PART I

The Balance of Payments

2

Time-Series Estimation of Import and Export Demand Relationships

In the present chapter we shall discuss the application of traditional multivariate least squares regression methods to the analysis of import and export time series. Our object is not to review regression methods in detail, since there are many standard books on the subject.[1] Rather, our primary concern will be to bring together the many questions with which a researcher must cope in planning his own work and in evaluating the work of others.

The foundation of the statistical research to be discussed in this chapter is a hypothesized behavioral relationship on the demand side between the level of imports (exports) of goods and services and several explanatory variables.[2] This relationship is assumed to have held consistently throughout the data period. When forecasting is performed, the relationship is extended by assumption into the future as well.

The central problem at issue is the specification of the import (export) demand relationship in a form suitable for statistical fitting. In this regard, we will be concerned with selection of appropriate dependent and independent variables, choice of functional form, and method for handling response lags. We shall also discuss certain special problems of estimation that may be relevant for particular countries and classes of goods, and appropriate formats for the presentation of results.

Some readers may be disturbed by the multitude of questions that will be posed in our discussion in comparison with the sparseness of definite answers. If this is indeed the case, we will have succeeded in our purpose, for it is far too common for the important questions to be glossed over and for research

[1] For introductory, intermediate, and advanced discussions of regression analysis, see for example the works by Suits [71], Johnston [29], and Goldberger [19].

[2] There will be a behavioral relationship between imports (exports) and several explanatory variables on the supply side as well. Much of the discussion to follow is relevant to the supply side. Unfortunately, the rather meagre empirical attempts to uncover supply relationships do not warrant extensive comment. See Rhomberg [62, 65] for examples of supply functions.

decisions to be fallen into rather than arrived at by explicit design. The fact that we shall concentrate mainly on statistical questions should not be interpreted to mean that knowledge of the institutions and the economic characteristics of the markets to be analyzed is unimportant. This surely is not the case, and any research design will be measurably improved if the investigator takes the time to learn about these matters.

CHOICE OF VARIABLES

The Dependent Variable As already mentioned, our focus here will be on import and export demand relationships. The most readily available data on imports (exports) are in value rather than quantity terms. However, the theory of demand suggests that quantity is the appropriate dependent variable, and we will have to divide or deflate the value series by a measure of prices to obtain the proper dependent variable. Thus the dependent variables are given by

$$M = \frac{V_M}{p_M} \tag{2.1}$$

and

$$X = \frac{V_X}{p_X} \tag{2.2}$$

where M = quantity of imports of some commodity class; p_M = price of imports; V_M = value of imports and X, p_X, and V_X are defined analogously for exports.

It should be obvious that for goods that are homogeneous in quality, M and X will be accurate measures of quantity. But when goods differ in quality and when classes of goods are combined into aggregates, M and X may bear little or no relation to real quantity.[3] In such cases the price variables are index numbers and the quantity variables are values in constant dollars. The defining relationship (2.1) is expressed explicitly by

$$M_t = \sum_i p_{i0} q_{it} = \frac{\Sigma_i p_{it} q_{it}}{\Sigma_i [(p_{i0} q_{it})/(\Sigma_i p_{i0} q_{it})](p_{it}/p_{i0})} = \frac{V_M}{p_M} \tag{2.3}$$

[3] Since countries import and export typically thousands of different types of goods, it is necessary to employ index numbers of prices for purposes of deflation and, as will be noted shortly, as explanatory variables in the analysis. This suggests the importance of precisely defining the characteristics of the goods to be included in the index in order to minimize the variations in prices due to quality differences. Needless to say, this is a very difficult thing to accomplish and one should be continually on guard therefore in the use of published price indexes and in the calculation of average prices or unit values. This point is further discussed in the appendix. A useful general source on index number construction is Mills [47]. See also Kravis and Lipsey [41] and Lipsey [44].

where p_{i0}, q_{i0}, p_{it}, q_{it} are the price and quantity of imports of commodity i in the base period and period t respectively. Interpreting this relationship, we observe that the dependent variable M_t is the value of imports at base-year prices, which may be expressed as current value divided by a price index. M and X are interpreted as the "real values" of imports and exports in the same way that the deflated value of gross national product (GNP) is called "real income."

Occasionally, researchers have used the value of imports in current dollars as the dependent variable. On theoretical grounds this should be avoided. In the appendix to this chapter we demonstrate that if the definition (2.3) is used, then there is a macro relation between M_t and several explanatory variables, including income and price indexes. For other definitions of M, the proper weights to be used in the import price index become somewhat more complicated.[4] However, when only a crudely constructed price index is available, it may be preferable to use the current-value variable and avoid the error introduced by deflating by the crude price variable. In addition, there will be cases when it is extremely difficult to express imports and exports in real terms, as for example with services, tourism, and banking charges. It is common in these circumstances to carry out the analysis using current-value data.[5]

Independent or Explanatory Variables The basic explanatory variables are suggested by the theory of demand, according to which the consumer allocates his income among consumable commodities in an effort to achieve maximum satisfaction. The quantity of imports purchased by any consumer will thus depend on his income, the price of imports, and the price of other consumable commodities. This suggests that for an economy we may write import demand as

$$M = \frac{V_M}{p_M} = f(p_M, p_Y, Y) \tag{2.4}$$

[4] The reader may verify this by demonstrating that (2.A.19) in the appendix would include a q_j' term.

[5] It might be argued, as Branson [7] does, that our ultimate interest is in the balance of payments and consequently in the *value* of imports. When the constant-dollar variable is used, a second equation to explain prices will be needed to predict the import value. It remains an open empirical question whether a single equation with value as the dependent variable would provide a better prediction than the two-equation explanation that is theoretically preferable.

There is one theoretically acceptable alternative to the variable defined by (2.3). It may be the case that disturbances to our hypothesized behavioral relationship will grow in size as M grows, perhaps maintaining a proportionality to M. Least squares procedure, in contrast, assumes that the disturbances maintain their absolute size. This conflict can perhaps be dealt with by dividing M by real income. The ratio of imports to real income is not likely to change significantly, and the assumption of constant-sized disturbances to a relationship explaining this ratio thus seems acceptable. The importance of the foregoing point will depend on the growth of M over the data period. We will not dwell on it further.

where Y is domestic money income, p_M is the price level of imports, and p_Y is the price level of other goods, in this case domestic goods. The fact that the demand relations for individual consumers can be aggregated over individuals and over commodities to yield (2.4) is supported by theorems on aggregation that are presented in the appendix to this chapter.

The theory of demand proceeds one step further in suggesting that the demand relationship (2.4) may actually be written as

$$M = f\left(\frac{p_M}{p_Y}, \frac{Y}{p_Y}\right) \tag{2.5}$$

or

$$M = g\left(\frac{p_Y}{p_M}, \frac{Y}{p_M}\right) \tag{2.6}$$

This transformation is based on the assumption that individual consumers display the absence of money illusion; that is, a doubling of all prices and money income will leave the quantity demanded unchanged. In that event, either (2.5) or (2.6) should be preferred to (2.4). The point at issue is whether we are so confident concerning the absence of money illusion that we will impose this presumption on the data, or whether the data should be allowed to support or to contradict the absence of money illusion hypothesis. In our judgment, the theoretical support for the absence of money illusion is not sufficiently strong to justify Equation (2.5) or (2.6), and we therefore prefer Equation (2.4).[6] It may be noted nonetheless that Equation (2.5) is the form which has traditionally been employed in demand analysis in international trade. In deference to tradition and with simplification as a side product, we will employ Equation (2.5) as the basic description of demand throughout the remainder of the chapter.

The export-demand function can be written analogously as

$$X = \frac{V_X}{p_X} = g\left(\frac{Y'}{p_{Y'}}, \frac{p_X}{p_{Y'}}\right) \tag{2.7}$$

where the primed values refer to income and prices in the rest of the world.

[6] Researchers will sometimes shy away from the more general form (2.4) due to the enlarged standard errors inherent in using a third explanatory variable that is likely to be correlated with the other two. While such reasoning is understandable, it is in our judgment unacceptable. In the first place, the proper statistical method does not allow for the choice of specification on the basis of standard errors. Secondly, the comparison of the standard errors obtained via Equations (2.4) and (2.5) is not particularly relevant to the question at hand. That is, the coefficient on the price term from (2.5) describes the effect of a change in p_M/p_Y holding Y/p_Y constant. A comparable statistic from (2.4) will be a combination of all three of the basic coefficients, the standard error of which will include covariance terms and cannot be read directly from the standard errors of the basic coefficients.

The theory of import demand we have just discussed is based on the proposition that imports and domestic goods are not perfect substitutes. Suppose, however, as illustrated in Figure 2.1, that imports and domestic goods are perfect substitutes, or equivalently that the price elasticities are very large. In Figure 2.1, DD is the domestic demand for some good while SS is the domestic supply. The difference between these schedules MM represents an excess demand or equivalently a demand for imports, where imports are the same good as that produced domestically. From an empirical point of view, the very important difference between the situation discussed earlier in connection with Equations (2.4) et seq. and that depicted in Figure 2.1 is that in the former case domestic supply will influence imports only through its impact on domestic prices, while in the latter case domestic supply will directly influence imports. Figure 2.1 suggests accordingly that the import demand function should include domestic supply variables.

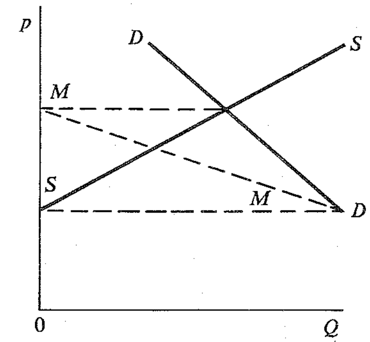


FIGURE 2.1

Demand for Imports: Perfect Substitutability Between Home and Foreign Goods

To illustrate this point further, suppose that the international supply of the good in question is infinitely elastic at the given price and that domestic investment increases the capacity of import-competing industries. In our first situation, domestic prices will fall and imports will be reduced. In the second situation, the domestic and import prices must be the same as long as some of both goods is being sold. Accordingly, no price change is observed, yet at the same time imports will be reduced. The only way to account for this is to include the capacity of the import-competing industries as an explanatory variable in the import-demand function. The basic import-demand function suggested by our second situation is given by

$$M = f(S, Y, p, p_A) \tag{2.8}$$

where S is a variable that shifts the domestic supply function, Y is money income, p is the common price of the good from domestic and foreign sources of supply, and p_A is the price of an alternative domestic good that is not a perfect substitute for the good in question. If M is total aggregate imports, then there will be no p_A at all.

We have therefore two quite different descriptions of import demand. In the first instance, prices will be observed to move and will allocate income among the alternative consumable commodities, imports and domestic goods. In the second instance, the effect of a price change is so great that no price change can occur and the only way to account for changes in imports due to competition between the sources of supply is through supply variables.[7]

Virtually all of the statistical studies of import demand that have come to our attention have used some variant of Equation (2.5).[8] What our discussion suggests, however, is that a more proper approach would be to distinguish commodities that are perfect substitutes from those that are not and to adopt the appropriate specification of the import function in each case. Any particular commodity may be placed in one class or the other according to how well the competing specifications in Equations (2.5) and (2.8) explain the data.

Inasmuch as there have been no efforts, of which we are aware, to explore relationship (2.8), we shall not pursue it further at this point, beyond indicating the sort of variable that is associated with S. The variable S should represent those factors that will affect the supply of import-competing goods. What comes to mind particularly in this regard is the capacity of the import-competing industries that may be reflected in recent investment. Other possible factors include the cost of inputs such as labor or raw materials.

A similar description of import demand will be needed to describe imports of raw materials and unfinished goods. These are inputs into productive activities and may be explained by

$$M = f(p_M, p_A, O) \tag{2.9}$$

where p_M is the price of the import, p_A is the price of an alternative domestic input, and O is the level of production or output of the industries in question. This is seen to be analogous to the finished-good demand function (2.4). In the case of infinite price elasticities, there will also be a function analogous to Equation (2.8).

[7] This discussion illustrates, incidentally, a consideration that will work against relatively high estimates of price elasticities. When elasticities are in fact high, little or no price movement is observed, making estimation of the elasticities subject to inaccuracy.

[8] Branson's [7] theory suggests that he is describing imports as an excess demand function, but it should be noted that he has in addition the relative price variable associated with the other theory.

Both the consumer and the producer items may be further divided into durables and nondurables. There are quite a number of studies that have examined the demand for consumer durables and investment goods (producer durables). While in principle these studies might well be adapted in dealing with import demand, they have unfortunately not yet had much impact on the international trade field. The importance of these considerations will of course depend on the extent to which the demand for durables is motivated differently from nondurables and also on the importance of durables in total imports (exports).

Let us now consider other possible explanatory variables. Some of these will be seen simply to represent variations of the basic real income and relative price variables to make allowance for particular categories of imports, while others are acknowledgements on the part of the researcher that understanding a complex demand phenomenon requires more variables than the usual two. Examples of variables commonly used in import-demand analysis are indicated in Table 2.1.

