Agriculture Handbook No. 141

Distributed Lags and Demand Analysis

for Agricultural and Other Commodities

Agricultural Marketing Service
UNITED STATES DEPARTMENT OF AGRICULTURE

June 1958

PREFACE

This publication summarizes available literature on the use of distributed lags in the analysis of demand for individual commodities and contributes a substantial amount of new material to the problem of estimating dynamic demand relationships. Although distributed lags have been used in the analysis of some economic problems, they have been used to date only rarely in connection with demand analyses. These techniques appear to be particularly useful in measuring long-run and short-run demand functions. Several applied examples are discussed in detail in this publication. These techniques are currently being used by the Agricultural Marketing Service in an analysis of the demand for textile products; results on this application will be reported in later publications. In this report considerable emphasis is placed on methodological aspects, in part to call attention to problem areas. It is hoped that this publication will result both in a stimulation of additional research on methodology and in increased use of this technique in applied areas.

For sale by the Superintendent of Documents, U. S. Government Printing Office
Washington 25, D. C. - Price 60 cents

CONTENTS

DISTRIBUTED LAGS AND DEMAND ANALYSIS FOR AGRICULTURAL AND OTHER COMMODITIES

by

20

Marc Nerlove,
Agricultural Economic Statistician
Agricultural Marketing Service

SUMMARY

In economics, a cause often produces its effect only after a lapse of time. For example, a drop in the price of potatoes in the fall cannot affect potato acreage until the following spring, nor can it decrease potato production until the following fall. The lapse of time between a cause and its effect is called a lag. The lag may be a specific time, say three months, or one year. But in many cases, the effects of an economic cause are spread over many months, or even many years. In such cases, we have a distributed lag. Distributed lags with respect to variables that affect consumption may arise for the following reasons: (1) Psychological, under which we include forces of habit and assumptions on the part of consumers that changes may be only temporary; (2) technological, which include factors such as, in a general case, lack of knowledge about possible substitutes or, in a specific case, the inability to increase greatly the use of frozen foods without first acquiring adequate freezer storage space; (3) institutional, which include situations in which certain contractual items of expenditure or savings may need to be adjusted before shifts can be made in consumption patterns, and the fact that some markets, particularly for durable goods, are imperfect in an economic sense.

One way to measure the degree of lag with respect to a particular variable is to find by statistical analysis that distribution of lag which maximizes the effect of the causal factor. Empirical analyses based on this approach have been run which (1) make no assumption as to the form of the distribution of lag, (2) are designed to estimate certain characteristics of an assumed general form for the distribution, or (3) derive and statistically fit a model based on the fundamental cause of the distributed lag but which yield a specific distribution of lag only incidentally. Examples of each are discussed in some detail in this report.

One way to formulate models for generating distributed lags is to assume that the lags arise chiefly because of technological and institutional rigidities. The traditional distinction between long- and short-run elasticities rests largely on causes of this sort. Although observable points each lie on a short-run curve, the coefficients of the long-run curve can be estimated if the model is properly formulated. Relations between long- and short-run demand curves depend chiefly on the path which observed consumption would follow if it moved directly toward its long-run equilibrium level following an initial change in a causal variable. The shape and form of such paths are determined by the type of institutional and technological rigidities that exist. Several alternative paths are described in this report in terms of difference or differential equations.

Another way to formulate models that generate distributed lags is to assume that technological and institutional rigidities are absent but that uncertainty about the future exists, and that habit is a powerful force. Under conditions of uncertainty, a change in price or income may be thought of by consumers as divided into two components--one permanent and one transitory. The permanent part changes expectations in all relevant future years, whereas the transitory part changes expectations in only some future years, or perhaps in none. The average level about which future prices or incomes are expected to fluctuate is called the "expected normal;" this level is affected only by that part of the initial change in a causal factor that is considered permanent. If forces of habit are strong, the effects of a change in current price or income on consumer behavior are slight compared with the effects of a change in the expected normal. Hence, models need to be formulated so as to emphasize the effects of changes in the expected normal.

Hicks (18, p. 205) 1/ defines "the elasticity of a particular person's expectations of the price of a commodity x as the ratio of the proportional rise in expected future prices of x to the proportional rise in its current price." The elasticity can range between 0 and 1. 2/ Many factors affect the elasticity of expectations; some affect all commodities equally, whereas others affect different commodities differently. Among the latter is the typical variance of prices of the commodity. For any given commodity the elasticity of expectations need not be stable over time, but models are simpler if we assume that it is stable. A number of applications based on this assumption are discussed. Frequently an equation that contains one or more distributed lags can be reduced by algebraic manipulation to an equation that does not contain such lags. The reduced equation then can be fitted statistically and the results used to obtain certain characteristics of the lag. A difficulty with this approach is that the number of variables added in the reduced equation is greater than or equal to the square of the number of variables with distributed lags that enter the initial equation. Thus the method is not feasible when the initial equation contains more than two or three variables with distributed lags.

In working with commodities that substitute for, or complement each other, we generally are interested in a system of equations rather than a single equation. Each equation frequently contains lagged values of the same variables. The multiple equation method of reduction of a demand equation which involves distributed lags takes advantage of the existence of all of the interrelated equations. By way of contrast with the single equation method of

---

1/ Underlined numbers in parentheses refer to Literature Cited, page 116.
2/ That the elasticity lies between zero and one is not necessary to Hicks' definition nor to much of the analysis contained in subsequent pages. The chief justifications for assuming the elasticity to lie between zero and one are: (1) Its interpretation as a proportion (see page 23); (2) mathematical convenience; (3) empirical validity. The last is, of course, the most compelling. In all statistical investigations in which the elasticity of expectations has been estimated without restricting it to lie between zero and one, it has been found to do so.

reduction described in the preceding paragraph, if each original demand equation contains n expected normal prices and expected normal income, each reduced demand equation in the multiple set includes, in addition to n + 1 current values of prices and income, n + 1 lagged values of the quantities demanded and aggregate consumption of all commodities. Thus, whereas the single equation method leads to one reduced demand equation containing $(n + 1) + (n + 1)^2$ independent variables, the multiple equation method leads to n reduced demand equations each containing $2(n + 1)$ independent variables. For statistical purposes, the latter are clearly preferable when n exceeds 1 or 2. Some systems of demand equations are of such a nature as to be "separable" in a mathematical sense defined in this paper. An example is a system containing several demand equations and a consumption function; here we may (1) reduce the whole system, including the consumption function or (2) solve the consumption function for expected income, substitute this into the demand equations, and then reduce only the demand equations. This is the sense in which the term "separable" is used.

Equations that contain distributed lags that are due only to technological or institutional rigidities can be reduced easily. Reduction becomes more complex when the lags are due to uncertainty about the future, but a considerable degree of simplification may be obtained by the multiple equation method of reduction. When the lags are due both to uncertainty and rigidities, in general no simplification of reduction is possible. But if the distributions of lag are of a special form, simple reduction is possible. Several examples are discussed.

