2	*
Korea, 1963-65	218	.92	6.48 (2.47)	.61 (17.2)	.58	297.3	*
Malaysia, 1963	45	.73	3.00 (.50)	.48 (1.44)	.05**	2.1	
Malaysia, 1965	45	.86	2.19 (.43)	.76 (3.07)	.18	9.5	*
Netherlands, 1959	29	.87	-11.75 (-5.95)	2.29 (8.76)	.74	76.7	*
Pakistan prices, 1963/64	32	.49	53.1 (2.01)	.36 (1.92)	.11**	3.7	
Pakistan, 1963/64	32	.72	- .88 (-.08)	.64 (4.96)	.45	24.6	

APPENDIX A (continued)

Country and Year	Number of Industries	Spearman Rank Correlation	Tariffs U = a + b N: a (t ratio)	b (t ratio)	R^2	F
	(1)	(2)	(3)	(4)	(5)	(6)
Philippines, 1961-65	89	.92	13.93 (1.63)	.84 (9.08)	.49	82.4 *
Sweden, 1962	36	.81	1.20 (.67)	1.49 (8.08)	.66	65.3 *
Taiwan, 1965	37	.80	26.8 (6.05)	.53 (5.68)	.48	32.3 *
Turkey, 1960's	7	.06	68.3 (3.01)	-.05 (-.17)	.01**	.03
U.K., 1962	36	.76	3.03 (.86)	1.13 (5.57)	.48	31.0 *
U.S.A., 1962	36	.84	.89 (.33)	1.24 (6.54)	.56	42.7
U.S.A., 1958-60	281	.37	2.00 (1.98)	1.14 (17.88)	.53	319.6 *
West Germany, 1959	29	.93	-10.70 (-12.0)	1.60 (15.1)	.89	228.7 *

** Indicates significance at 1 percent level.

* Indicates countries selected for sample of 16; see Appendix C.

Sources: Cohen [2]

APPENDIX B

Tests for homogeneity of regression parameters among groups of observations have been extensively used in time series analysis where the question often arises whether cross-section data for different time periods can be pooled. The problem discussed in the body of this note is logically identical: we wish to know whether these parameters vary from country to country rather than from time period to time period. Tests for homogeneity are discussed in Wallace and Toro [1]] and Huang [6]. We shall not attempt a full description of these tests, but only outline briefly the method.

Consider the following model:

1. $Z = B_1 + B_2 \cdot t + \varepsilon \qquad \varepsilon \sim N(0,\sigma^2 I)$

where: Z is the effective rate of protection (equal to the percentage increase in domestic per unit value added over world price per unit value added

t is the nominal tariff or tariff equivalent

Auume further that there are five such regression models, one each for each of the four countries--France, Netherlands, Belgium, and Italy--and one such equation for the pooled observations. Letting the superscripts F,N,B,I and T identify each of these regressions, we can formulate the null hypothesis as follows:

$H_0 : B^F = B^N = B^B = B^I = B^T$

where: B is the 2 x 1 vector, $B = \begin{vmatrix} B_1 \\ B_2 \end{vmatrix}$

To determine whether the null hypothesis is to be rejected or accepted we need to compare two kinds of sums of squared errors. Letting S^T represent the sum of squared errors (SSE) calculated from the pooled regression and S^I represent the <u>sum</u> of the SSE's for each

of the four country equations, we can perform the test for overall homogeneity by computing:

$$2. \quad F_1 = \frac{\frac{S^T - S^I}{2(T-1)}}{\frac{S^I}{T(J-2)}}$$

where: T is the number of groups (countries)

J is the number of observations in each group

If F_1 exceeds the F-statistic for the prespecified level of confidence, we can reject the null hypothesis.

Table B-1 shows the parameters and characteristics of each of the five equations in lines 1-5. Since T=4 and J=29, F_1 is calculated as:

$$3, \quad F_1 = \frac{\frac{8584 - 6019}{6}}{\frac{6019}{108}} = 7.67$$

Since the non-central F for a level of confidence of .05 for 6 and 108 degrees of freedom is less than 2.53, the observed level of F_1 clearly suggests that the null hypothesis, H_o, should be rejected.[1/]

Given that there is a lack of overall homogeneity, due either to differing slopes or differing intercepts or possibly both, the next step is to determine its source. A null hypothesis to test for the similarity

---

1/ The rationale for the use of the non-central F in tests of this type is developed in Wallace and Toro [12]; the F-values shown in this appendix were drawn from their table which computes the non-central F with a centrality parameter of .5.