It will be noted in line 1 of Table 2.1 that the income or "activity" variable is chosen to conform with the particular import category and that further subdivisions of these variables can be used for more refined categories. This suggests that the use of an aggregate income or an aggregate production term

TABLE 2.1

Explanatory Variables Used in Import-Demand Analysis

Total Imports	Imports of Finished Goods	Imports of Unfinished Goods
1. Real GNP; degree of capacity utilization	1. Real disposable income; real expenditure components; degree of capacity utilization	1. Industrial production; real change in inventories; degree of capacity utilization
2. Relative price of imports [a]	2. Relative price of imports [a]	2. Relative price of imports [a]
3. Dummy variables for unusual periods	3. Dummy variables for unusual periods	3. Dummy variables for unusual periods
4. Dummy variables for seasonal variation	4. Dummy variables for seasonal variation	4. Dummy variables for seasonal variation
5. Lagged variables	5. Lagged variables	5. Lagged variables
6. Foreign exchange reserves	6. Foreign exchange reserves	6. Foreign exchange reserves
7. Credit	7. Credit	7. Credit

[a] Measured as import price divided by the price of domestic goods in general or as import price divided by the price of close domestic substitutes. The specification will depend especially on the level of aggregation employed.

in the aggregate import function is improper insofar as increases in particular categories of income or output may generate significantly more imports than similar increases in other categories. This is especially true for the aggregate production variable. The differences in the import requirements of the various productive activities are likely to be quite substantial. In order to take such differences into account, as noted in the detailed discussion of aggregation in the appendix to this chapter, the components of the aggregate activity variables ought to be reweighted according to their marginal contribution to imports. These weights may be approximated in the case of the production variable by the average import content of the commodity class. The use of the inventory variable separately may be interpreted to reflect the assumption that increases in GNP due to changes in inventories have a unique and separable impact on imports.

The capacity-utilization variable represents an amendment to the traditional theory of demand insofar as it gives cognizance to the idea that queues as well as prices may be used to allocate goods among consumers. Thus, an increase in domestic demand may not be met immediately by price increases. Rather, domestic producers may ration the available supply by delaying deliveries or, in other words, forcing the consumer into a queue to await servicing of his order. In such a period, the consumer may look to foreign sources of supply to avoid the delay in delivery. The consumer therefore pays two prices for the good he desires: the quoted price as well as an imputed cost of waiting in a queue. He will seek the supply source that provides the good at minimum total cost, including the cost of waiting.

What this suggests accordingly is that the import-demand function should include variables that reflect the length of queues at home and abroad. Capacity utilization is a proxy for queue-length. When production is close to capacity, queues are likely to be long. In periods of excess capacity, orders are likely to be filled rapidly. The inventory variable may be interpreted as a proxy for queue-length as well. Increases in demand may be met initially by a drawing down of inventories and later by increases in queue-length. Accordingly, a disinvestment in inventories in one period may be a signal for increased queue-length in the next.

The two-good description of demand is misleading in discussion of the relative price term noted in line 2 of Table 2.1. The dependent variable, imports, typically represents, as we have noted, a conglomerate of goods that substitute freely for some domestic goods, not so freely for others, and not at all for still others. Ideally one would use a price relative for each of the first two classes of goods. The specification of these two classes of goods will of course depend on the goods that make up the dependent variable, imports, and consequently on the level of aggregation. When aggregate imports is the dependent variable, the use of only the price index for GNP in the price relative involves the assumption that imports substitute generally the same with

all domestic goods. Such an assumption is rarely warranted, and the procedure will surely impair the price relative's explanatory power. This is again a problem of aggregation, which is treated extensively in the appendix.

Another complication is introduced when price indexes are not available, and unit-value indexes are used instead. A unit value may be calculated for particular commodity classes of imports (exports) by dividing the value of imports (exports) of that class by the unweighted sum of the quantities imported. A unit-value index is thus a weighted average of such unit values, and it may change because of the commodity composition of any of the commodity classes quite independently of any price changes. Since a unit-value index is thus not ordinarily a true price index, one must exercise great care in its use and interpretation.

The dummy variables in line 3 are designed to allow for the effects upon imports of unusual occurrences such as a strike, war, or natural disaster. Such variables would assume a value of one for the duration of the unusual period and zero otherwise. If quarterly or monthly data are being used, it may be desirable, as indicated in line 4, to employ dummy variables to reflect the seasonal variation in the relationship.[9] The alternative here would be to use data that were already seasonally adjusted. However, this might have the disadvantage of imposing possibly arbitrary regularities on the data, which might be inappropriate for the particular demand function being estimated.[10]

Line 5 of Table 2.1 referring to lagged variables is of particular importance in measuring the influences of past changes in the independent variables on the current behavior of imports. The effects of lags will be more important the shorter the time-period units utilized in the analysis. Thus, for example, in a quarterly analysis current imports may be influenced more by the prices of preceding quarters than by current prices. This is because of such factors as the lag between orders and shipments and the speed with which imports are adjusted to changes in prices. A common way of introducing lagged influences is by including as an independent variable the value of the dependent variable lagged one quarter.[11]

The level of foreign exchange reserves indicated in line 6 may be relevant in particular to less developed countries, where the reserve position can be considered indicative of the strictness of controls affecting imports. That is,

[9] See Suits [72] or Johnston [29] for a discussion of the uses and interpretation of dummy variables.

[10] A point worth mentioning here is that seasonal adjustment will use up the same number of degrees of freedom whether performed before or during the import regression. Thus, those who use adjusted data should reduce their degrees of freedom by the appropriate numbers. It is then clear that any argument in favor of adjusted data may not appeal to degrees-of-freedom restrictions.

[11] This would correspond to the so-called Koyck distributed lag, in which the weights of past influences are assumed to decline geometrically. We shall have much more to say on these matters later in this chapter.

imports may respond in these countries more to foreign exchange availability than to the level of real income. A similar variable would be foreign exchange earnings. In the case of developed countries, the reserves variable can be used as a proxy for the degree of official restraints. Thus, for example, as the U.S. has lost more and more reserves, government officials spend more and more time persuading consumers and producers to "buy American."

The "credit" variable indicated in the last line of Table 2.1 has been neglected in most studies to date. This variable is meant to indicate the availability and terms at which credit is provided for the financing of imports. Such a variable will play an important role especially in linking the current and capital accounts of the balance of payments. Increasing interest in the capital account will surely provide a stimulus to increased examination of the effect of credit on imports and exports.[12]

A table could be constructed for exports that would be analogous in principle to Table 2.1. Data do not exist, however, for world income and world prices, which would be determining variables. In cases when the country in question exports primarily to countries or regions that have published and reliable income and price data, export demand may then be viewed as primarily the demand for imports by these important countries or regions.[13] It is necessary otherwise to seek proxy measures for world income and world prices.

The most common choices are to use real world exports (less those of the country in question) as a proxy for world income and a measure of the price of world exports as a proxy for the price of world goods. This amounts roughly to assuming that the country exports a certain fraction of world exports and that this fraction is altered by the country's relative export prices.[14] The interpretation of the export-demand relationship will thus differ from the one for imports because in the case of exports neither world income nor world prices is being measured directly. There may in addition be a complication introduced by relying on unit values rather than on actual prices in the construction of the indexes used to measure relative prices.

The variables discussed above are in general the most important ones, although the list is by no means all-inclusive. Many other explanatory variables will suggest themselves in particular situations, and the choice among these

[12] See Houthakker and Magee [27] and Prachowny [57] for an analysis that incorporates credit considerations.

[13] Examples of this approach are to be found in the work of Rhomberg and Boissonneault [64] for the United States and Davis [11] for Canada.

[14] Polak [56, especially pp. 47–51] has argued that the aggregate of world exports is in fact preferable to the use of aggregate world income on two grounds: (1) it makes allowance for intercountry differences in marginal import propensities; and (2) it is not affected by any general shift in the relationship between income and imports vis-à-vis the country in question. See also Prachowny [57]. Houthakker and Magee [27] have recently constructed a measure of world income that apparently gives reasonably good results.

variables will depend significantly on the intimacy of the researcher's knowledge concerning the demand phenomena he is attempting to explain. Thus, the importance of close contact with the markets being investigated cannot be overemphasized.[15]

FUNCTIONAL FORM

The general functional relations for import and export demands ignoring all but the two basic variables have already been noted in Equations (2.4) and (2.7). Thus, the import-demand function was

$$M = f\left(\frac{Y}{p_Y}, \frac{p_M}{p_Y}\right)$$

where M is the quantity of imports, Y/p_Y is real income, and p_M/p_Y is the relative price of imports. In order to fit such a relationship statistically using least squares regression, a particular functional form must be chosen. The most common forms are linear, as in Equation (2.10), and log-linear, as in Equation (2.11)

$$M = a + b\frac{Y}{p_Y} + c\frac{p_M}{p_Y} + u \qquad (2.10)$$

$$\log M = \log a_1 + b_1 \log \frac{Y}{p_Y} + c_1 \log \frac{p_M}{p_Y} + \log u \qquad (2.11)$$

which is equivalent to

$$M = a_1\left(\frac{Y}{p_Y}\right)^{b_1}\left(\frac{p_M}{p_Y}\right)^{c_1} u \qquad (2.12)$$

In Equation (2.10), a is the constant term in the regression, b is the marginal propensity to import, c is the import coefficient of relative prices, and u

[15] The following quotation from Lewis [43, p. 579] is particularly suggestive in the present context:

> Every successful drive for trade in manufactures contains five elements: keen prices, a flood of salesmen, large scale organization of selling, attention to customers' wishes and liberal credit. Econometricians usually put all the emphasis on the first element, prices, because this is easiest to measure. Businessmen, however, usually attribute much greater importance to sales effort, and it is clear that success or failure in selling is not always attributable to prices.

It will be noted in our later discussion, particularly of factors determining export-market shares, that Ginsburg [18] has attempted quantitatively to assess the importance of price and nonprice factors and found the latter to be especially important in the case of American and British exports of manufactures in the interwar and postwar periods.

is an error term reflecting other minor influences, which is assumed to be uncorrelated with the explanatory variables.[16] In the linear form the income and price elasticities of import demand will depend on the levels of these variables,[17] while in the log-linear form the income and price elasticities will be measured by the constants b_1 and c_1, which are read directly from the regression result.

There are unfortunately no clear-cut criteria that can be relied on in choosing a functional form. The researcher is more or less left to select a functional form according to his own theoretical leanings with the hope that his choice does not adversely affect his result. Some comments on the linear and log-linear forms may nevertheless be in order. One drawback of the linear form is that the price elasticity will diminish as income grows. Under such circumstances the log-linear form, which constrains the elasticities to be constant, might be preferred.[18] A characteristic of the linear and log-linear forms worth noting is that they both presume the basic demand relationship to be linear. Thus, consider Figure 2.2, in which the true underlying demand schedule *DD* is curvilinear. Assuming income effects already to have been removed, a linear or log-linear relation fitted to the data points on price and quantity would produce a demand schedule like *EE*. Statistical tests might indicate a poor fit and possibly a regression coefficient on price that was not statistically significant. The use of the linear or log-linear form might therefore be looked on as testing the significance of a particular functional form rather than the significance of the particular explanatory variable.[19]

[16] It will be noted below that if this assumption is violated, the least squares estimates of b and c will be biased.

[17] This is the same thing as saying that the price elasticity of demand will vary along a straight-line demand schedule. When the linear form is used, the relevant elasticities are often computed at the point of the sample means.

[18] From time to time some investigators have used an inverted price term (p_Y/p_M) instead of the usual one (p_M/p_Y). The coefficients and elasticities reported are therefore positive in such cases rather than negative as we generally think of them. This need not create confusion, however, since the elasticity with respect to the inverted price term is simply the negative of the usual elasticity. This can be seen as follows by considering the demand function $q = f(p)$. Letting $p^* = p^{-1}$ we would like to calculate the following demand elasticities

$$\eta = \frac{dq}{dp} \cdot \frac{p}{q} \quad \text{(usual elasticity)}$$

$$\eta^* = \frac{dq}{dp^*} \cdot \frac{p^*}{q} \quad \text{(inverted elasticity)}$$

We can write $\eta = f'p/q$, where $f' = dq/dp$. Note also that $q = f(p) = f[g(p^*)]$, where $g(p^*) = 1/p^*$. Then we can write

$$\eta^* = f'\left(-\frac{1}{p^{*2}}\right)\frac{p^*}{q} = \frac{-f'}{p^*q} = -\frac{dq}{dp} \cdot \frac{p}{q} = -\eta$$

[19] The problem of functional form is common to all econometric research and has unfortunately not been handled adequately, perhaps because it has not been thought to

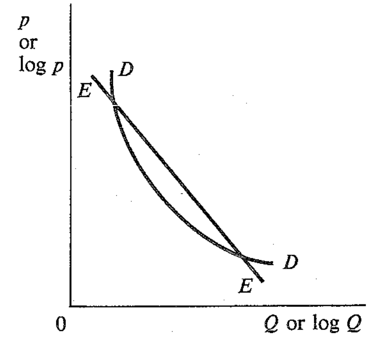


FIGURE 2.2

Linear Regression Fit to Curvilinear Demand Schedule

TIME DIMENSION

The time dimension poses quite serious problems generally in economic analysis. The econometric analysis of demand is no exception. We can begin our discussion by recalling the well-known distinction between short-run and long-run elasticities of demand. That is, whenever a demand schedule is drawn in theory, it refers to some specified period of time. In the very short run when habits are persistent, the demand schedule will be completely inelastic with respect to changes in price. The more time we permit for adjustment to price changes, the more elastic the demand schedule will be. What time period should we have in mind therefore in demand analysis?

be very important. Thus, for example, even the comparatively simple problem of choosing between the linear and log-linear forms is often decided in an ad hoc manner. The only statistically proper method currently available to make this selection is presented by Zarembka [79]. This approach however discards the niceties of least squares.

In practice, once a particular form is chosen, the investigator is often content to explore the residuals from the estimating equation for regularities that would suggest other forms. If ambiguity is revealed in the choice of function, it is often not considered important. This is surely ascientific in character.

It might be desirable therefore to introduce special variables into the function that are designed to reflect curvilinearity in the relationship. Some suggestions in this regard are given in the appendix to this chapter.