Statistical estimation of the coefficients in demand equations that contain distributed lags can be done theoretically in two ways--(1) by dealing directly with the equation that involves the distributed lags, or (2) by using one or more reduced equations. Maximum likelihood procedures under the first approach are given and a likelihood ratio test is described. This approach requires a large number of repeated steps; for this reason, it is called the "iterative" method of estimation. In order to use the iterative method, the unexplained residuals in each equation that contains a distributed lag must be normally and independently distributed. If the lags arise solely because of technological and institutional rigidities, this will be true only under special conditions. If the lags are caused by uncertainty, problems of serial correlation in the residuals do not arise but the iterative procedure is computationally feasible only in the simplest cases. Thus the iterative approach can be used only under special circumstances.

If a distributed lag is due only to rigidities of a technological or institutional nature, the coefficients can be estimated easily by using a single reduced equation fitted by least squares. Estimation becomes more complicated if the lags are caused by uncertainty about the future. If three or more variables with distributed lags are involved and we use the single-equation method of reduction, the number of variables in the reduced equation becomes so large as to make statistical fitting virtually impossible with time series of normal length. Complications from serial correlation of the residuals also enter. Only in the simplest of cases should the non-iterative method be based on a single reduced equation. In using the multiple reduced equation approach, we must assume that the distributions of lag for the same variable in each of the equations is the same. If the multiple equation method can be used, it is computationally much simpler than the single equation method.

In analyzing the demand for commodities in general, if anticipated distributed lags result from uncertainty about the future, we should specify a system of equations before proceeding with any estimation. If the distributed lags are believed to result only from technological or institutional rigidities, only a single demand equation for the particular commodity need be specified.

In a concluding section, the permanent income hypothesis, advanced by Friedman (16), is discussed in some detail. Two empirical tests of this hypothesis are given. Although a firm conclusion cannot be reached based on these, they suggest that technological and institutional rigidities may play a greater role in demand analysis than does uncertainty about the future. If this is true, estimation of distributed lags by the methods discussed in this paper is simplified.

INTRODUCTION

Irving Fisher (13) was the first to use and discuss the concept of a distributed lag. Although the idea that one economic variable depends on another variable lagged in time is an old one, the use of lagged variables in empirical research has been restricted primarily to areas other than demand analysis.

Distributed lags arise in theory when any economic cause (for example, a price change or an income change) produces its effect (for example, on the quantity demanded) only after some lag in time; so that this effect is not felt all at once, at a single point of time, but is distributed over a period of time. Thus, when we say that the quantity of cigarettes demanded is a function of the price of cigarettes taken with a distributed lag, we mean, essentially, that the full effects of a change in the price of cigarettes is not felt immediately, and that only after some passage of time does the quantity of cigarettes demanded show the full effect of the change in the price of cigarettes.

This example may be made more concrete: Let time be divided into discrete periods and let q_t = the quantity demanded during period t, p_t = the price during period t, p_{t-1} = the price during period t-1, and so on. Then, in line with the above, q_t may be written as a function of past prices:

$$q_t = f(p_t, p_{t-1}, p_{t-2}, \dots) \tag{1}$$

Suppose that a tax is levied which raises the price of the commodity by an amount Δp. 3/ Then, in period t+1, equation (1) yields

3/ We assume that the supply of the commodity is perfectly elastic.

Essentially ... during period t+1

$$q_{t+1} = f(p_{t+1} + \Delta p, p_t, p_{t-1}, \cdots) \qquad (2)$$

In period t+2, we have

$$q_{t+2} = f(p_{t+2} + \Delta p, p_{t+1} + \Delta p, p_t, p_{t-1} \cdots) \qquad (3)$$

so that q_{t+2} is not generally the same as q_{t+1} even though no further change in price has occurred. Thus, the effect of a price change may be distributed over many periods of time.

CAUSES OF DISTRIBUTED LAGS

The economic variables which affect demand may do so with a distributed lag for a variety of reasons. These reasons fall into three broad groups: (1) Psychological, (2) technological and (3) institutional. Typically, some conjunction of factors falling in all three groups operate to produce a distributed lag.

Psychological

Whether we consider the demand for individual commodities or consumption expenditures as a whole, we can reasonably suppose that people are loath to change their pattern of consumption or their level of living radically in response to changes in prices or income. Two basic reasons cause this rigidity of behavior: (1) Habit is a powerful force; the process of change is an activity to which disutility may be attached, particularly when the change tends to result in a reduction in the standard of living. (2) Changes in economic variables may be considered only temporary; hence, the disutility involved in adjustment and readjustment may more than offset the utility to be gained by maintaining a continuous position of equilibrium even at a higher level of living. Consequently, habit and uncertainty about the future make for rigidity in consumer behavior. If a change in price or income persists for a sufficiently long period, consumers may become convinced of its permanence and act accordingly.

Changes in total expenditures for consumption in response to changes in income may be considered from a slightly different approach. It is reasonable to suppose that consumers wish to even out their consumption to a certain extent over their "lifetimes," or at least over the forseeable future. 4/ If this is true, an individual consumer tends to save when his income is temporarily high and to dissave when his income is temporarily low; his total consumption during any period of time thus is related to his expected longer-term earnings, not to his current income. Consequently, total consumption expenditures tend to be stable relative to current incomes, and a change in current income tends to affect consumption only insofar as it affects consumers' notions of their "lifetime" or "permanent" incomes. If these notions are relatively

4/ See Friedman (16, Chap. II) and Fisher (14, pp. 73-76, 231-262).

inelastic with respect to current income, changes in total consumption expenditures lag behind changes in current income, and this lag is likely to be of a distributed nature for, if a change in income persists for a sufficiently long period, people come to believe in its permanence. If total consumption is related to income taken with a distributed lag, at least some individual items must be also. 5/

Technological

The economic theory of the individual consumer or family unit is similar to the theory of the individual firm: a firm maximizes profits subject to various restraints and a consumer maximizes satisfaction or utility subject to restraints. Just as a firm produces a product with fixed as well as variable factors, so a consumer produces satisfaction with stocks of durable or semi-durable goods as well as with goods of a more perishable nature. The existence of consumption goods of a durable nature leads to a lag in the reactions of consumers for technological reasons: the shift from one perishable commodity to another may be delayed because of the existence of complementary goods of a durable nature. For example, the introduction of frozen foods was probably delayed because consumers did not have proper storage facilities. Manufacturers of refrigerators took time to recognize the need for larger freezer compartments and to produce such refrigerators. Furthermore, consumers took time to adjust their stock of refrigerators, given the existence of those of the older type.

The ability of consumers to substitute one commodity for another depends on the passage of time for another reason--imperfect knowledge. A rise in the price of one commodity may provide the incentive to consumers to substitute others for it, but time may be required to discover what commodities may be easily substituted and how to substitute them. Similarly, a fall in the price of a commodity may lead to a desire, on the part of consumers, to substitute it for another commodity, but time is required to make the relevant substitutions.