TABLE B-1

Regression	No. of Observations	Intercept (B_1)	Slope (B_2)	R^2	SSE	Identification
1. France	29	-20.8	1.88	.88	2692	
2. Netherlands	29	-12.3	2.57	.73	798	
3. Belgium	29	-10.53	2.01	.66	1532	
4. Italy	29	-18.51	2.21	.91	997	
Total					6019	$= S^I$
5. Pooled	116	- 8.32	1.65	.843	8585	$= S^T$
6. Pooled -- varying intercepts	116	F = -24.4 N = - 8.3 B = -10.4 I = -14.7	2.00	.89	6292	$= S^D$
7. Pooled-- varying slopes	116	-15.85	F = 1.75 N = 2.99 O = 2.35 I = 2.10	.88	6340	$= S^C$

Source: Data taken from Grubel and Johnson [3]

in slopes can be formulated as follows:

$$H'_o: \quad B_2^F = B_2^N = B_2^B = B_2^I = B_2^T$$

The test of this hypothesis requires that the following regression be computed:

4. $$Z = B_1^F \cdot D^F + B_1^N \cdot D^N + B_1^B \cdot D^B + B_1^I \cdot D^I + B_2 \cdot t + \varepsilon$$

where: D is a dummy variable with 1,0 values

Letting the SSE of the regression estimated from equation 4 be represented by S^D, the appropriate F-test on which to accept or reject H'_o is:

5. $$F_2 = \frac{\frac{S^D - S^I}{T-1}}{\frac{S^I}{T(J-2)}}$$

Substituting the appropriate values in equation 5:

6. $$F_2 = \frac{\frac{6292-6019}{3}}{\frac{6019}{108}} = 1.633$$

Since the non-central F for 3 and 120 degrees of freedom at even a 25% level of confidence is 1.856, we cannot reject H'_o.

Another possible cause for lack of homogeneity is, of course, differing intercepts. The appropriate null hypothesis can be formulated as:

$$H''_o: \quad B_1^F = B_1^H = B_1^B = B_1^I = B_1^T$$

The SSE of the appropriate regression for testing H''_o is taken from line 7 in Table B-1 and is represented by S^C.

The appropriate F-test is:

7. $$F_3 = \frac{\frac{S^C - S^I}{T-1}}{\frac{S^I}{T(J-2)}}$$

Substituting values from Table B-1 in equation 7.

$$8. \quad F_3 = \frac{\frac{6340-6019}{3}}{\frac{6019}{108}} = 1.92$$

This F-statistic is not significant at the 10% level.

Thus, by elimination the lack of homogeneity can be traced to a joint effect of both slopes and intercepts.

APPENDIX C

Cross-Section Analysis of Regression Parameters

The three dependent variables analyzed are the slopes (the "b's" of Cohen's table of results), the intercepts (the "a's") and the coefficients of determination (the "R^2's"). Four independent variables were used in the regression analysis: the number of industrial classifications defined in each country's study of effective protection (N); the average nominal tariff of the country (T); the size of the manufacturing sector in relation to total GDP (M); and the country's per capita GDP measured in U.S. dollars (Y).

Data from twenty-two of the twenty-four countries in Cohen's sample were used to estimate each regression equation. Turkey was excluded because of the extremely low number of industries (7) contained in the study of effective rates for that country. One study for Pakistan was excluded because it omitted the effects of quotas in estimating effective rates of protection. Data for the three dependent variables and two of the independent variables--number of industrial classification defined (N) and the average (unweighted) tariff (T)-- were taken from Cohen [2]. The data on manufacturing shares (M) and per capita income (Y) were gathered from World Bank sources.

The results of the estimated regressions below show only the "best fits"; variables with t ratios less than one do not contribute to the corrected R^2 and therefore have been dropped. The R^2 reported for each equation is uncorrected for degrees of freedom. The t ratios for coefficients are shown in parentheses.

1. $a = 6.9 + .053 \cdot N - .50 \cdot M + .145 \cdot T$ $\quad R^2 = .44$
$\quad (1.89) \quad (-2.56) \quad (1.06)$ $\quad F = 5.27$

2. $b = .787 + .026 \cdot M - .018 \cdot T$ $\quad R^2 = .70$
$\quad (4.05) \quad (-4.14)$ $\quad F = 24.4$

3. $R^2 = .025 - .00007 \cdot Y + .022 \cdot M + .0038 \cdot T$ $\quad R^2 = .67$
$\quad (-1.31) \quad (5.80) \quad (1.49)$ $\quad F = 13.485$

Note that the size of the manufacturing sector is a consistently strong explanatory variable in the three equations. In none of the equations, however, is more than 70 percent of the variance in the dependent variable explained.