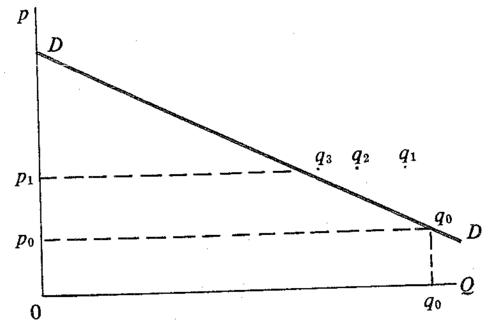


FIGURE 2.3

Adjustment of Quantity to a Change in Price

Suppose that point q_0 at price p_0 in Figure 2.3 lies on the long-run demand schedule DD, which indicates the level of demand after complete adjustment has occurred. If the price now rises from p_0 to p_1, we could imagine a sequence of decisions affecting quantity such as q_1, q_2, and q_3, where the subscripts indicate the appropriate time period. It would appear that short-run elasticities can be calculated for each of three periods of varying length. We would expect these elasticities to increase as the length of period was increased.

Let us now consider Figure 2.4, in which two successive price changes

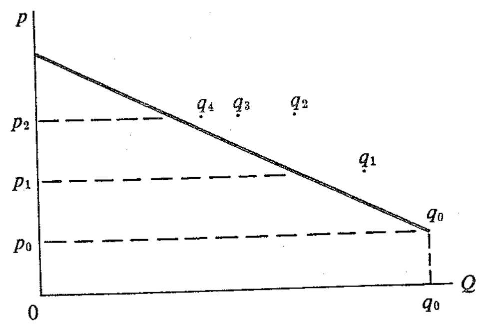


FIGURE 2.4

Adjustment of Quantity to Successive Price Changes

occur from p_0 to p_1 to p_2. What is noteworthy now is that the adjustment indicated by q_2–q_3–q_4 results from the first price change as well as the second. The q_1–q_2 adjustment appears for this reason to involve a larger elasticity than the q_0–q_1 adjustment. The point we are trying to illustrate is that the adjustment of quantity depends on the past history of price changes and the sequence of price changes within the relevant period as well as the total price change within the period. The concept of short-run elasticity is not meaningful in this context unless we can assume that the initial period's price–quantity point falls on the underlying long-run demand schedule and that the relevant price changes occur totally before any adjustment begins to take place.

Should the assumptions just mentioned be invalid, we may conclude that the concept of short-run elasticity is misleading and, regardless of the assumptions made, unnecessary. It is misleading in the sense that it implies the existence of several demand schedules that differ by the amount of time allowed for quantity to adjust, when in actuality there is only a single long-run demand schedule.[20] The concept is unnecessary insofar as the long-run demand schedule and the associated adjustment process fully describe the system.

The central question that concerns us here is the effect the time pattern of adjustment of quantity to changes in price has on the statistical estimating procedures. Thus, suppose our estimating equation is

$$M_t = f\left[\left(\frac{Y}{p_Y}\right)_t, \left(\frac{p_M}{p_Y}\right)_t, \cdots\right] \tag{2.13}$$

[20] There is one sense, however, in which time should be associated with a demand schedule. The term long-run demand schedule has been used to indicate the quantity demanded that would occur if a price change persisted indefinitely and allowance was made for complete adjustment. The quantity demanded in this equilibrium situation is a flow that will clearly depend on the length of the time period. That is, if we double the time period, we will double the quantity demanded and therefore move the long-run demand schedule proportionately to the right. This will not affect the elasticity, however. To illustrate this, let us compute the elasticity as a function of time

$$\frac{q}{t} = D(p)$$

or

$$q = tD(p)$$

where q refers to quantity, t to time, and p to price. Differentiating the foregoing expression, we have

$$\frac{dq}{dp} = tD'(p)$$

which in elasticity form is

$$\frac{dq/q}{dp/p} = \frac{pD'(p)}{D(p)}$$

which is independent of time t. Thus at any level of p, the greater slope just cancels the increased quantity, thereby maintaining the same elasticity. This conclusion is of course not startling when we recall that an elasticity is a dimensionless measure.

where each period's quantity of imports depends on the levels of the explanatory variables in the same period. Assume an adjustment process such that any period's quantity purchased is midway between the quantity in the last period and the quantity indicated by the long-run demand schedule. This is the situation in Figure 2.5, which depicts a price change from p_0 to p_1 that persists indefinitely. The dots in Figure 2.5 indicate the price–quantity points

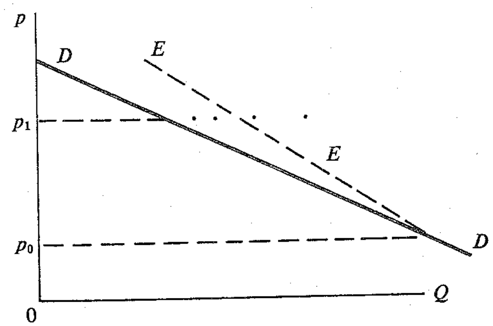


FIGURE 2.5

Linear Regression Fit to Quantities Adjusted to a Change in Price

for each period and the dashed line *EE* represents the regression estimate fitted to these points. This regression line conveys little or no information about the underlying demand schedule or the process of adjustment, since it is an unknown mixture of these two things.[21]

We must endeavor in the light of the foregoing discussion to grapple explicitly with the process of adjustment. The proper approach will depend on the length of the data period. While time lags may be unimportant when annual data are used, this will not be the case when quarterly or monthly data are used. Our discussion will thus presume data periods of relatively short length. The most straightforward procedure in dealing with the adjustment process is to add lagged explanatory variables

$$M_t = \gamma + \sum_{i=0}^{n} \alpha_i \left(\frac{Y}{p_Y}\right)_{t-i} + \sum_{i=0}^{n} \beta_i \left(\frac{p_M}{p_Y}\right)_{t-i} \tag{2.14}$$

[21] We would also get a lower elasticity estimate if we assumed p_0 to move to p_1 in a series of steps and that quantity adjusted by half the distance between last period's quantity and the quantity measured along the true long-run demand schedule *DD*.

In this formulation, an income change at time t induces a quantity change of $\alpha_i \Delta Y/p_Y$ in time period $(t+i)$. If the income change is maintained, the total or long-run effect is $(\Sigma\alpha_i)\Delta Y/p_Y$. Implicitly we have assumed, and shall continue to assume, that the nature of the response lag is fixed. Other than for this assumption, Equation (2.14) is completely general.

Unfortunately, Equation (2.14) meets with two rather severe difficulties. The economic researcher typically has a scarcity of observations and an equation such as (2.14) requires many observations because of the number of explanatory variables. More important, there is almost surely substantial collinearity among the various lagged explanatory variables. The result of such "multicollinearity" is well known: large standard errors on the coefficients.

Some investigators have approached this problem by including lagged values until they assumed the incorrect sign or had statistically insignificant coefficients. Such a procedure is improper, however, for it is sure to exclude lagged values that should properly be included, to overestimate the coefficients on the more recent terms, and furthermore to yield misleading statistical tests. Alternatively, it might be observed that if we are interested in the long-run response and not the distribution of the response over time, then large standard errors on individual coefficients are unimportant. We should be interested instead in standard errors on $\Sigma\hat{\alpha}_i$ and $\Sigma\hat{\beta}_i$ which would generally be smaller because of the negative covariances between the individual estimators associated with the multicollinearity of the lagged variables.

The two procedures just mentioned represent attempts to work within the framework of Equation (2.14). However, given the generality involved in that equation, it can be argued that it does not adequately represent our theoretical knowledge of the lag structure. Rather, if we restrict the values which the α_i and β_i can assume, smaller standard errors will be obtained. The simplest procedure in this regard is to replace the terms in (2.14) with weighted averages such as

$$\alpha_0 \left(\frac{Y}{p_Y}\right)_t + \alpha_1 \left(\frac{Y}{p_Y}\right)_{t-1} + \alpha_2 \frac{(Y/p_Y)_{t-2} + \cdots + (Y/p_Y)_{t-5}}{4}$$

This embodies the idea that most of the response occurs in the first two periods, with the remainder being so small that the last four periods can be treated identically.

A currently very popular procedure is to describe the demand phenomenon in "stock-adjustment" terms.[22] Thus, the long-run demand is

$$M_t^* = a + b\left(\frac{Y}{p_Y}\right)_t + c\left(\frac{p_M}{p_Y}\right)_t + \epsilon_t \tag{2.15}$$

[22] See Nerlove [52] for further discussion.

The current demand adjusts only δ 100% to the long-run level, with $0 \leq \delta \leq 1$:

$$M_t = M_{t-1} + \delta(M_t^* - M_{t-1}) = \delta M_t^* + (1 - \delta)M_{t-1} \tag{2.16}$$

Substituting (2.15) into (2.16) we have

$$M_t = \delta a + \delta b\left(\frac{Y}{p_Y}\right)_t + \delta c\left(\frac{p_M}{p_Y}\right)_t + (1 - \delta)M_{t-1} + \delta\epsilon_t \tag{2.17}$$

The constant δ is to be interpreted as a coefficient of adjustment and should fall between zero and one. If δ is equal or close to zero, the adjustment process is very slow as is evident from Equation (2.16). Rapid adjustment is associated with values of δ close to one.

When Equation (2.17) is used as a regression, the long-run coefficients a, b, and c are easily calculated by dividing the regression coefficient by one minus the coefficient of M_{t-1}. As argued earlier, the results are most meaningfully reported in terms of the long-run coefficients and δ, the coefficient of adjustment. Some researchers, such as Rhomberg and Boissonneault [64], report the unadjusted or impact elasticities on the grounds that the assumptions underlying Equation (2.16) are untested. However, it should be noted that if these assumptions are improper, the impact as well as the long-run elasticities will be incorrectly estimated.

Equation (2.17) can be derived in still another way that is illuminating. Suppose

$$M_t = \sum_{i=0}^{\infty} \gamma_i f\left[\left(\frac{Y}{p_Y}\right)_{t-i}, \left(\frac{p_M}{p_Y}\right)_{t-i}\right] + \epsilon_t \tag{2.18}$$

with the long-run demand

$$M_t^* = (\Sigma\gamma_i) f\left[\left(\frac{Y}{p_Y}\right), \left(\frac{p_M}{p_Y}\right)\right]$$

and an adjustment γ_i in the $(t+i)$th period.

Assume the γ_i to decline geometrically

$$\gamma_{i+n} = (1 - \delta)^n \gamma_i \tag{2.19}$$

Thus

$$\begin{aligned}(1 - \delta)M_{t-1} &= \sum_{i=0}^{\infty} (1 - \delta)\gamma_i f\left[\left(\frac{Y}{p_Y}\right)_{t-i-1}, \left(\frac{p_M}{p_Y}\right)_{t-i-1}\right] + (1 - \delta)\epsilon_{t-1} \\ &= \sum_{i=0}^{\infty} \gamma_{i+1} f\left[\left(\frac{Y}{p_Y}\right)_{t-i-1}, \left(\frac{p_M}{p_Y}\right)_{t-i-1}\right] + (1 - \delta)\epsilon_{t-1} \\ &= \sum_{i=1}^{\infty} \gamma_i f\left[\left(\frac{Y}{p_Y}\right)_{t-i}, \left(\frac{p_M}{p_Y}\right)_{t-i}\right] + (1 - \delta)\epsilon_{t-1} \\ &= M_t - \gamma_0 f\left[\left(\frac{Y}{p_Y}\right)_t, \left(\frac{p_M}{p_Y}\right)_t\right] - \epsilon_t + (1 - \delta)\epsilon_{t-1}\end{aligned} \tag{2.20}$$

and

$$M_t = \gamma_0 f\left[\left(\frac{Y}{p_Y}\right)_t, \left(\frac{p_M}{p_Y}\right)_t\right] + (1 - \delta)M_{t-1} + \epsilon_t - (1 - \delta)\epsilon_{t-1} \tag{2.21}$$

which except for the error term is the same as Equation (2.17).

This derivation illustrates three points. Evidently from Equation (2.18) the nature of the adjustment lag is the same for both variables. This is a disturbing assumption for it may well be that adjustment to income changes will be more rapid than adjustment to price changes. Thus, for example, information concerning price changes is disseminated relatively slowly and the alteration of buying habits may be quite slow. In contrast, the awareness of an income change must occur at receipt and adjustment to new income levels may be rapid, perhaps even anticipatory.

The second observation is that the adjustment has been assumed to decay in a geometric fashion given by Equation (2.19). This is a very restrictive form, which is open to criticism. It seems reasonable that adjustment to a price change would build up slowly rather than decay. A price change is probably met at first by a refusal to alter buying habits on the grounds that the price change may be transitory. Only after the permanence of the change has been accepted will adjustment begin. When the perfect substitutability description of demand is used and import demand is the difference between home demand and home supply, adjustment to price changes will occur partly on the supply side, which will require the transfer of resources between sectors of the economy. Geometric decay is most unlikely to be an adequate description of such a supply response.

Finally, we should notice the difference in the error terms in Equations (2.17) and (2.21). The error $\delta\epsilon_t$ in (2.17) has all the classical properties except that it is correlated with succeeding values of the explanatory variable M_{t-1}: $M_t, M_{t+1}, \ldots$. This will cause bias in the estimates, although the desirable properties of unbiasedness and minimum variance are preserved for large samples. Somewhat worse is that the error $\epsilon_t - (1 - \delta)\epsilon_{t-1}$ in Equation (2.21) is correlated with earlier error terms, with later values of the explanatory variable M_{t-1}, and with the explanatory variable M_{t-1} as well. In such a situation, the estimators will not retain even their large sample properties. More complicated estimating procedures have been suggested to deal with this situation, as in Johnston [29, pp. 211–12]. The question as to whether the stock-adjustment or the geometrically decaying lag description is preferable is best left open. It should be emphasized, however, that the use of lagged dependent variables as explanatory variables is an extremely tricky business that requires knowledgeable exploitation in order to be effective.