Institutional

Institutional factors may also produce a certain rigidity in consumer behavior for the following reasons: (1) Many contractual items of expenditure or saving exist, for example, insurance premiums, installment credit, and the like. Given time, all these may be adjusted, but in the short run they may cause consumers to adjust purchases less than they otherwise would.
(2) Markets for many goods are imperfect, and large costs to the consumer may result from changes in the pattern of expenditure. For example, the

---

5/ In an empirical context, the accuracy of these statements depends upon the distinction between savings and consumption. Purchases of some goods, such as refrigerators or automobiles, reflect saving more than they do consumption. This problem is discussed in subsequent paragraphs.

secondhand market for many durable goods, say refrigerators, tends to be imperfect, so that buyers and sellers have difficulty in getting together. Costs involved in selling or negotiating for a sale may persuade a consumer to retain his old refrigerator longer than he would in a more perfect market.

The distinction between the various causes of distributed lags is never clear-cut. In practice one must decide on the basis of available evidence whether the behavior of the consumer is likely to be affected by one or more of these considerations; if so, a distributed lag might properly be introduced.

Effects of Uncertainty

When we discuss models leading to distributed lags, we condense the three categories of causes into two categories. Even if no uncertainty about the future exists, rigidities in consumer behavior might arise for all three types of reasons, but the distributed lags in such a case might be quite different from those arising in a world of uncertainty. A practical example of this distinction is as follows: Suppose the Government announces a fixed price for bread, higher than the current price, at which the Government is willing to buy or sell bread in unlimited quantities. Contrast this with a more usual case in our economy in which the price of bread rises because of a short wheat crop. In each situation people tend to consume less bread and, if the new price persists, an equilibrium will eventually be reached. Even if the equilibrium position is the same in each case, the approach to equilibrium may be quite different, since in one case the change in price may be considered permanent whereas in the other it may be considered temporary, at least for a while.

In many practical uses to which demand analyses may be put, the distinction between lags in adjustment under more-or-less certain conditions and lags under conditions of uncertainty is of great importance. The discussion beginning on page 14 is therefore concentrated on this twofold distinction rather than on the three categories discussed in preceding paragraphs.

EMPIRICAL ANALYSES IN WHICH DISTRIBUTED LAGS HAVE BEEN USED

According to Fisher (15, p. 323) the basic problem in applying the theory of distributed lags "... is to find the 'best' distribution of lag, by which is meant the distribution such that ... the total combined effect [of the lagged values of the variable taken with a distributed lag] ... [has] the highest possible correlation with the actual statistical series with which we wish to compare it." That is, we wish to find that distribution of lag which maximizes the explanation of "effect" by "cause" in a statistical sense.

This problem of finding the "best" distribution of lag may be attacked in several ways: (1) We may make no assumption as to the form of the distribution; this approach is taken by Alt (2) and Tinbergen (31). (2) We may assume a general form for the distribution of the lag and estimate certain of its particular characteristics; this approach is taken by Fisher (13, 15) and Koyck (22) and is suggested by Alt (2). (3) We may develop a specific model based on the considerations suggested in the preceding section; this model yields a specific distribution of lag only incidentally. This approach is examined further in the sections that begin on page 14. Approaches (2) and (3) do not yield the "best" explanation in a statistical sense, but they do offer certain advantages in terms of ease of computation and interpretation.

When No Assumption Is Made As To the Form of the Distribution

Tinbergen (31) suggests that the typical form of a demand equation is

$$x_t = a + b_0 p_t + b_1 p_{t-1} + \dots = a + \sum_{i=0}^{\infty} b_i p_{t-i} \tag{4}$$

where x_t is the quantity demanded in period t; p_t, the price in period t; p_{t-1}, the price in t-1; and so on, and the b_0, b_1, ... are constants. Let ϵ_S be the short-run elasticity of demand (that is, the immediate effect of a one percent change in price), and let ϵ_L be the long-run elasticity of demand (that is, the eventual effect of a one percent change in price). Tinbergen proposed to interpret the short-run elasticity as

$$\epsilon_S = b_0 \frac{\bar{p}}{\bar{x}} \tag{5}$$

and the long-run elasticity as

$$\epsilon_L = \left(\sum_{i=0}^{\infty} b_i \right) \frac{\bar{p}}{\bar{x}} \tag{6}$$

where $(\bar{p}, \bar{x})$ is the point on the demand function at which we wish to evaluate the elasticity.

The method proposed by Tinbergen for estimating an equation such as (4) involves no assumption as to the relation among the b_i, that is, no assumption regarding the distribution of the lag effect of past prices on the quantity demanded. Tinbergen proposed to add lagged prices successively to the least squares regression of quantity on price. This procedure stops when the signs of the coefficients become erratic and cease to make sense. Such behavior of the coefficients is usually a consequence of the intercorrelation of the lagged explanatory variables. Sizable standard errors of the coefficients, even where the signs of the coefficients have not yet become erratic, may suggest stopping the procedure at an earlier stage.

An example of this method is given by Alt (2, p. 116). Let y_t represent fuel oil consumption in quarter t, and x_t represent new orders in quarter t; Alt's calculations for the period 1930-39 are

$$y_t = 8.37 + 0.171\, x_t \tag{7a}$$

$$y_t = 8.27 + 0.111\, x_t + 0.064\, x_{t-1} \tag{7b}$$

$$y_t = 8.27 + 0.109\, x_t + 0.071\, x_{t-1} - 0.005\, x_{t-2} \tag{7c}$$

$$y_t = 8.32 + 0.108\, x_t + 0.063\, x_{t-1} + 0.022\, x_{t-2} - 0.020\, x_{t-3} \tag{7d}$$

Alt chooses equation (7d) because, at the stages represented by (7c-d), the coefficients of x_{t-2} and x_{t-3} were erratic. Hence, (7b) is chosen, not because the influence of new orders two and more quarters removed is zero, but because the influence cannot be safely ascertained. Two points should be noted: (1) The determination of whether the signs of the coefficients are erratic depends on some *a priori* notion of the distribution of lag, although alternating signs in successive equations certainly suggests an erratic situation; and (2) the fact that a variable taken with a given lag does not have a determinable influence (in a least squares regression) does not mean that it has no influence. When the intercorrelations among lagged values of the independent variables are high, the method suggested by Tinbergen is not likely to prove useful.

When the General Form of the Distribution Is Assumed and Estimation Is Made of Certain Characteristics

Fisher (13, 15), Alt (2), and Koyck (22) suggest that the form of the distribution of the lag may be appropriately assumed and that specific characteristics of the distribution may then be estimated, given the form.

Use of a logarithmic normal curve time path.--Fisher (15, p. 323) writes: "This best distribution [see the quotation from Fisher on page 7] presumably has the general form of some type of probability curve. The working assumption is made that, immediately following the cause ..., the effect is very small, but that it rises fast, reaching its mode after a very short interval and then tapering off slowly. More specifically, the type of distribution assumed is the 'normal' probability curve, provided, however, that not time but the *logarithm* of time is used as the abscissa. (If time itself is plotted, the curve is evidently very skew.) That is, the type is 'normal' logarithmically."