Because the use of Cohen's sample which contains several studies for the same country--e.g. Brazil, studies which are aggregations of other studies--e.g. the two EEC studies, and regression parameters for some countries which are statistically insignificant could lead to biases, we estimated the three regression equations from an "improved" sub-sample of the 16 "best" equations in Cohen's sample (those marked with an asterisk in Appendix A). The use of this sub-sample did not, however, improve the corrected R^2's.

One important qualification to the statistical procedures employed in this analysis deserves special emphasis. The use of the regression parameters and the coefficient of determination as variables in a regression analysis leads to an "errors in variables" bias since each observation itself is subject to sampling error. No attempt was made to correct for this bias since these results are intended only to be suggestive.

FOOTNOTES

The authors are Assistant Professors at Southern Methodist University and Harvard University, respectively. We would like to thank Thomas Johnson for his comments on an earlier draft. Richard Stakem performed the statistical computations and made several useful suggestions. Portions of the research involved in this study were supported by the Harvard Development Advisory Service through funds provided by The Agency for International Development.

1. Cohen [2], p. 14.

2. ibid, p.3

3. For discussions of the static welfare costs of tariff systems using the concept of the effective rate of protection see Johnson [7], Balassa [1], Snape [10] and Krueger [9].

4. The effective rate of protection can be defined as either the tariff-created subsidy as a share of domestic value added or the subsidy as a share of world-priced value added. This point is discussed in greater detail in a later section.

5. These four countries were part of a five country study of effective protection conducted by Grubel and Johnson [3]. The fifth country (Germany) included in the Grubel and Johnson study was excluded because of the dissimilarity of coverage; effective rates were calculated for only 15 separate industries in Germany whereas 29 different effective rates were calculated for the other four countries.

6. Multi-collinearity between this variable and the other independent variables may be partly responsible for the low t statistic, but this seems unlikely.

7. The R^2's shown in Appendix A are uncorrected but even the mean of corrected R^2's would exceed the corrected R^2 for these three countries.

8. One reason why only an approximation is being made is that in equation 1 above, K is assumed equal for each category in the aggregated data set. This assumption is not met by the data in Table 1.

9. It can be shown that $Z = \frac{U}{1-U}$ and $U = \frac{Z}{1+Z}$.

10. Assuming that "negative value added" industries have been excluded.

11. One recent study does, however, suggest that even at highly disaggregated industrial classifications, the association of nominal and effective rates <u>over time</u> may still be quite strong. In a study of the U.S. tariff system during the period 1870-1914, Sundararajan [11] calculated both nominal and effective rates annually for the Pig Iron industry over the 45 year period. We performed a regression analysis between these two variables and found the following relationship:

$$Z = -.40 + \underset{(16.5)}{3.6 \cdot t} \qquad R^2 = .87$$

REFERENCES

1. Balassa, B., (ed.), "The Structure of Protection in Developing Countries," World Bank, (June, 1970), mimeo.
2. Cohen, Benjamin, "The Use of Effective Tariffs," Discussion Paper No. 62, Economic Growth Centre, Yale University, (Feb., 1969).
3. Grubel, H.G. and Johnson, H.G., "Nominal Tariffs, Indirect Taxes and Effective Rates of Protection: The Common Market Countries 1959," Economic Journal, LXXVII, 308, (December, 1967).
4. Grunfeld, Y. and Griliches, Z., "Is Aggregation Necessarily Bad?," The Review of Economics and Statistics, (February, 1960)
5. Guisinger, S.E., "Negative Value Added and the Theory of Effective Protection," Quarterly Journal of Economics (August, 1969).
6. Huang, David, Regression and Econometric Methods, John Wiley and Sons, 1970.
7. Johnson, H.G., "The Cost of Protection and the Scientific Tariff," Journal of Political Economy, (August, 1960).
8. Korean Development Association, Effective Rates of Korean Industries, 1967.
9. Krueger, A.O., "Some Economic Costs of Exchange Control: the Turkish Case," Journal of Political Economy, (October, 1966).
10. Snape, R.H., "Sugar: Costs of Protection and Taxation," Economica, (February, 1969).
11. Sundararajan, V., "The Impact of the Tariff on Some Selected Products of the U.S. Iron and Steel Industry, 1870-1914," The Quarterly Journal of Economics, (November, 1970).
12. Wallace, T.D. and Toro, C.E., "Tables for the Mean Square Error Test for Exact Linear Restrictions in Regression," The Journal of the American Statistical Association, (December, 1969).