After the discussion of the question of lags in terms of stock adjustment and geometric decay, it may be useful to consider the somewhat more general approaches attributable to Jorgenson [31] and Almon [3]. According to

Jorgenson, if we include lagged values of the explanatory variables and additional lagged values of the dependent variables, we can produce a very general lagged response. Thus, for example

$$M_t = a + b_0\left(\frac{Y}{p_Y}\right)_t + b_1\left(\frac{Y}{p_Y}\right)_{t-1} + c_0\left(\frac{p_M}{p_Y}\right)_t + c_1\left(\frac{p_M}{p_Y}\right)_{t-1} + \delta M_{t-1} \quad (2.22)$$

allows the current coefficients to be completely free, while all lagged responses form a geometrically declining series governed by the adjustment coefficient δ. Clearly both the multicollinearity and the degrees-of-freedom pinch have been mitigated. Unfortunately, the problems concerning the error term and the associated statistical properties of the estimators discussed above still apply.

The procedure suggested by Almon [3] is to observe that the perfectly general form

$$M_t = \gamma + \sum_{i=0} \alpha_i\left(\frac{Y}{p_Y}\right)_{t-i} + \sum_{i=0} \beta_i\left(\frac{p_M}{p_Y}\right)_{t-i} \quad (2.23)$$

allows the coefficients an amount of freedom that is quite unjustified. For instance the α_i can assume a sawtooth of values as in Figure 2.6. But if this

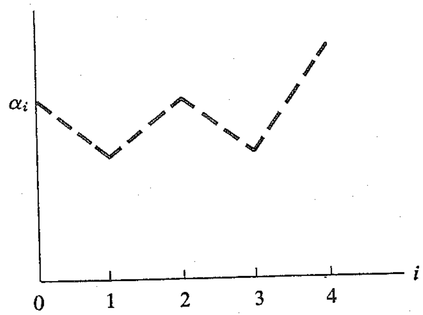


FIGURE 2.6

Unacceptable Configuration of α_i

is not believed to be possible, we ought to restrict the values of α_i. If we use this extra information, we will improve our estimates of the α_i's.

Suppose, for instance, that we expect the α_i to look something like Figure 2.7. We observe that such a configuration may be well approximated by a quadratic form: [23]

$$\alpha_i = a + bi + ci^2 \quad (2.24)$$

[23] This derivation of the Almon technique is based on remarks made by D. B. Suits.

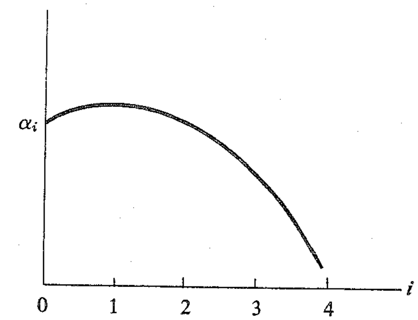


FIGURE 2.7

Alternative Configuration of α_i

We observe further that assumption (2.24) is not particularly restrictive. All it says is that α_i must be relatively close to α_{i-1} and α_{i+1}, and that if there is a hump (or a trough, for that matter) in the distribution of the α_i, it is the only one. Also, the response after five periods is negligible.

We can now use assumption (2.24) to express Equation (2.25)

$$\begin{aligned}\sum_{i=0}^{4} \alpha_i\left(\frac{Y}{p_Y}\right)_{t-i} &= a\left(\frac{Y}{p_Y}\right)_t + (a+b+c)\left(\frac{Y}{p_Y}\right)_{t-1} \\ &\quad + (a+2b+4c)\left(\frac{Y}{p_Y}\right)_{t-2} + \cdots \\ &\quad + (a+4b+16c)\left(\frac{Y}{p_Y}\right)_{t-4} \qquad (2.25) \\ &= a\left[\sum_{i=0}^{4}\left(\frac{Y}{p_Y}\right)_{t-i}\right] + b\left[\sum_{i=0}^{4} i\left(\frac{Y}{p_Y}\right)_{t-i}\right] \\ &\quad + c\left[\sum_{i=0}^{4} i^2\left(\frac{Y}{p_Y}\right)_{t-i}\right]\end{aligned}$$

We see, therefore, that by using a very plausible assumption we have reduced the number of explanatory variables from $n+1$ to three. The number of variables can be further reduced by adding other assumptions such as $\alpha_4 = 0$. Alternatively, a larger number of variables will be required if a higher-degree polynomial for α_i is used. Precisely the same procedure can, of course, be used for the β_i.

In theory this Almon technique is the most promising method for attacking the problem of lagged responses. The researcher is free to choose from a wide variety of polynomials to restrict the α_i. Even when such an assumption is made, the distribution of the α_i is usually quite free. Finally, the error term remains agreeable to ordinary least squares. However, experimentation with the Almon technique is required before its efficacy can be established in the kinds of analysis we have been discussing.

One final approach worth mentioning is to express the variables in terms of first differences

$$\Delta M = f\left[\Delta\left(\frac{Y}{p_Y}\right), \Delta\left(\frac{p_M}{p_Y}\right)\right] \tag{2.26}$$

where

$$\Delta M = M_t - M_{t-1}, \Delta\left(\frac{Y}{p_Y}\right) = \left(\frac{Y}{p_Y}\right)_t - \left(\frac{Y}{p_Y}\right)_{t-1}, \ldots$$

The use of first differences is often recommended when it is desired to reduce the effects of serial correlation.[24] First differences do not constitute a solution to the adjustment problem, however, since no allowance is made for the adjustment of quantity in relation to the changes that have occurred historically in income and prices. The first-difference equation form is consequently no different from the naive form in which no explicit allowance is made for the process of adjustment.

In our earlier discussion of the choice of explanatory variables, extensive familiarization with the details of the relationships under study was urged in order that a judicious selection of explanatory variables could be made. The same point applies to the selection of adjustment processes.

SPECIAL PROBLEMS IN ESTIMATION

There appeared during the 1940's a number of studies in which statistical estimates, using ordinary least squares regression methods, were made of the income and price elasticities of demand for the imports and exports of individual countries during the interwar period.[25] The price elasticities in particular were estimated to be substantially less than unity. This suggested in the context of exchange-rate adjustment that a devaluation would tend to worsen rather than improve the trade balance because the sum of the elasticities of demand for an individual country's imports and exports might together

[24] If such effects are believed to be serious, it may be desirable to employ some type of adjustment scheme. Some success with the first-order Cochran-Orcutt iterative technique is reported by Houthakker and Magee [27].

[25] See Cheng [8] for a description of these studies.

add up to less than unity.[26] This "elasticity pessimism" suggested that measures other than changes in relative prices might have to be relied upon for purposes of adjusting the balance of trade.

In a pathbreaking article published in 1950, Orcutt [53] sought to demonstrate that the statistical results obtained in the aforementioned studies were subject to very serious reservations because the method and data employed tended to bias the calculated price elasticities towards zero. He gave five reasons why this might be the case:

(1) Lack of independence between relative prices and the random deviation in the import-demand function.

(2) The data may reflect errors of observation.

(3) The use of data aggregates may give undue weight to goods with relatively low elasticities.

(4) Short-run elasticities were measured and these are typically lower than long-run elasticities.

(5) Devaluation elasticities were larger than the estimated short-period elasticities, which reflect adjustment to small price changes.

It is worthwhile to discuss each of these points in turn. In doing so, we shall find it convenient to use the linear form of the regression equation for import demand noted earlier in Equation (2.10):[27]

$$M = a + b\left(\frac{Y}{p_Y}\right) + c\left(\frac{p_M}{p_Y}\right) + u$$

where M is the quantity of imports, Y/p_Y is real income, p_M/p_Y is the relative price of imports, and u is a random deviation associated perhaps with variables that have been inadvertently excluded. According to the theory underlying ordinary least squares regression, the estimate of the relative price coefficient c will be unbiased only if the random deviation u is independent of p_M/p_Y.

Since income and relative prices tend to move together, the estimate of the price coefficient c will be biased if real income is excluded from the relationship. In this case the random disturbance includes the excluded income term and would necessarily be correlated with the only explanatory variable, relative prices. The problem is not solved, however, by including both real income and relative prices in the relationship because the relative price variable and the random variable u may continue to move together rather than independently. This can be seen with the aid of Figure 2.8.

Thus suppose that we have a demand DD and supply schedule SS as in Figure 2.8, in which the quantities indicated are net of what can be explained

[26] That is, if we assume infinite supply elasticities, the Marshall–Lerner condition that is necessary for a devaluation to improve the trade balance would not be satisfied.

[27] The remarks to be made apply equally to the log-linear form or any other function employed.

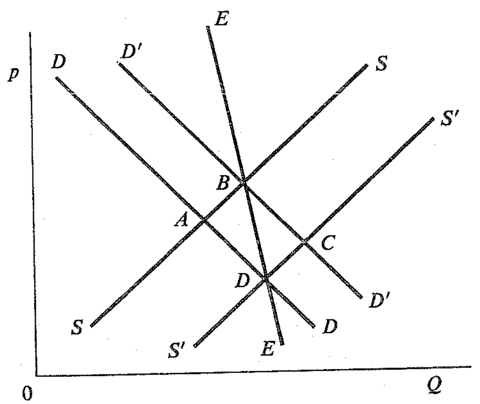


FIGURE 2.8

Downward Bias in the Estimated Price Coefficient †

† Adapted from G. H. Orcutt, "Measurement of Price Elasticities in International Trade," *Review of Economics and Statistics*, XXXII (May 1950), 123.

linearly by variations in income. Now if we assume that there is a random disturbance that results in a shift of the demand schedule up to the right to *D′D′*, there will be an increase in price with *SS* unchanged. What this means therefore is that high values of the random variable u are associated with high values of price. This violates the requirement that u be independent of p_M/p_Y, with the result that the estimate of the price coefficient c will be between the true negative elasticity of demand and the positive elasticity of supply.[28]

The bias in the estimate can be illustrated by supposing that there are random disturbances that cause the demand schedule to shift up and down between *DD* and *D′D′*, and similarly that there are random disturbances that cause the supply schedule to shift up and down between *SS* and *S′S′*. This will yield the parallelogram *ABCD*, within which the data points for prices and quantities will be confined. If we were now to fit a regression line through these points that would minimize the deviations in a horizontal direction with respect to the dependent variable, quantity demanded, we would get a line such as *EE*. It will be evident that the elasticity calculated from *EE* will be underestimated in comparison to the true elasticity on the underlying demand schedule.

[28] For a rigorous demonstration of the downward bias in the price as well as in the income coefficients, see Orcutt [53].

A bit of experimenting with Figure 2.8 will show that if the shifts of the demand schedule were large relative to those of the supply schedule, the regression line fitted would have a positive slope approximating that of the supply schedule. By the same token, if the shifts of the supply schedule were large relative to those of the demand schedule, a regression line closer to the demand schedule would be obtained. In either case the estimate of the price coefficient c would be less than the true one. If, however, the elasticity of supply can be taken as infinite, the estimated and true coefficients will coincide. This situation is depicted in Figure 2.9.

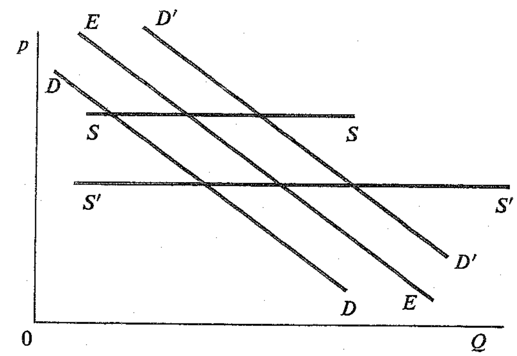


FIGURE 2.9

Unbiased Estimate of Price Coefficient

This discussion suggests that the use of ordinary least squares regression may be appropriate when the shifts in the supply schedule are large relative to those of the demand schedule, and/or when the supply schedule is highly elastic. In employing ordinary least squares, care should be taken therefore of the particular economic conditions affecting the relationship. Thus, in the case of a small country that imports only a relatively small fraction of total world exports, it may be quite realistic to assume an infinitely elastic supply schedule. In contrast, a country like the United States may face a rising supply schedule because of its relatively large size.[29]

[29] An obvious way to deal with the problem of biased estimates is to use simultaneous equation estimating techniques. The problem here, however, is that we have very little knowledge of supply relations. The little experimentation that has been done along these lines—see especially Morgan and Corlett [49]—has not been particularly successful. Alternatively, one can seek to approximate the shifts in the schedules as Harberger [24] has done, although it is by no means clear that this can be accomplished in a nonarbitrary manner.