If time is treated as a continuous variable, rather than a discrete variable, equation (4), which represents a demand function, might be written

$$x(t) = a + \int_0^{\infty} b(u)\, p(t-u)\,du \tag{8}$$

where the quantity demanded, $x(t)$, and the price, $p(t)$, are continuous functions of time, and $b(u)$ is the "weight" attached to price at the instant of time $t-u$. Fisher's assumption was that the weights $b(u)$ follow a logarithmic normal distribution, such as that shown in figure 1. Clearly, the initial effect of a price change is small, but as time passes the cumulative effect of the change becomes greater. The full effects of a change are realized theoretically, however, only after the passage of an infinite amount of time. All this, of course, is on the assumption that the change in price persists. If the change occurs only instantaneously, the variable $x(t)$ itself follows a path such as that shown in figure 1. The maximum effect of an instantaneous change occurs only at time u_m after the event, where u_m is the mode of the distribution of $b(u)$. Koyck (22, pp. 9-10) calls the distribution of $b(u)$ the "time-shape of an economic reaction" and the cumulative distribution of $b(u)$, "the adjustment path."

Fisher (13) used a distributed lag of the form shown in figure 1 to study the relation between changes in the general price level and changes in business activity. The logarithmic normal distribution of figure 1 was approximated by a discrete distribution, and the correlation between a weighted average of current and past price levels and the level of business activity was maximized with respect to changes in the variance of the distribution of weights.

Use of an approximation to the logarithmic normal curve.--Later Fisher (15) suggested an approximation to the distribution shown in figure 1 which involves considerably less computational labor. Fisher (15, p. 324) describes his short-cut method as follows: "By ... [the] short-cut method the influence of any cause ... is assumed to be the greatest at the very next time unit (month or quarter) and then to taper off by equal decrements for each successive time unit. That is, the distribution curve becomes a straight line beginning one month, or other time unit, after the 'cause.' There is only one parameter to determine, namely the length of time elapsing to the end of the distribution." The weight given to the value of the independent variable lagged i periods by Fisher's short-cut method is proportional to

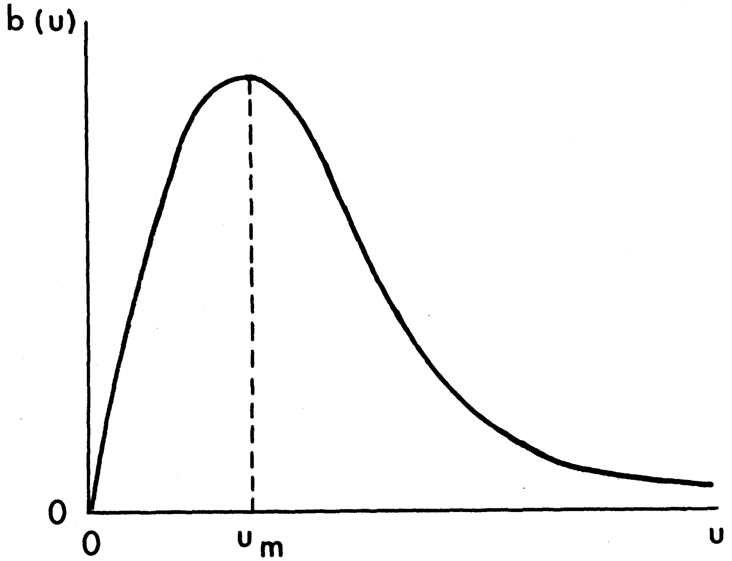
$$\frac{N-i}{\sum\limits_{n=1}^{N} n} \qquad i < N \tag{9}$$

where N is the total number of periods to the point at which the effect becomes zero. If time is taken as a continuous variable, the weight for the value lagged an amount of time t is proportional to

$$\frac{1}{N} - \frac{1}{N^2}\,t \qquad t < N \tag{10}$$

The form of Fisher's second distribution of lag is shown in figure 2.

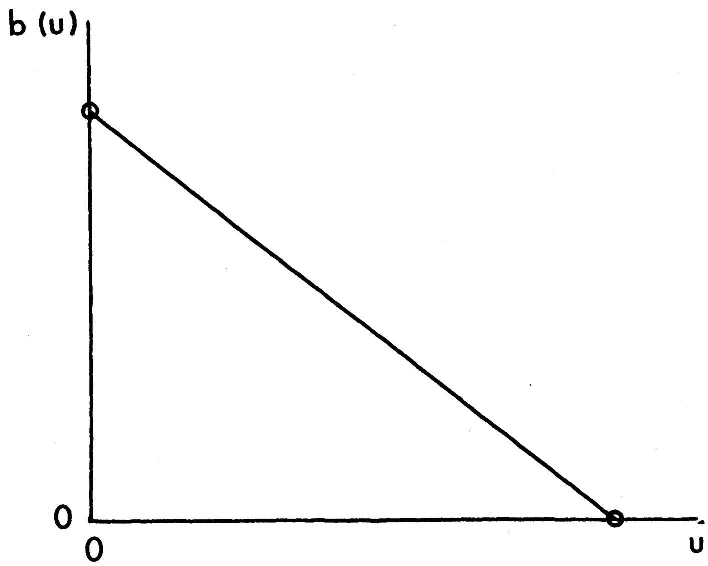
FISHER'S FIRST DISTRIBUTION OF LAG



U. S. DEPARTMENT OF AGRICULTURE NEG. 4408-57 (8) AGRICULTURAL MARKETING SERVICE

Figure 1

FISHER'S SECOND DISTRIBUTION OF LAG



U. S. DEPARTMENT OF AGRICULTURE NEG. 4409-57 (8) AGRICULTURAL MARKETING SERVICE

Figure 2

Fisher's short-cut method may be illustrated by Berger's study of India's glass imports from the United Kingdom (4). Berger ran the following least squares regressions:

$$x_t = a + b\,\frac{(3p_t + 2p_{t-1} + p_{t-2})}{6} \tag{11a}$$

$$x_t = a + b\,\frac{(4p_t + 3p_{t-1} + 2p_{t-2} + p_{t-3})}{10} \tag{11b}$$

$$x_t = a + b\,\frac{(5p_t + 4p_{t-1} + \cdots + p_{t-4})}{15} \tag{11c}$$

$$x_t = a + b\,\frac{(6p_t + 5p_{t-1} + \cdots + p_{t-5})}{21} \tag{11d}$$

where x_t = the ratio of glass imports from the United Kingdom to total glass imports during period t, and p_t = the ratio of British glass prices to prices of competing glass. The simple correlations between the dependent variable and the weighted average independent variable were 0.858, 0.881, 0.836, and 0.751 for regressions (11a), (11b), (11c), and (11d), respectively. Regression (11b), with the largest correlation, was selected as showing the "best" distribution of lag.