Orcutt's second point was that when the data contain errors of measurement due to misclassification, falsification, and faulty methods of index-number construction, the effect may be to bias the coefficients toward zero. This point is illustrated in Figures 2.10 and 2.11, in which the underlying demand *DD* has no error term associated with it. We may contrast Figure 2.10,

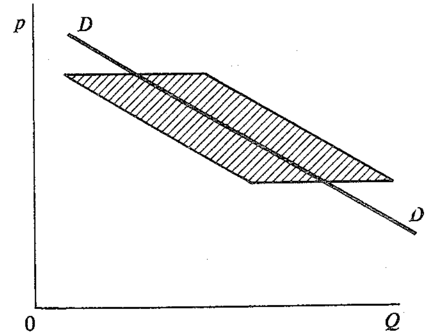


FIGURE 2.10

Errors of Observation in Quantity †

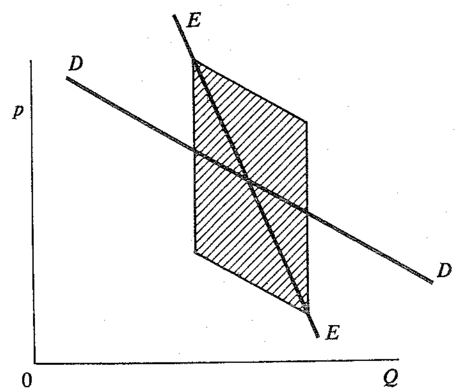


FIGURE 2.11

Errors of Observation in Price ‡

† Adapted from G. H. Orcutt, "Measurement of Price Elasticities in International Trade," *Review of Economics and Statistics*, XXXII (May 1950), 124.

‡ Adapted from G. H. Orcutt, "Measurement of Price Elasticities in International Trade," *Review of Economics and Statistics*, XXXII (May 1950), 124.

in which it is assumed that there are errors of observation in quantity and not in price, with Figure 2.11, in which the errors are in price but not in quantity. In Figure 2.10 the true demand schedule would be measured by the regression line. This would not be the case, however, with Figure 2.11, in which the regression line implies a lower elasticity than the true underlying demand schedule. Observation errors in the explanatory variables are thus seen to produce biased estimates. The importance of this point is essentially an empirical question and is therefore hard to assess. Moreover, we have implicitly assumed a certain relationship between the observation error and the true value. Other assumptions about this relationship will lead to other conclusions.[30]

In his third point Orcutt argued that goods with relatively low elasticities may exhibit the greatest price variation and will therefore exert a predominant effect on the aggregative price indexes. The use of such aggregative indexes may thus understate the true elasticity to the extent that goods with lower elasticities are given undue weight.[31] It would thus appear on these grounds that a strong argument can be made against using aggregative

[30] Suppose that the true description of demand is

$$Q = \alpha + \beta p + \epsilon$$

where ϵ is the usual error. Suppose further that we do not observe p but rather Π, which is p plus an independent error term u

$$\Pi = p + u$$

A regression of Q on Π

$$Q = \alpha + \beta\Pi + (\epsilon - \beta u)$$

will yield a biased estimate of β because of the correlation of the independent variable Π and the error term $(\epsilon - \beta u)$. This is clear, since Π is the sum of the true value p plus an error u, and must be correlated with that error u.

If, on the other hand, the measurement error is independent of the observed value we can write

$$p = \Pi + u$$

to express the fact that the observed value Π and the measurement error u are independently generated and add to p, the true value. A regression of Q on Π in this case

$$Q = \alpha + \beta\Pi + (\epsilon + \beta u)$$

will yield an unbiased estimate of β, since Π and $(\epsilon + \beta u)$ are independent.

Observation error is typically described by the first model and is therefore associated with biased coefficients. Although the second model casts some doubt on such statements, it does not seem to make much sense. In Prais's [58, p. 576] words, however, ". . . it is not easy to know which of these hypotheses is closer to the truth, but the case for supposing a downward bias to be inevitable is clearly not as strong as at one time thought." Johnston [29, Chap. 6] provides an extensive discussion of the treatment of errors in variables.

[31] This point is demonstrated in the appendix to this chapter, in which we indicate that the price indexes should be reweighted proportionally to the individual demand elasticities.

indexes. This may be reinforced more generally by the fact that a price index, being a weighted average, tends to show less variation than any of its components insofar as price increases are offset against price declines. As a consequence, there will tend to be an increase in the estimated standard error. An aggregative price index is likely, moreover, to be highly correlated with income.

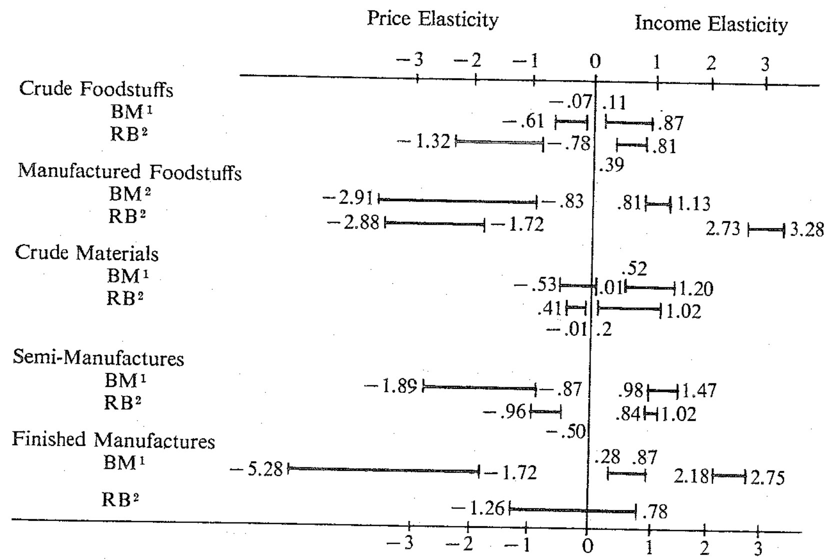
The foregoing points suggest that disaggregation may be desirable, although whether or not the results will be improved significantly is an empirical question. What little evidence is available—see Ball and Marwah [4], DaCosta [10], and Dutta [13]—indicates that the returns to the use of fine subcategories of data may be limited.

Orcutt's fourth point related to the fact that what was calculated in most studies was a short-run elasticity that would be expected to be lower than the long-run elasticity. We shall not dwell on this point here since it was already argued in the preceding section dealing with the time dimension that the concept of the short-run elasticity is not particularly meaningful, and, further, that ignoring the time dimension in the analysis would bias downward the estimate of the long-run price elasticity.

Orcutt's final point was that the price elasticity of demand for large price changes will generally be higher than for small price changes. His reasons were that it takes time for habits to adjust and that the price changes must be large enough to overcome the costs of switching. This point is in direct opposition either to our assumption of a long-run demand function that depends only on current prices and not on the history of prices, or to our assumption that the nature of the response lag is fixed independent of the type of price movement that is occurring. Since in our judgment the assumption of an underlying long-run demand relationship seems absolutely basic, we assume Orcutt's point is really that adjustment to large price changes is more rapid than adjustment to small changes. This will be especially true in the case of a devaluation when the price changes are clearly going to be permanent and there will be no adjustment delay in anticipation of a reversal of the price change. However, if this is the case, we will underestimate the speed of adjustment but not the total adjustment. The long-run elasticity will be properly estimated.

The upshot of Orcutt's five points was to cast grave doubts on the usefulness of least squares procedures for the time-series analysis of demand. It led Neisser [51, p. 130], for example, to declare in 1958 that: "The traditional multiple regression analysis of time series . . . is dead." This conclusion was overly pessimistic, however, since Orcutt's reservations about least squares procedure are not quite as devastating as they may appear. That is, there may be many cases in which this procedure is reasonably applicable. This will be true when countries are relatively small and also when demand is relatively stable. It is also possible to use data specifications that will avoid lumping

together commodities with widely varying elasticities. Explicit allowance can be made, moreover, for lags in the adjustment process. Finally, it may well be that the interwar period had special characteristics that made for unreliable statistical results. The postwar period has in contrast been much more stable, with the result that the effects of changes in income and prices may be more readily isolated. These conclusions seem to be borne out, it may be noted, in a number of studies using postwar data. Typical results from two of these studies are given in confidence-interval form in Figure 2.12.[32]



¹ BM refers to Ball and Marwah [4].
² RB refers to Rhomberg and Boissonneault [64].
† Based on Harley [25, p. 19]; original data from R. J. Ball and K. Marwah, "The U. S. Demand for Imports, 1948–1958," *Review of Economics and Statistics*, XLIV (November 1962), and from R. R. Rhomberg and L. Boissonneault, "The Foreign Sector," in J. S. Duesenberry et al., eds., *The Brookings Quarterly Econometric Model of the United States*. Chicago: Rand McNally & Company, 1965.

FIGURE 2.12

Confidence Intervals for Income and Price Elasticities for U.S. Import Demand †

[32] The evident differences in results noted may be due largely to differences in the time periods and data classifications employed in the two studies cited. Results for fifteen major industrialized countries and twelve developing countries for 1951–66 are to be found in Houthakker and Magee [27]. Alternative formats for presentation of data are treated in the next section, which deals with the reporting of results.

In closing, two additional matters deserve comment. These concern the presence of multicollinearity in the explanatory variables and the types of statistical tests performed in conjunction with least squares regression. When two explanatory variables are statistically correlated, least squares regression divides the explanatory power between them and both may take on relatively large standard errors. When one variable is removed from the regression, the other will gain in significance but the coefficient becomes biased. A situation of this kind is due largely to data inadequacy, and it is very difficult to overcome since explanatory variables respond jointly to similar underlying forces, as for example when income and prices tend to move together over time. Multicollinearity thus will plague most time-series analysis of demand relations and there is often not much that can be done to cope with it. One possibility worth mentioning, however, is to employ extraneous estimates—see Kaliski [33]—which amounts to fixing a particular coefficient at a level determined by exterior knowledge, obtained perhaps from a cross-section estimate.

Finally, some caution appears necessary in placing excessive faith in the statistical tests employed in least squares regression. It is of course quite important to attach levels of confidence to any estimates that are made, provided that the assumptions upon which the tests are predicated are reasonably good approximations.[33] But it should be noted that the final results in any study often represent the best of many experiments that are all based upon the same data. This means that out of the many experiments that may have been done, some of the significant results may only be random. This might be taken into account when specifying the width of confidence intervals.

REPORTING RESULTS

Two points that have been mentioned just briefly are worthy of additional comment. These involve the reporting of experimentation and the use of significance tests.

It is altogether too common for researchers to report only their best results without indicating the trial-and-error process by which they were obtained. It is statistically improper in principle to experiment freely, select the best fit, and report confidence intervals and/or significance tests. When none of the experimentation is reported, it becomes very difficult to assess the quality of the research effort in terms of its approach to the many important

[33] For example, standard errors will not be reliable when there is serial correlation in the residuals of the regression.

methodological issues we have discussed earlier. It is also the case that one researcher's experimental failures are of considerable importance in the design of future research by others.

As far as significance tests are concerned, the economist's general obsession with them leads one to the belief that often sight has been lost of just exactly what significance means. A thorough discussion of this subject is well outside the scope of the present undertaking, but it is hoped that the brief discussion below will serve to clarify the issues involved.

We will first consider the question of inference within the classical statistical framework. When, for example, a price elasticity is reported to be statistically significant at the 95 percent level, the following statement is being made: "A procedure has been followed such that if the true value of the price elasticity were zero, 95 percent of the time we would say the value is not significantly different from zero, while the other 5 percent of the time we would say that it *is* significantly different from zero, that is, we would be making a mistake." This is all that is being asserted. For instance, the statement says nothing about the situation when the price elasticity is different from zero. Nor does it allow us to accept any useful hypothesis; thus a rejection of the value zero does not imply acceptance of some value different from zero, nor does failure to reject zero imply acceptance of zero.

The obvious weakness in significance tests has led enlightened researchers to report confidence intervals. They have argued that there are really two distinct reasons why an estimate is deemed not significant. In the one case, standard errors are small but the point estimate is near zero. In the other case, standard errors are large and the point estimate may be anywhere. The general feeling is that in the first case the value of the elasticity is truly near zero while in the second case its value is unknown. This same idea is sometimes expressed by pointing out that "not significant" does not mean "insignificant." In other words, "not significant" does not mean "near zero." On the contrary, "not significant" means that the data have been unable to answer the question that has been posed. This is not to say that the data can answer no questions. It is up to the researcher to ask other questions and thereby get "significant" answers. Confidence intervals are meant to contain all the answers the data will yield. It is well known that the set of points exterior to the confidence interval includes all those points that would be rejected by a significance test.

At this juncture we should ask just what a confidence interval means. The appropriate description is: "A procedure has been followed to generate a set of points that 95 percent of the time will include the true value of the elasticity." This should not be taken to mean that the true value is a member of the set of points; nor should we necessarily believe that it is. For instance, the interpretation that a short interval "nails down" the estimate is inappropriate in this setting. Such a statement requires that we turn the above statement around to "the probability that the true value lies in the interval is

0.95." To the classical statistician the probability that the true value lies in the interval is zero or one. In other words, the true value is a parameter and does not have a probability distribution.[34]

Economic researchers generally do not behave in this classical manner, as they are more than willing to bestow the reward of belief on their confidence intervals. This puts them, perhaps unknowingly, into the Bayesian school of inference. The Bayesian has a probability distribution of the true parameter and does make the statement that the probability is 0.95 that the true value lies in the confidence (credence, to him) interval. Perhaps, more important, he can make statements such as, "The probability that the price elasticity exceeds one is 0.85." This is the kind of information, for example, that is needed by the policymaker who must choose to devalue or not.

Quite apart from one's statistical school, the central issue in reporting results is how to convey the information contained in the study as completely and concisely as possible. Confidence intervals are clearly the proper vehicle. To expand on this let us consider five possible formats for reporting on a simplistic aggregate import function for the U.S.