Lag distributions used by Alt and related distributions.--Alt (2) suggested special lag distributions of the general form

$$A\ e^{-B(N-i-c)^2} \tag{12}$$

or

$$A(N-i)\ e^{-B(N-i-c)^2} \tag{13}$$

where N and i have the same meaning as before. The derivatives of most of the commonly used growth curves, such as the Gompertz or the Logistic, also provide reasonable distributions of lag. All of these distributions are difficult to estimate in practice due to the number of parameters involved.

Lag distributions that involve only a single parameter.--Koyck (22) suggests an ingenious distribution of lag which involves only one parameter and which lends itself readily to statistical application. Let time be measured as a discrete variable and consider a demand equation such as (4). Koyck's assumption is that after a certain point, say i=k, the series of coefficients b_i, i=0, 1, ..., can be approximated by a convergent geometric series, so that

$$b_{k+m} = \delta\, b_{k+m-1} \tag{14}$$

where $m \geq 0$ and $0 \leq \delta < 1$. From (4) and (14) it follows that

$$x_t = a + b_0 p_t + \dots + b_{k-1} p_{t-k+1} +$$

$$b_k p_{t-k} + b_k \delta p_{t-k-1} + b_k \delta^2 p_{t-k-2} +$$

$$b_k \delta^3 p_{t-k-2} + \dots + b_k \delta^m p_{t-k-m} + \dots .$$

$$= a + b_0 p_t + \dots + b_{k-1} p_{t-k+1} + b_k \sum_{m=0}^{\infty} \delta^m p_{t-k-m} \qquad (15)$$

Thus, x_t is a function of k-1 unweighted lagged prices and a geometrically weighted average of all other past prices. If time is treated as a continuous variable, Koyck's distribution of lag has the form shown in figure 3. Figure 3 shows the distribution plotted for different values of the parameter δ.

If k=0, the long- and short-run elasticities and the exact distribution of lag are particularly easy to estimate if the distribution has the general form assumed by Koyck. Consider equation (15) with k=0. Then

$$x_t = a + b_0 p_t + b_0 \delta p_{t-1} + b_0 \delta^2 p_{t-2} + \dots . \qquad (16)$$

If we lag (16) one period and multiply by δ, we get:

$$\delta x_{t-1} = a \delta + b_0 \delta p_{t-1} + b_0 \delta^2 p_{t-2} + \dots . \qquad (17)$$

Now subtract (17) from (16) to get:

$$x_t = a(1-\delta) + b_0 p_t + \delta x_{t-1} \qquad (18)$$

The distribution of lag is given by the estimate of δ, and the short-run elasticity of demand is given by $b_0 \frac{\bar{p}}{\bar{x}}$. The cumulative effect of a maintained price change is

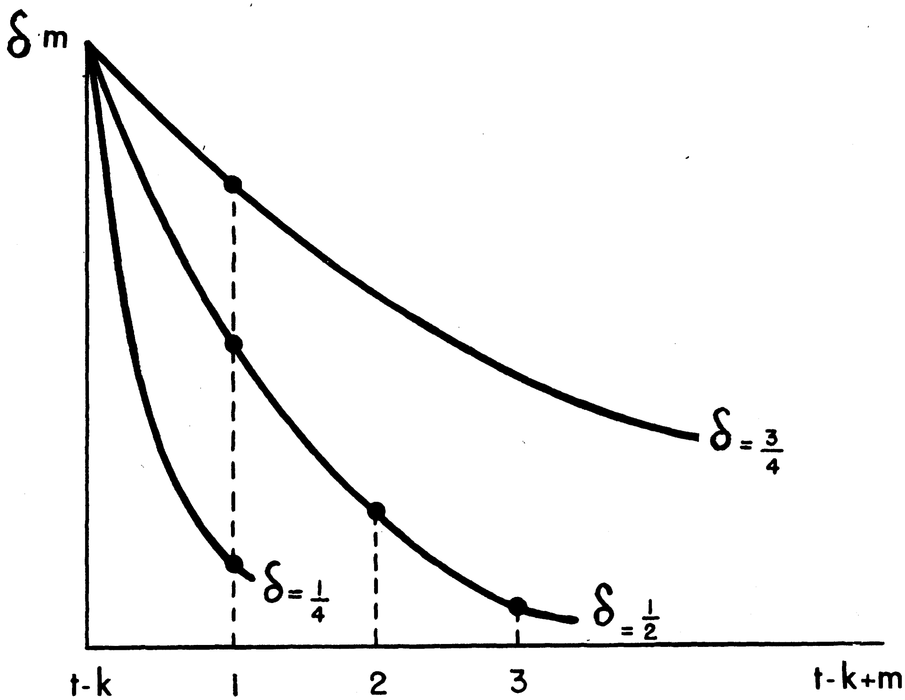
$$b_0 \sum_{m=0}^{\infty} \delta^m = \frac{b_0}{1-\delta} \qquad (19)$$

if $0 \leq \delta < 1$. Hence, the long-run elasticity of demand is given by

$$\frac{b_0}{1-\delta} \quad \frac{\bar{p}}{\bar{x}} \qquad (20)$$

The statistical difficulties which arise in the use of Koyck's distribution of lag are discussed below.

KOYCK'S DISTRIBUTION OF LAG



U. S. DEPARTMENT OF AGRICULTURE NEG. 4410-57 (8) AGRICULTURAL MARKETING SERVICE

Figure 3

Based on Models of Expectation

Friedman (16) and Cagan (5) each suggest a model of expectation formation which leads to a distribution of lag remarkably similar to Koyck's distribution. This model is discussed beginning on page 20 after certain simpler models have been described.

MODELS FOR GENERATING DISTRIBUTED LAGS BASED ON TECHNOLOGICAL AND INSTITUTIONAL RIGIDITIES

A variable with a distributed lag can be introduced into the demand function in many alternative ways; only a few are considered here. As indicated on page 7, factors that cause distributed lags can be divided into two categories--(1) those related to uncertainty about the future, and (2) those related to technological or institutional rigidities. Alternative models under each of these categories are discussed in detail in this and the following section.

Distinction Between Long- and Short-Run Elasticities

The traditional distinction between long- and short-run elasticities of demand rests primarily on the consideration of technological or institutional rigidities. 6/ Here we implicitly assume that changes in prices or income are expected to be and are permanent changes. In the real world, any change may be thought of as having a permanent and a transitory component. Only the permanent component is actually treated in the traditional analysis of the difference between long-run and short-run demand.

To illustrate, let us begin with the simplest case: Suppose that expectations of prices and income are static, that is, consumers expect current prices and incomes to persist indefinitely. Under an expectation of this kind, all changes in prices or in income are expected to be permanent. Given sufficient time to adjust, for every set of prices and incomes one and only one equilibrium quantity will be demanded. Call this the long-run equilibrium quantity demanded and denote it by x_t^*. The subscript t indicates that the quantity x_t^* is that which would emerge in the long-run if the prices and the income during period t were to persist indefinitely.