The formats are listed in the increasing order of our preference. Thus Format 2 conveys much more information than Format 1. Although confidence intervals could be calculated from the information in Format 2, Format 3 saves the reader the trouble. The summary statistics of Format 2 can easily be added to 3, 4, and 5 if necessary. Format 4 is preferred to Format 3 only because it is more readable. Although Format 5 may be prohibitively space-consuming, it contains an additional but quite important bit of information. A high price elasticity can be traded for a low income elasticity because of the correlation between the two. This is evident from the slanted confidence area.[35]

34 If this discussion is unconvincing, consider the following situation. Two researchers take independent samples from the same process. Each provides us with a confidence interval for the value of the price elasticity. We must choose between these two intervals and are told only that one is longer than the other. Which should we choose? If you prefer the shorter one, then you have improperly turned the confidence statement around. It is easy to show that the longer interval is more likely to fall over the true value than the shorter one (from independence of the position and length of the interval). Accordingly, one should be quite indifferent between the two intervals. The length of the interval in this example should not be associated with the quality of the information afforded by the interval.

35 Goldberger [20, p. 175] shows that an α-confidence set for a subset of the coefficients, denoted by vector β_2, is given by

$$\frac{(b_2 - \beta_2)' D (b_2 - \beta_2)}{(K - H)s^2} \leq F^{K-H}_{T-K-1}(\alpha)$$

where T is the number of observations; K is the number of explanatory variables; H is the number of variables in β_2; b_2 is the estimated value of β_2; D is the inverse of the appropriate submatrix of $(X'X)^{-1}$; s^2 is the estimated standard error; and $F(\alpha)$ is the value of the Snedecor F distribution at level α.

FORMAT 1

Price Elasticity	−0.77*
Income Elasticity	1.27*

* Significant at the 95 percent level.

FORMAT 2

$$\log M = 15.2 - \underset{(.13)}{.77} \log \frac{p_M}{p_Y} + \underset{(.08)}{1.27} \log \frac{Y}{p_Y}$$

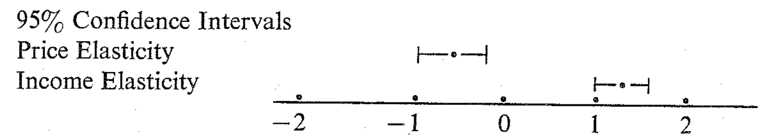
d.w. = .56 $\quad \sigma = .04 \quad R^2 = .997$

FORMAT 3

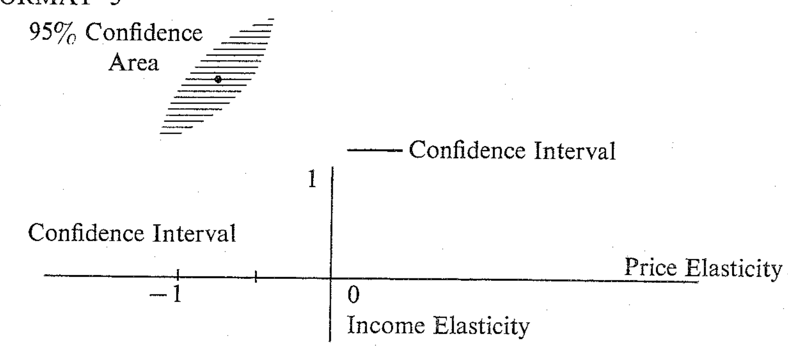
95% Confidence Intervals

	Lower Limit	*Point Estimate*	*Upper Limit*
Price Elasticity	−0.99	−0.77	−0.55
Income Elasticity	1.14	1.27	1.40

FORMAT 4



FORMAT 5



In the example in the text, β_2 has two components, D is a two-by-two matrix, and the formula above describes an ellipse centered at the estimates of the elasticities. The slope of the major axis of the ellipse is given by θ, where

$$\tan \theta = \frac{2d_{12}}{(d_{11} - d_{22}) - \sqrt{(d_{11} - d_{22})^2 + 4d_{12}^2}}$$

and d_{ij} are the elements of D. The denominator is necessarily negative; thus the slope of the major axis will be positive when d_{12} is negative, and the converse. In our example, d_{12} is simply the covariance between the price and income terms, and is negative. The slope of the confidence region is consequently positive. If the price and income terms are positively correlated, the confidence region would slope the other way.

There are two reasons why the confidence-region information is important. Low price elasticities are often attributed to the collinearity between the price and income terms. This statement is supported by the slope of the confidence region. Secondly, inferences and decisions such as a devaluation will involve the price elasticity and income elasticity jointly. Such a decision will be incorrectly made if it is based on the two intervals separately rather than on the confidence region. The 95 percent confidence region implied by the two intervals is a box bordered by the interval limits. This is evidently quite different from the true confidence ellipse in Format 5. If the researcher prefers one of the other formats, he should by all means report the covariance terms between the estimates. This will allow the reader to perform any sort of test he might desire.

CONCLUSION

In conclusion, we may briefly review the salient points of this chapter. From a theoretical point of view, import (export) demand should be distinguished according to (1) consumer demand and (2) producer demand. In addition, these may both be subdivided into durables and nondurables. Each of these categories represents a unique economic phenomenon that requires a unique set of explanatory variables. We may disaggregate further within each of these groups according to the nature of domestic substitutes, with a special treatment when there is an almost perfect substitute available domestically.

In addition to deciding on the level and nature of the disaggregation and the proper explanatory variables, the researcher will have to select a functional form and a method for handling response lags. The choice of functional form is largely arbitrary. Current thinking on response lags tends to favor the Almon approach.

Finally, the researcher must decide how best to convey the results of his efforts to the reader. We have argued that any and all experimentation should be reported, and that the statistical evidence is best summarized in confidence intervals and regions.

We have seen that a long list of indictments of the traditional least squares method drawn up by Orcutt caused considerable skepticism during the 1950's concerning the reliance on this method. Since then, however, reconsideration of Orcutt's points has suggested that they were not as damaging to the use of least squares procedures as was initially believed. Thus, used with proper care, these procedures may yield valid and meaningful measures of the income, relative price, and other influences on the demand for a country's imports (exports).

APPENDIX TO CHAPTER 2

INDEX NUMBERS AND THE PROBLEM OF AGGREGATION

The validity and relevance of all that has been discussed in Chapter 2 rest fundamentally on the existence of the macroeconomic relationship (2.4) describing the demand for a quantity of imports as a function of a price and an aggregate income term. We have mentioned in the text that the quantity and price variables are not raw observations but rather index numbers made necessary by the aggregation over commodities. In this appendix we will be interested in defining precisely the index numbers and discussing the effects of aggregation, with an aim toward understanding the occasions when aggregation should be avoided and the procedures by which its bad effects might be alleviated.

Index Numbers The point of an index number is to reduce a large set of data describing a complex economic event into a single number that somehow captures the essential features of that event. Thus, for example, we would like to express a set of price observations p_{io} and p_{it}, $(i = 1, \ldots, n)$ in a single number or index, which is able to indicate whether prices have generally risen or fallen from year o to year t. An apparently good candidate for such a number is a weighted average of the price relatives

$$P_{to} = \sum_i w_i \frac{p_{it}}{p_{io}}, \qquad w_i \geq 0, \qquad \sum_i w_i = 1 \tag{2.A.1}$$

The index P_{to} has the property that if all prices fall, remain the same, or rise, P_{to} will be less than, equal to, or greater than one. We might hope therefore that the index P_{to} captures the average movement of prices. We have not as yet selected the appropriate values of the weights and might do so by observing that for the ith commodity

$$\frac{V_{it}}{V_{io}} = \frac{p_{it}}{p_{io}} \times \frac{q_{it}}{q_{io}} \tag{2.A.2}$$

That is, the increase in the value from year o to year t has been separated into the price change p_{it}/p_{io} and the quantity change q_{it}/q_{io}. The same thing can be done for the aggregates

$$V_{to} = \frac{V_t}{V_o} = \frac{\sum_i V_{it}}{\sum_i V_{io}} = \frac{\Sigma p_{it} q_{io}}{\Sigma p_{io} q_{io}} \times \frac{\Sigma p_{it} q_{it}}{\Sigma p_{it} q_{io}}$$
$$= \left(\Sigma w_{io} \frac{p_{it}}{p_{io}}\right) \times \left(\Sigma w_{it} \frac{q_{io}}{q_{it}}\right)^{-1} \tag{2.A.3}$$

where

$$w_{io} = \frac{p_{io} q_{io}}{\Sigma p_{io} q_{io}} \quad \text{and} \quad w_{it} = \frac{p_{it} q_{it}}{\Sigma p_{it} q_{it}}$$

The first term reflects the change in prices while the second reflects the change in quantities. The price term using base year weights is a Laspeyres price index and the quantity term using current year weights is a Paasche quantity index.

An alternative index scheme is

$$V_{to} = \frac{V_t}{V_o} = \frac{\sum_i V_{it}}{\sum_i V_{io}} = \frac{\Sigma p_{it} q_{it}}{\Sigma p_{io} q_{it}} \times \frac{\Sigma p_{io} q_{it}}{\Sigma p_{io} q_{io}}$$
$$= \left(\Sigma w_{it} \frac{p_{io}}{p_{it}}\right)^{-1} \times \left(\Sigma w_{io} \frac{q_{it}}{q_{io}}\right) \tag{2.A.4}$$

where the first term using current year weights is a Paasche price index and the second term using base year weights is a Laspeyres quantity index.

There is of course a wide variety of other possible weights and other possible formulas for the construction of index numbers. While it is not our intention to review index numbers in general, we do wish nevertheless to impress on the reader the often rather arbitrary methods used in constructing index numbers. The properties that make an index number useful in some contexts may work to a disadvantage in other contexts. We will presently be discussing the properties that seem desirable in the time-series analysis of demand.

The Problem of Aggregation An aggregate demand relationship such as Equation (2.5) is meant to reflect the reaction of the constant dollar value of imports M to the changes of aggregate income Y/p_Y and changes of a price index p_M/p_Y. It accordingly involves the reactions of rather diverse individuals purchasing rather diverse commodities. We therefore must ask if it is possible to condense such a heterogeneous collection of responses into a single function such as Equation (2.5).

Let us begin by expressing the dependent variable, imports, in the following manner

$$M = \sum_j p_{jo} q_j = M(y_1, \ldots, y_m, p_1, \ldots, p_n, \pi_1, \ldots, \pi_p) \tag{2.A.5}$$

This expresses the fact that the value of imports at base year prices depends on $y_i (i = 1, \ldots, m)$, the income going to the ith individual, $p_j (j = 1, \ldots, n)$, the price of the jth import, and π_k $(k = 1, \ldots, p)$, the price of the kth domestic good. The income terms y_i can also be thought of as including the productive activity of various industries, thereby allowing for the importation of raw materials and unfinished goods. Henceforth we will call the y_i "activity variables" to indicate either income or output.

Equation (2.A.5) is quite general and is hardly disputable. Unfortunately it is not manageable as the number of explanatory variables will ordinarily far exceed the number of available observations. To deal with this problem we will summarize the effects of income and prices in a set of index numbers.

Equation (2.A.5) can be linearized (using the Taylor series approximation) to yield

$$M = k + \left[\sum_i \frac{\partial M}{\partial y_i} y_i\right] + \left[\sum_j \frac{\partial M}{\partial p_j} p_j + \sum_k \frac{\partial M}{\partial \pi_k} \pi_k\right] \tag{2.A.6}$$

To a first-order approximation, the level of imports is seen to depend on an activity term and on a price term. This suggests that an appropriate aggregate activity term is

$$Y = \sum_i \frac{\partial M}{\partial y_i} y_i \tag{2.A.7}$$

One-unit increments in Y will induce one-unit increments in M. We may wish to normalize this expression so that the weights add to one

$$Y_N = \sum_i \frac{\partial M / \partial y_i}{\sum_i \partial M / \partial y_i} y_i \tag{2.A.8}$$

In this case a one-unit increase in Y_N results in a $\sum_i \partial M / \partial y_i$ increment in M.

The important point to notice is that the appropriate aggregate activity term is not the simple unweighted Σy_i, but is rather a summation weighted by the marginal contribution to imports $\partial M / \partial y_i$. Although the marginal responses are unknown, there are three possible approaches that may be followed with regard to them. In the case of the importation of inputs, it may be argued that the marginal responses are closely approximated by the import content of the GNP final-demand components, which are then used as weights. Alternatively one can argue that aggregate output and income are distributed among producers and individuals according to a rule such as

$$y_i = f(\sum_i y_i) = \alpha + \beta \Sigma y_i \tag{2.A.9}$$

In other words, as aggregate activity increases, every component of it increases in a regular fashion. This can be substituted into (2.A.7) to yield

$$Y' = \left(\sum_i \frac{\partial M}{\partial y_i}\right)\left(\alpha + \beta \sum_i y_i\right) \tag{2.A.10}$$

The unweighted activity term would then be sufficient to indicate changes in imports induced by activity changes. Finally, if the marginal responses are similar for all activity components, the unweighted aggregate can be used. This suggests that we use several activity terms, indicating the output/incomes accruing to producers/consumers with widely different marginal import coefficients. The choice among these three alternatives is an empirical issue that has not, as yet, been resolved. Any answer that might be obtained would unfortunately apply only to that particular set of data.