The quantity x_t^* is never observed, because another change always occurs before full adjustment to an initial change can take place. Nonetheless, the quantity x_t^* is important to consider because it is the only quantity uniquely determined by prices and income. Figure 4 illustrates this point. Suppose real income and the prices of commodities other than the one in question are both held constant. Suppose, however, that the price of the commodity is varied. Given sufficient time for adjustment, the quantities demanded trace out the curve D_LD_L in figure 4. This curve shows x_t^* as a function of p_t, the price of the commodity in question. Consider the two points on D_LD_L corresponding to the prices p_0 and p_1. Given sufficient time for adjustment, x_0^* is consumed at a price p_0 and x_1^*, at a price p_1. If, however, we suddenly change the price from p_0 to p_1 consumers increase their consumption, not from x_0^* to x_1^*, but from x_0^* to x_1, where $x_1 \leq x_1^*$. Since x_1 differs from x_1^* only on account of the difference in the time allowed for consumers to adjust, it cannot exceed x_1^*. 7/ The point (x_1, p_1) lies on <u>a</u> short-run demand curve D_SD_S. If consumers fully adjust to the change in price, and then the price returns to a level p_0, the quantity demanded will not be x_0^* but $x_0 \geq x_0^*$. The point (x_0, p_0) lies on a different short-run demand curve $D_S'D_S'$. The curve D_SD_S is appropriate only if we start from a position where x_0^* is demanded; the curve $D_S'D_S'$ is appropriate only if we start from a position x_1^*. A different short-run

---

6/ See Marshall (24, pp. 378-379).

7/ The curves D_LD_L and D_SD_S are obtained for an individual consumer by assuming that he maximizes a utility function with certain properties subject to some set of restraints. Under our assumptions the set of restraints is larger the less the time allowed for adjustment. x_1^* represents the <u>maximum</u> amount consumers would demand at a price p_1 provided they have time to adjust; they always have the alternative of consuming less than x_1^*, say x_1.

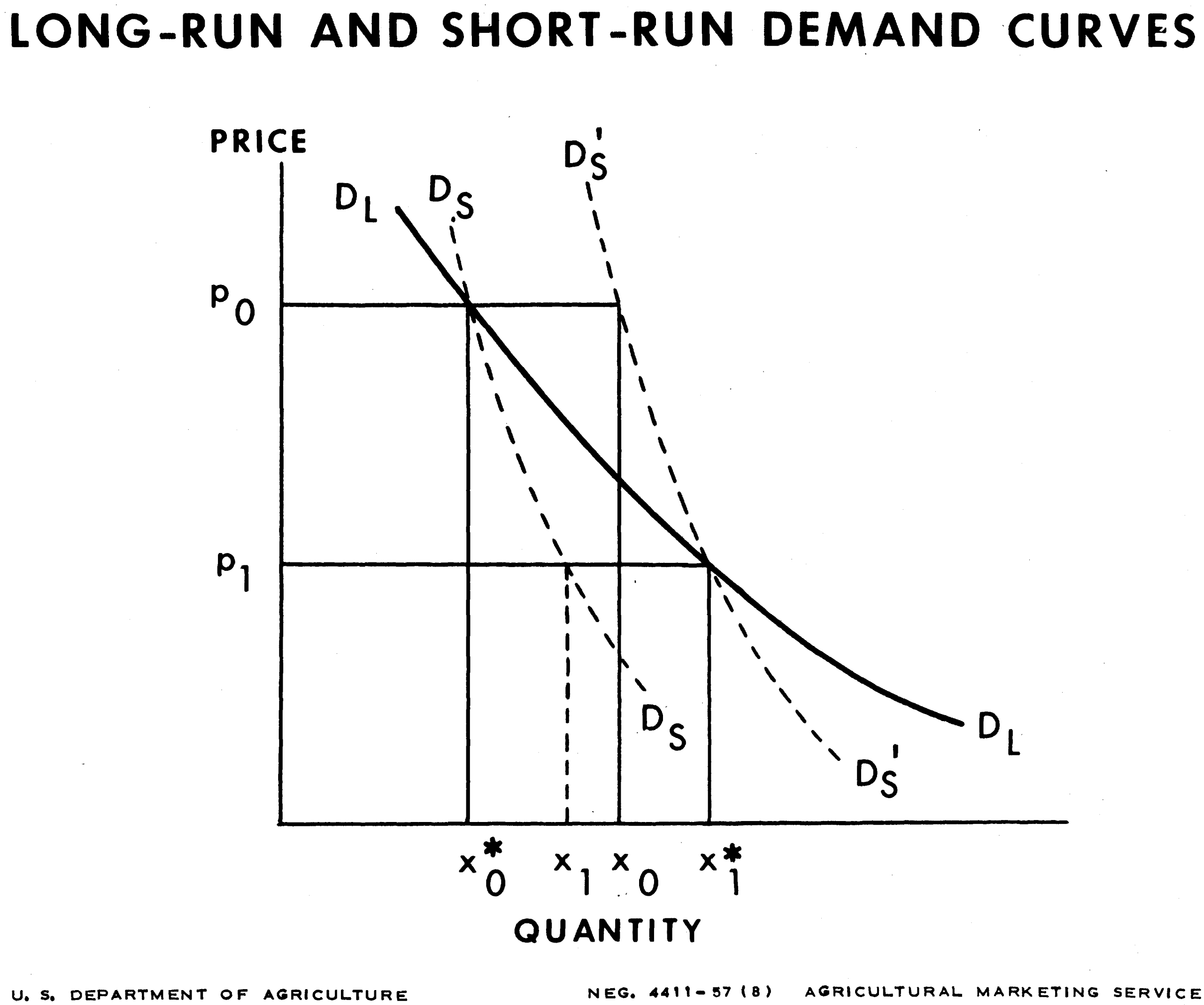


Figure 4.--Diagram that illustrates the relation between the long-run demand curve, D_LD_L, and various short-run demand curves, D_SD_S and $D_S'D_S'$.

demand curve corresponds to each and every point along the long-run demand curve. Although we observe only points lying on short-run demand curves and these curves all differ, each curve is related to the unique long-run curve. Normally we wish to estimate the coefficients of the long-run curve rather than the coefficients of the many varying short-run curves. We can do this if we formulate our relationship properly.

Time Paths Followed in Moving Toward Equilibrium

If we hold real income and other prices constant, we may write

$$x_t^* = a + b\ p_t\ , \tag{21}$$

as a linear approximation to the long-run demand function. The relation between this function and the short-run demand functions may be derived from an assumption about the path which observed consumption would follow as it moves toward its long-run equilibrium level if no further changes in price occurred. Two such paths are plotted in figure 5. If at time 0, x^* becomes the new long-run equilibrium and if $x_0 = x_0^*$, then a curve such as A or B describes the path of x_t to the new equilibrium position. The shape and form of this path is determined by the type of institutional or technological rigidities which exist.

If the rigidity is due primarily to technological factors, we might expect x_t to start rising immediately toward x^*. As x^* is approached, easy substitutions are exhausted so that x_t proceeds more slowly towards x^*. Such a path is described by curve A in figure 5. After t periods have elapsed, consumption is $x_t^{(A)}$ if A is the appropriate time-path. On the other hand, if rigidity is due primarily to factors of an institutional nature, such as impediments to the spread of information concerning the price change, we might expect consumption to change only slowly at first, then more rapidly, and finally slowly again; such a path is described by curve B in figure 5. After t periods, consumption is $x_t^{(B)}$ if B is the appropriate time path. $x_t^{(B)}$ is less than $x_t^{(A)}$, in this case, but after a certain point in time it may exceed $x_t^{(A)}$. The case of a fixed lag of t periods in the adjustment of x_t to x^* could be described by a curve with a horizontal segment lying along the time-axis, a vertical segment at the point t, and another horizontal segment at a level x^* from t on out.

Difference or Differential Equations that Characterize Alternative Time Paths

The time-paths under discussion may be characterized by the nature of their derivatives, $\frac{dx}{dt}$, if time is considered as a continuous variable. If time is considered discrete, analogous difference equations may be formulated. The basic differential equation which we may consider is

$$\frac{dx}{dt} = \delta(t)\,[x^* - x_t] \tag{22}$$

where $\delta(t)$ is some function of time whose range is the interval between zero and one. By varying the function $\delta(t)$ we can produce time-paths of different forms.

For example, let $\delta(t)$ be equal to a constant δ for all t; then the solution to (22) is

$$x_t = (x_0^* - x^*)e^{-\delta t} + x^* \tag{22a}$$

where x_0^* is the initial equilibrium position. The curve described by (22a) is of the same general form as the curve A in figure 5.

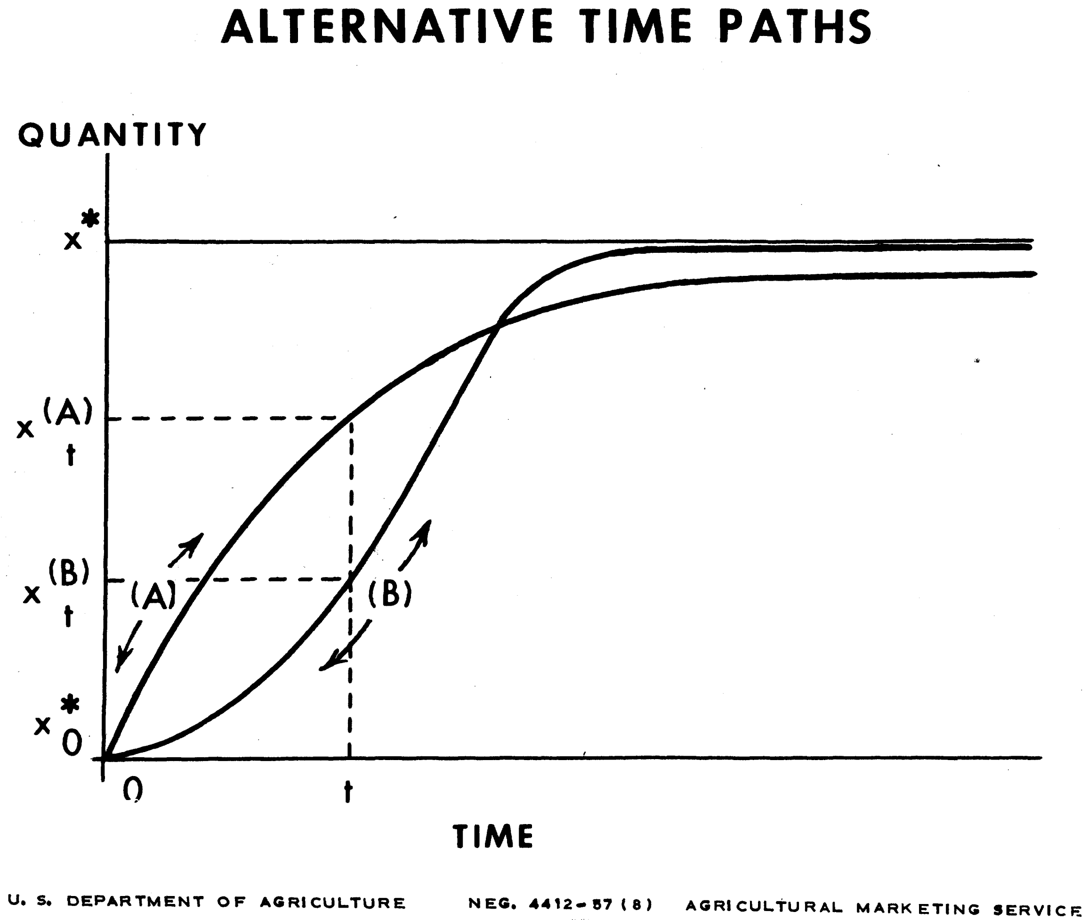


Figure 5.--Alternative time-paths, A and B, of the movement of the observed quantity consumed, x_t, toward the long-run equilibrium quantity, x^*.

The functions $\delta(t)$ which produce curves of the same general form as B in figure 5 are typically rather complicated. One such, for example, is the probability density function of t where t is logarithmically normally distributed. We also could choose $\delta(t)$ in such a way as to obtain the Pearl-Reed logistic curve, which is of the same general form as B. We shall deal here only with the simplest case in which $\delta(t)$ is a constant. It should be emphasized, however, that the appropriate form of the function $\delta(t)$ should be selected on the basis of evidence related to the character of the rigidities in consumer behavior.

Derivation of the Lag Distribution Assumed by Koyck

When $\delta(t)$ is a constant δ, the difference equation analogue of (22) is

$$x_t - x_{t-1} = \delta\,[x_t^* - x_{t-1}],\quad 0 \leq \delta < 1 \tag{23}$$

In this formulation we allow x_t^* to be a function of time. Clearly prices and income are continually changing over time, so that x_t^* is also changing over time. Equation (23) may be solved for x_t as a function of the function x_t^* and the parameter δ; the solution is

$$x_t = \sum_{\lambda=0}^{t} \delta(1-\delta)^{t-\lambda} x_\lambda^* \tag{24}$$

provided period 0 is in the sufficiently distant past. 8/ Equation (24) shows x_t to be a weighted average of the past values of x_t^*, where the weights decline as one goes back in time. The sum of the weights

$$\sum_{\lambda=0}^{t} \delta(1-\delta)^{t-\lambda}$$

equals one since $0 \leq \delta < 1$.