By a similar argument the price term suggested by (2.A.6) is

$$P = \sum_j \frac{\partial M}{\partial p_j} p_j + \sum_k \frac{\partial M}{\partial \pi_k} \pi_k \tag{2.A.11}$$

One-unit increments in P will induce one-unit increments in M. This expression can be made more familiar in the following manner

$$\begin{aligned} P_t &= M_o\left[-\sum_j\left(-\frac{\partial M}{\partial p_j}\frac{p_{jo}}{M_o}\right)\left(\frac{p_{jt}}{p_{jo}}\right) + \sum_k\left(\frac{\partial M}{\partial \pi_k}\frac{\pi_{ko}}{M_o}\right)\left(\frac{\pi_{kt}}{\pi_{ko}}\right)\right] \\ &= M_o\left[-\left(\sum_j - e_j\right)\sum_j\left(\frac{-e_j}{\sum_j - e_j}\right)\left(\frac{p_{jt}}{p_{jo}}\right)\right. \\ &\qquad \left. + \left(\sum_k e_k\right)\sum_k\left(\frac{e_k}{\sum_k e_k}\right)\left(\frac{\pi_{kt}}{\pi_{ko}}\right)\right] \end{aligned} \tag{2.A.12}$$

where the subscripts t and o indicate the time period, and where the letter e indicates the elasticity of M with respect to the appropriate price term. We therefore have P_t expressed as the weighted difference in two price indexes for import and domestic goods

$$\begin{aligned} P_t &= -M_o\left(\sum_j - e_j\right)\sum_j w_j\frac{p_{jt}}{p_{jo}} + M_o\left(\sum_k e_k\right)\sum_k w_k\frac{\pi_{kt}}{\pi_{ko}} \\ &= -M_o(\Sigma_j - e_j)P_M + M_o(\Sigma e_k)P_Y \end{aligned} \tag{2.A.13}$$

where the constants w are the appropriate elasticities normalized to sum to one. It should be noted that in order to combine P_M and P_Y into a single index, we will have to know $\Sigma - e_j$ and Σe_k or be confident that they are equal. We will soon be evaluating these expressions and will argue that they cannot be known a priori. The implication of this remark is that our aggregate import equation must have two price terms, one for imports and one for domestic goods.

This discussion implies that the appropriate price term is a weighted sum of the individual prices for home and foreign goods. Actually a price ratio is more often used. We can produce a price-ratio term by changing (2.A.7) into a multiplicative relation. The linear relation that we have constructed to this point is

$$M = k + \left(\sum_i \frac{\partial M}{\partial y_i}\right) Y_N - M_o\left(\sum_j - e_j\right)P_M + M_o\left(\sum_k e_k\right)P_Y \tag{2.A.14}$$

We can easily calculate the following elasticities

$$\frac{\partial M}{\partial Y_N}\frac{Y_{No}}{M_o} = \left(\sum_i \frac{\partial M}{\partial y_i}\right)\frac{Y_{No}}{M_o} = \sum_i \frac{\partial M}{\partial y_i}\frac{y_{io}}{M_o} = \sum_i e_i$$

$$\frac{\partial M}{\partial P_M}\frac{P_{Mo}}{M_o} = -M_o\left(\sum_j - e_j\right)\frac{P_{Mo}}{M_o} = -\sum_j - e_j \qquad (\text{since } P_{Mo} = 1)$$

$$\frac{\partial M}{\partial P_Y}\frac{P_{Yo}}{M_o} = \sum_k e_k$$

These elasticities imply the multiplicative approximation

$$M = k Y_N^{\Sigma e_i} P_M^{-(\Sigma - e_j)} P_Y^{\Sigma e_k} \tag{2.A.15}$$

At this point it is typical to appeal to the absence of money illusion to argue that the function is linearly homogeneous of degree zero, and therefore

$$\sum_i e_i + \sum_j e_j + \sum_k e_k = 0 \tag{2.A.16}$$

If this is true, we can alter (2.A.15) to

$$M = k\left(\frac{Y_N}{P_Y}\right)^{\Sigma e_i}\left(\frac{P_M}{P_Y}\right)^{-(\Sigma - e_j)} \tag{2.A.17}$$

These two explanatory variables, real income and relative prices, are the ones typically chosen to explain imports. If, however, condition (2.A.16) does not hold, then three explanatory variables must be used: *money* income, import price, and domestic price. We will argue shortly that the absence of money illusion, condition (2.A.16), is too strong a proposition to be known a priori and imposed on the data. All three explanatory variables should therefore be used.

Let us now turn to the evaluation of the elasticities e_j and e_k, which are also the appropriate weights in the price indexes. From (2.A.5) we have

$$M = \sum_j p_{jo}q_j = \sum_j p_{jo}q_j(y_1, \ldots, y_m, p_1, \ldots, p_n, \pi_1, \ldots, \pi_p) \tag{2.A.18}$$

where q_j is the demand function for commodity j.

Thus we have

$$
\begin{aligned}
e_{j'} &= \frac{\partial M}{\partial p_{j'}}\frac{p_{j'o}}{M_o} = p_{j'o}\frac{\partial q_{j'}}{\partial p_{j'}}\frac{p_{j'o}}{M_o} + \sum_{j\neq j'} p_{jo}\frac{\partial q_j}{\partial p_{j'}}\frac{p_{j'o}}{M_o} \\
&= \left(\frac{p_{j'o}q_{j'o}}{M_o}\right)\left(\frac{\partial q_{j'}p_{j'o}}{\partial p_{j'}q_{j'o}}\right) + \sum_{j\neq j'}\left(\frac{p_{jo}q_{jo}}{M_o}\right)\left(\frac{\partial q_j}{\partial p_{j'}}\frac{p_{j'o}}{q_{jo}}\right)
\end{aligned}
\tag{2.A.19}
$$

The proper weights to use in the import price index are seen to be the sum of direct and cross elasticities of demand weighted by the share of the commodity in total imports. A judicious mixture of skill and luck may allow us to define the commodity classes such that the cross elasticities are negligible. In this case all but the first term would drop out and the appropriate import price index can be seen to be the common Laspeyres price index, reweighted by the direct elasticities. Of course these elasticities are not known, else we would not be concerned with estimating them, and we shall have to content ourselves with the standard Laspeyres price index. If, however, we are careful to include commodities in the aggregate import variable with roughly similar elasticities, then the Laspeyres index is the appropriate choice. Alternatively, if such disaggregation is undesirable, we may mitigate the bad effects by including several price terms, one for each disaggregated class of commodities. These classes are to be chosen so that the elasticities differ between classes much more than within classes.

In the main text we mentioned that when an import good and a home good are perfect substitutes, we have to handle them somewhat differently. The same point can be made in the context of the present discussion. A perfectly substitutable good would have an infinite or very large price elasticity. To conform with the weights indicated by (2.A.19), this good would receive all or most of the weight in the index. The appropriate approach would be in this instance to exclude this commodity's direct elasticity from the index and include in the relationship variables that explain the domestic supply. Alternatively, as suggested in the text, we should run separate regressions for these commodities.

Similarly we may evaluate e'_k

$$
\begin{aligned}
e'_k &= \frac{\partial M}{\partial \pi_{k'}}\frac{\pi_{k'o}}{M_o} = \sum_j p_{jo}\frac{\partial q_j}{\partial \pi_{k'}}\frac{\pi_{k'o}}{M_o} \\
&= \sum_j \left(\frac{p_{jo}q_{jo}}{M_o}\right)\left(\frac{\partial q_j}{\partial \pi_{k'}}\frac{\pi_{k'o}}{q_{jo}}\right)
\end{aligned}
\tag{2.A.20}
$$

Once again we have an inner product of a vector of elasticities and a vector of import shares. This time, however, we do not have a direct elasticity that is likely to dominate the expression. We might nonetheless be able to specify a group of imports that are close substitutes for the particular domestic good. Ignoring the effect of the elasticities as we did to obtain the Laspeyres index earlier, the weight we give to the kth domestic good would be the import

share of foreign substitutes for the kth domestic good. If the commodity composition of imports were similar to the commodity composition of domestic goods, we would be able to use the usual GNP price deflator, with domestic weights. It would be clearly advantageous, however, to remove from the index prices of goods that do not substitute for imports, for instance, services.

We pointed out earlier that the specification of imports as a function of real output and relative prices rests on the absence of money illusion in the aggregate relation (2.A.15). This event is expressed quite formally by condition (2.A.16). To evaluate this expression we shall have to define demand functions for individuals as

$$
q_{ij} = q_{ij}(y_i, p_1, \ldots, p_n, \pi_1, \ldots, \pi_p) \tag{2.A.21}
$$

which expresses the purchases of imports of the jth commodity by the ith individual as a function of his own income y_i and the set of prices. The commodity demand function q_j, which we have used earlier, is simply the sum of q_{ij} over all i. It is tedious, but not difficult, to show that if the individual demand functions q_{ij} are homogeneous of degree zero, then the aggregate relation (2.A.15) is likewise, or equivalently that condition (2.A.16) holds. We will content ourselves with the intuitive proof that if no individual suffers from money illusion (that is, if no individual alters his buying habits in response to a doubling of his income and all prices), then the sum of all individuals must behave similarly.

We have seen, therefore, that the homogeneity postulate in the aggregate rests upon the homogeneity postulate for individuals. But can we be so certain of that postulate that it can be imposed on the data by using the relative price term P_M/P_Y and Y/P_Y? We think not. In the first place, the homogeneity postulate is only a postulate and not a necessary description of reality. It describes how people, in seeking to maximize a utility function, ought to behave. Whether they do that or not is quite another question. Secondly, import demand includes the importation of raw materials and unfinished goods. Whereas the postulate has some appeal on the demand side, it is quite a bit more doubtful for the demand for inputs. Finally, even if we firmly believe in the homogeneity postulate, we can use it to justify the relative prices and real income only when the three measures are weighted properly. As the weights are quite impossible to know, we can expect that the approximations that are used will affect the three terms differently. We therefore ought to use the price terms and money income individually.

At this point it is interesting to compare the effect of aggregation on the income and price terms. We have seen with regard to the income term that a rather reasonable assumption concerning income distribution was sufficient to eliminate aggregation effects. When this assumption is unwarranted, comparatively simple cures are available. This is not the case when we explore the price term. We have already discussed the difficult problems and complex solutions surrounding the price term. There is a strong presumption that the price term is more severely damaged by aggregation than the income term.

Conclusion The points made in this appendix are summarized below for reference:

I. Income Term
 A. Weight income components by marginal contribution to imports. Import content may be used for weights on the assumption that the marginal propensity is well approximated by the average import content.
 B. Weighting is unnecessary when all income components behave similarly over time.
 C. Weighting is unnecessary when all income components have similar marginal propensities. Alternatively, use several income terms reflecting incomes accruing to groups with widely different marginal import propensities.

II. Import Price Term
 A. Proper weights are quite complicated; see (2.A.19).
  1. Laspeyres price index is probably the best available approximation.
  2. Disaggregation and/or the inclusion of several price terms will mitigate the bad effects.

III. Domestic Price Term
 A. Proper weights are quite complicated; see (2.A.20).
  1. The readily available Laspeyres index with domestic weights is unlikely to be a good price index.
  2. In the absence of information to the contrary, an unweighted index is as good as any other.
  3. A minimal step toward a proper weighting scheme would exclude prices of domestic goods that clearly do not substitute with imports.

THE USE OF DUMMY VARIABLES IN THE ANALYSIS OF CURVILINEARITY

Some suggestions for handling curvilinearity have been given by Ginsburg [18] and Goldberger [19]. Let us define the following variable

$$W_i = \begin{cases} 1, \text{ if relative price is in interval } i \\ 0, \text{ otherwise} \end{cases}$$

The regression equation would then be

$$M = f\left(\frac{Y}{p_Y}, \frac{p_M}{p_Y}\right) + \sum_i a_i W_i + \sum_i b_i W_i \frac{p_M}{p_Y} \qquad (2.A.22)$$

This function has different levels a_i and different marginal response to price b_i in each of the selected price intervals W_i. The functional form f in this expression becomes the dominant characteristic of the demand equation while the dummy variables indicate the interval-by-interval deviation from that dominant form. In a sense, the data determine the proper functional form. It would appear, especially on the basis of Ginsburg's interesting results [18], that this method of handling curvilinearity should be explored more fully.

In practical application the dummy-variable method for handling curvilinearity may be severely limited by lack of data in much the same sense that the number of explanatory variables is often limited in time-series analysis. An alternative is to use the Durbin–Watson test for serial correlation in the residuals of the regression equation. Usually the residuals are arranged by time period and the Durbin–Watson test compares successive time periods for correlation. This procedure will be ineffective for discovering problems of functional form, however, unless the explanatory variables exhibit a trend over time. An alternative to the Durbin–Watson test would be to specify domains of the explanatory variables and with the use of variance analysis to test for internal homogeneity and external heterogeneity of the residuals within the domains. This alternative is in effect the same procedure as the dummy variable method discussed above.

It is also worth noting that the dummy variable method described does not make the determination of the functional form completely dependent on the data. Cross-product or interaction terms have been excluded. Let us now indicate how such terms can be taken into account.

THE INTRODUCTION OF INTERACTION TERMS IN THE ANALYSIS OF CURVILINEARITY BY MEANS OF DUMMY VARIABLES

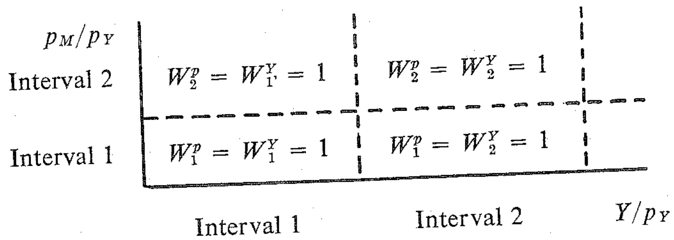
We have just noted that curvilinearity in the import demand function may be handled by specifying dummy variables W_i, which permit the function to have different slopes or elasticities for the particular price intervals chosen. This formulation made no allowance, however, for cross-product interaction terms, such as $(p_M/p_Y \cdot Y/p_Y)$. The situation can be remedied by defining dummy variables as indicated already, but specifying their values in domains of the $(p_M/p_Y \cdot Y/p_Y)$ space rather than in intervals along the p_M/p_Y and Y/p_Y axis.

It is of interest to compare the two approaches for the W variables that take on the values of zero or one. Thus, with no interaction we have

$$W_i^p = \begin{cases} 1, \text{ if } p_M/p_Y \text{ is in interval } i \\ 0, \text{ otherwise} \end{cases}$$

$$W_j^Y = \begin{cases} 1, \text{ if } Y/p_Y \text{ is in interval } j \\ 0, \text{ otherwise} \end{cases}$$

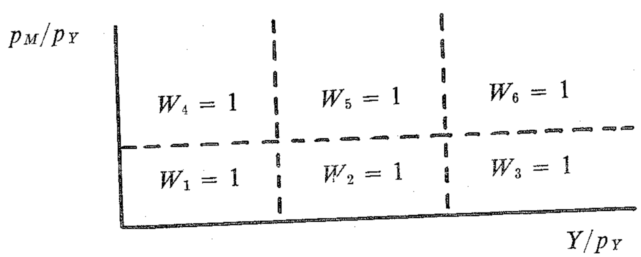
This can be seen diagrammatically as follows



Allowing for interaction we have

$$W_i = \begin{cases} 1, \text{ if the point } (p_M/p_Y,\ Y/p_Y) \text{ is in region } i \\ 0, \text{ otherwise} \end{cases}$$

This is seen diagrammatically as follows



If the price variable is divided into k intervals and the income variable into m intervals, then $k + m$ explanatory dummies are needed when there are no interaction terms present. When interaction is allowed for, $k \times m$ explanatory dummies are required. It should be evident that $k \times m$ will ordinarily be substantially larger than $k + m$, which implies that we may select either short intervals or make allowance for interactions, but not both.

REFERENCES

The following list of references contains most of the important works completed since 1958 on the subject of estimating demand relationships for imports and exports. Earlier works are summarized for convenient reference in Cheng [8]. The reader should also consult the references to Chapter 3 since there is some degree of overlap in the material in that chapter and in the present one. The most significant contributions (in our judgment) are denoted by *; and the principal surveys by **.

1. Allen, R. G. D., *Mathematical Economics*. London: Macmillan, 1965. Chap. 20.

2. Allen, R. G. D. and J. E. Ely, eds., *International Trade Statistics*. New York: John Wiley and Sons, 1953. Chap. 10.

*3. Almon, S., "The Distributed Lag Between Capital Appropriations and Expenditures," *Econometrica*, 53 (January 1965), 178–96.

*4. Ball, R. J. and K. Marwah, "The U.S. Demand for Imports, 1948–1958," *Review of Economics and Statistics*, XLIV (November 1962), 395–401.

5. Ball, R. J., et al., "The Relationship Between United Kingdom Export Performance in Manufactures and the Internal Pressure of Demand," *Economic Journal*, LXXVI (September 1966), 501–18.

6. Bannerji, H., "Analysis of Import Demand," *Artha Vijnana*, 1 (September 1959), 259–70.

7. Branson, W. H., "A Disaggregated Model of the U.S. Balance of Trade," Board of Governors of the Federal Reserve System, Staff Economic Studies, No. 44 (February 29, 1968).

**8. Cheng, H. S., "Statistical Estimates of Elasticities and Propensities in International Trade: A Survey of Published Studies," International Monetary Fund, *Staff Papers*, VII (August 1959), 107–58.

9. Choudry, N. K., "An Econometric Analysis of the Import Demand Function for Burlap (Hessian) in the U.S.A., 1919–53," *Econometrica*, 26 (July 1958), 416–28.

10. DaCosta, G. C., "Elasticities of Demand for Indian Exports—An Empirical Investigation," *Indian Economic Journal*, 13 (July–September 1965), 41–54.

11. Davis, T. E., "A Model of the Canadian Current Account," *The Canadian Journal of Economics and Political Science*, XXXII (November 1966), 468–88.

12. Detomasi, D. D., "Elasticity of Demand for Canadian Imports to the United States," *Canadian Journal of Economics*, II (August 1969), 416–26.

13. Dutta, M., "A Prototype Model of India's Foreign Sector," *International Economic Review*, 5 (January 1964), 82–103.

14. Dutta, M., "Measuring the Role of Price in International Trade: Some Further Tests," *Econometrica*, 33 (July 1965), 600–07.

15. Dutta, M., "Import Structure of India," *Review of Economics and Statistics*, XLVII (August 1965), 295–300.

16. Ferguson, C. E. and M. Polasek, "The Elasticity of Import Demand for Raw Apparel Wool in the United States," *Econometrica*, 30 (October 1962), 670–85.

17. Gehrels, F., "The Effect of Price on Europe's Exports to the United States," *Kyklos*, XXIV (Fasc. 1, 1961), 47–59.

*18. Ginsburg, A. L., *American and British Regional Export Determinants*. Amsterdam: North-Holland Publishing Company, 1969.

19. Goldberger, A. S., *Econometric Theory*. New York: John Wiley and Sons, 1964.

20. Goldberger, A. S., *Topics in Regression Analysis*. New York: The Macmillan Company, 1968.

21. Gray, H. P., "The Demand for International Travel by the United States and Canada," *International Economic Review*, 7 (January 1966), 83–92.

*22. Griliches, Z., "Distributed Lags: A Survey," *Econometrica*, 35 (January 1967), 16–49.

23. Haq, W., "An Econometric Model for International Trade," *Pakistan Economic Journal*, IX (March 1959), 1–13.

*24. Harberger, A. C., "A Structural Approach to the Problem of Import Demand," *American Economic Review*, Papers and Proceedings, XLIII (May 1953), 148–60.

**25. Harley, C. K., "Empirical Literature on the U.S. Balance of Trade," Board of Governors of the Federal Reserve System, Staff Economic Studies, No. 35 (July 1967).

26. Heien, D. M., "Structural Stability and the Estimation of International Import Price Elasticities," *Kyklos*, XXI (Fasc. 4, 1968), 695–711.

*27. Houthakker, H. S. and S. P. Magee, "Income and Price Elasticities in World Trade," *Review of Economics and Statistics*, LI (May 1969), 111–25.

28. Islam, N., "Experiments in Econometric Analysis of an Import Demand Function," *Pakistan Economic Journal*, XI (September 1961), 21–38; XI (December 1961), 1–19.

29. Johnston, J., *Econometric Methods*. New York: McGraw-Hill, 1963.

30. Johnston, J. and M. Henderson, "Assessing the Effects of the Import Surcharge," *Manchester School*, XLV (May 1967), 89–110.

31. Jorgenson, D. W., "A Rational Distributed Lag Function," *Econometrica*, 32 (January 1966), 135–49.

32. Kaliski, S. F., "Some Recent Estimates of the Elasticity of Demand for British Exports—An Appraisal and Reconciliation," *Manchester School*, XXXIX (January 1961), 23–42.

33. Kaliski, S. F., "Extraneous Estimates and Goodness of Fit: A Problem in Empirical Economics," *Manchester School*, XXXVIII (January 1960), 59–71.

34. Kemp, M. C., "Errors of Measurement and Bias in Estimates of Import Demand Parameters," *Economic Record*, 38 (September 1962), 369–72.

35. Kemp, M. C., *The Demand for Canadian Imports 1926–55*. Toronto: University of Toronto Press, 1962.

36. Klein, L. R. et al., *An Econometric Model of the United Kingdom*. Oxford: Oxford University Press, 1961.

37. Koo, A. Y. C., "Marginal Propensity to Import as a Forecaster," *Review of Economics and Statistics*, XLIV (May 1962), 215–17.

38. Koo, A. Y. C., "On Measuring European Tariff Changes and United States Exports Based on Cross Section Data," *Review of Economics and Statistics*, XLV (August 1963), 323–28.

39. Koyck, L. M., *Distributed Lags and Investment Analysis*. Amsterdam: North-Holland Publishing Company, 1954.

40. Krause, L. B., "United States Imports, 1947–1958," *Econometrica*, 30 (April 1962), 221–38.

41. Kravis, I. B. and R. E. Lipsey, *Price Competitiveness in World Trade*. New York: Columbia University Press, forthcoming.

42. Kreinin, M. E., "United States Imports and Income in the Postwar Period," *Review of Economics and Statistics*, LXII (May 1960), 223–25.

43. Lewis, W. A., "International Competition in Manufactures," *American Economic Review*, XLVII (May 1957), 579.

44. Lipsey, R. E., *Price and Quantity Trends in the Foreign Trade of the United States*. Princeton: Princeton University Press, 1963.

*45. Lovell, M. C., "Seasonal Adjustment of Economic Time Series and Multiple Regression Analysis," *Journal of the American Statistical Association*, 58 (December 1963), 993–1010.

**46. Magee, S. P., "Theoretical and Empirical Studies of Competition in International Trade: A Review," an unpublished study for the Council of Economic Advisers (1968).

47. Mills, F. C., *Statistical Methods*. New York: Henry Holt and Company, 1955. Chaps. 13 and 14.

48. Moore, L., "Factors Affecting the Demand for British Exports," *Bulletin of Oxford University Institute of Economics and Statistics*, 25 (1963), 343–59.

*49. Morgan, D. J. and W. J. Corlett, "The Influence of Price in International Trade: A Study in Method," *Journal of the Royal Statistical Society*, Series A (General), CXIV (Part III, 1951), 307–58.

50. Murakami, A., "Import Behavior in Asian Countries," *Kobe University Economic Review*, 7 (1961), 101–14.

51. Neisser, H., "Comment" on "Some Evidence on the International Price Mechanism," *Review of Economics and Statistics*, XL (February 1958, Supplement), 129–32.

52. Nerlove, M., "Distributed Lags and Estimation of Long-Run Supply and Demand Elasticities: Theoretical Considerations," *Journal of Farm Economics*, XL (May 1958), 301–11.

*53. Orcutt, G. H., "Measurement of Price Elasticities in International Trade," *Review of Economics and Statistics*, XXXII (May 1950), 117–32.

54. Organization for Economic Co-operation and Development, Department of Economics and Statistics, "Statistical Import Functions and Import Forecasting: An Empirical Study of Import Determination in Seven Member Countries," DES/NI/F(66)4 (17 October 1966).

55. Organization for Economic Co-operation and Development, F. G. Adams, et al., *An Econometric Analysis of International Trade*. Paris: OECD, 1969.

*56. Polak, J. J., *An International Economic System*. Chicago: University of Chicago Press, 1953.

*57. Prachowny, M. F. J., *A Structural Model of the U.S. Balance of Payments*. Amsterdam: North-Holland Publishing Company, 1969.

**58. Prais, S. J., "Econometric Research in International Trade: A Review," *Kyklos*, XV (Fasc. 3, 1962), 560–79.

59. Preeg, E. H., "Elasticity Optimism in International Trade," *Kyklos*, XX (Fasc. 4, 1967), 460–69.

60. Rao, S. V., "Elasticities of Demand for Imports During the Period 1920–21 to 1929–30," *Artha Vijnana*, 2 (December 1960), 307–12.

61. Reimer, R., "The United States Demand for Imports of Materials," *Review of Economics and Statistics*, XLVI (February 1964), 65–75.

*62. Rhomberg, R. R., "A Model of the Canadian Economy Under Fixed and Fluctuating Exchange Rates," *Journal of Political Economy*, LXXII (February 1964), 1–31.

63. Rhomberg, R. R., "Canada's Foreign Exchange Market," International Monetary Fund, *Staff Papers*, VII (April 1960), 439–56.

*64. Rhomberg, R. R. and L. Boissonneault, "The Foreign Sector," in J. S. Duesenberry et al., eds., *The Brookings Quarterly Econometric Model of the United States*. Chicago: Rand McNally & Company, 1965.

65. Rhomberg, R. R. and L. Boissonneault, "Effects of Income and Price Changes on the U.S. Balance of Payments," International Monetary Fund, *Staff Papers*, XI (1964), 59–124.

66. Robinson, T. R., "Canada's Imports and Economic Stability," *Canadian Journal of Economics*, I (May 1968), 401–28.

67. Roy, P. N., "Some Export Demand Functions for India," *Arthanti*, IV (July 1961), 129–37.

68. Sasaki, K., "Quantitative Effect of the United States Economy on Japan's Foreign Trade Between 1950–1956," *Review of Economics and Statistics*, XLI (August 1959), 320–24.

69. Scott, M. FG., *A Study of United Kingdom Imports*. Cambridge: Cambridge University Press, 1963.

70. Stern, R. M., *Foreign Trade and Economic Growth in Italy*. New York: Praeger, 1967. Chap. 2.

71. Suits, D. B., *Statistics: An Introduction to Quantitative Economic Research*. Chicago: Rand McNally, 1963.

*72. Suits, D. B., "Use of Dummy Variables in Regression Equations," *Journal of the American Statistical Association*, 52 (December 1957), 548–51.

73. Swamy, D. S., "A Quarterly Econometric Model of Demand for and Supply of Exports," *Indian Economic Review*, I (April 1966), 79–103.

74. Taylor, W. B., "Short Term Factors Influencing New Zealand Lamb Prices in the United Kingdom," *Economic Record*, XXXVI (December 1960), 568–80.

75. Turnovsky, S. J., "International Trading Relationships of a Small Country: The Case of New Zealand," *Canadian Journal of Economics*, I (November 1968), 772–90.

76. Von Böventer, E., "The Production Elasticity of U.S. Raw Material Imports," *Economia Internazionale*, XII (August 1959), 470–91.

77. Wemelsfelder, J., "The Short-term Effect of the Lowering of Import Duties in Germany," *Economic Journal*, LXX (March 1960), 94–104.

78. Wilkinson, B. W., *Canada's International Trade: An Analysis of Recent Trends and Patterns*. Montreal: Private Planning Association of Canada, for the Canadian Trade Committee, 1968.

79. Zarembka, P., "Functional Form in the Demand for Money," *Journal of the American Statistical Association*, 63 (June 1968), 502–11.