If we retain the simplifying assumption that real income and other prices are held constant, we have by equations (21) and (24)

$$x_t = a + b \sum_{\lambda=0}^{t} \delta(1-\delta)^{t-\lambda} p_\lambda \tag{25}$$

that is, the quantity demanded is a linear function of lagged prices. The distribution of lag is that assumed by Koyck (22), since the coefficient of price lagged k+1 years is $(1-\delta)$ times the coefficient of price lagged k years. (See pages 12-13.) If the demand equation is not (21) but

$$x_t^* = a + b\, p_t + c\, y_t \tag{26}$$

we have

$$x_t = a + b \sum_{\lambda=0}^{t} \delta(1-\delta)^{t-\lambda} p_\lambda + c \sum_{\lambda=0}^{t} \delta(1-\delta)^{t-\lambda} y_\lambda \tag{27}$$

The important thing to note about equation (27) is that the distribution of lag for <u>both</u> income and price is the <u>same</u>. No essential difficulties are introduced if additional prices are added to the demand equation. If the demand equation is

$$x_t^* = a + \sum_{i=1}^{n} b_i\, p_{it} + c\, y_t \tag{28}$$

this reduces to

8/ See Nerlove, Marc. Estimates of the Elasticities of Supply of Corn, Cotton, and Wheat. Ph.D. thesis, Johns Hopkins University, 1956, pp. 52-53.

$$x_t = a + \sum_{i=1}^{n} b_i \sum_{\lambda=0}^{t} \delta(1-\delta)^{t-\lambda} p_{i\lambda} + c \sum_{\lambda=0}^{t} \delta(1-\delta)^{t-\lambda} y_{\lambda} \quad (29)$$

Again all prices and income enter with the same distributed lag.

An even simpler formulation results from equations (23) and (28). If we substitute (28) into (23) we have

$$x_t = a\delta + \sum_{i=1}^{n} b_i \delta p_{it} + c\delta y_t + (1-\delta) x_{t-1} \quad (30)$$

This result is the same as that obtained by Koyck's method of reduction if applied to equation (29). (See pages 12-13.) The usefulness of our procedure, as opposed to Koyck's, is discussed in later paragraphs.

MODELS BASED ON UNCERTAINTY ABOUT THE FUTURE

Let us now assume instead that technological or institutional rigidities are absent, but that uncertainty about the future exists and that habit is a powerful force. Let us further suppose that expectations are not static.

Definition of Terms

Autonomous and induced price expectations.--For simplicity we begin by considering only price expectations and changes in those expectations. Any change in price expectations may be divided into two components: (1) autonomous and (2) induced. 9/

Autonomous changes in price expectations are the result of particular causal factors such as, for example, changes in the levels at which the prices of certain farm commodities are supported or changes in the taxes levied on commodities like liquor or cigarettes. Autonomous changes may also be due to wars, other social upheavals, or natural phenomena. Changes in price expectations due to sudden technological change may also be classed as autonomous.

Part of any change in expected future prices, however, typically is induced by a change in current prices. If people knew in every instance exactly what factors caused a change in current price, no changes in price expectations would be induced: any change in price expectations would be related to the specific causal factors. The typical consumer, however, is not generally aware of the specific reasons for a change in a particular price. He therefore uses past prices as his guide to the future, at least to some extent.

9/ See Enthoven and Arrow (11, p. 289).

Expectations as to the future values of the price of a commodity generally are not single-valued, that is, a consumer's expectations concerning the price of a particular commodity cannot, in general, be represented by a single number, but only by a function of time. Both current price and price expectations may be assumed to affect the quantity demanded for two reasons: (1) The consumer maximizes not only today's satisfaction, but also tomorrow's, and (2) the consumer may suffer a loss in utility or satisfaction if he alters his accustomed habits. In order to analyze consumer behavior, however, we must reduce the problem of price expectations to manageable proportions; single-valued expectations are more manageable than are multi-valued expectations.

Permanent and transitory components of price changes.--Consideration of induced changes in expectations of price leads to an interpretation in terms of single-valued expectations. Any change in current price can be thought of as divided into two components--(1) the permanent component; and (2) the transitory component. The permanent component of a change in current price affects all expected future prices; the transitory component affects only some of them and perhaps none of them. What is permanent and what is transitory clearly depends on the length of the consumer's economic horizon. If the horizon is short, almost all of a change in current price may be supposed to be permanent; on the other hand, if the horizon is long, almost all the change may be supposed to be transitory.

Let the average level about which future prices are expected to fluctuate be called "expected normal price." Expectations about normal price are, of course, single-valued. More important, however, is that, as defined, changes in expected normal price are induced only by the permanent components of changes in current price. The transitory components of changes in current price affect only the deviations of expected future prices about the expected normal price.

If the force of habit is strong, two things result: (1) Changes in the behavior of a consumer due to the permanent component of a change in current price are large compared with changes due to the transitory component. (2) The effects of changes in expected normal price are even stronger as compared with the effects of deviations of expected future prices about expected normal price. It follows that, if the force of habit is strong, the effects of a change in current price on consumer behavior are slight compared with the effects of a change in expected normal price, and the effects of deviations of expected future prices about expected normal price are negligible.

Permanent and transitory components of changes in income.--Similar conclusions apply when income is considered. The grounds are firmer, however, for supposing that the effects of changes in expected normal income are strong compared with changes in current income and even stronger compared with the effects of deviations of expected future incomes about expected normal income. Indeed Friedman (16, Chap. II) has gone so far as to say that changes in expected normal income are the only changes in income that affect aggregate consumption expenditures. 10/ The reason for Friedman's hypothesis is related to his discussion of the tendency of consumers to even out their consumption over long periods of time. (See pages 93-96.) The importance of Friedman's work on the aggregate consumption function goes beyond this point, however, and is discussed in detail beginning on page 93.

Induced changes in the expected normal price of a commodity may depend upon changes in the current prices of commodities other than the one in question. For example, a change in the current price of steel may affect consumers' expectations of the normal price of automobiles even more than a change in the current price of automobiles. For the present, however, we neglect the effect of changes in the current prices of commodities other than the one in question upon that commodity's expected normal price.

Definition of the coefficient or elasticity of expectations.--A model which shows the effect of a change in current price upon expected normal price may be derived from Hick's definition of the elasticity of expectations. 11/ Hicks defines "the elasticity of a particular person's expectations of the price of a commodity x as the ratio of the proportional rise in expected future prices of x to the proportional rise in its current price." Hicks distinguishes two limiting cases: an elasticity of zero, implying no effect of a change in current price upon expected future prices; and an elasticity of one, implying that if prices were previously expected to remain constant, that is, were at their long-run equilibrium level, they are now expected to remain constant at the level of current price. Clearly, an elasticity of expectations equal to one implies "static" expectations in the sense in which the term is used above. By allowing for a range of elasticities between the two extremes, Hicks implicitly recognizes that a particular past-price may have something but not everything to do with people's notion of normal.

Hicks' definition of the elasticity of expectations implies that prices have actually been "normal" until some change occurred, as the change in current price is expressed as a deviation from what prices were up to now. But, of course, we know that conditions are seldom if ever "normal" in the real world; and "normality" itself is a subjective matter. Instead of expressing a change in current price as a deviation from a level of past prices, let us express it as a deviation from what was previously the expected normal price. Let p_t^* be the expected normal price of a particular commodity during period t, let p_t be the actual price of the commodity during t, and let β be the elasticity of expectations. Hicks' definition of β implies

$$\frac{\log p_t^* - \log p_{t-1}^*}{\log p_t - \log p_{t-1}^*} = \beta \qquad (31)$$

10/ Friedman uses the term "permanent income" rather than our term "expected normal income." The use of the longer term in our case is related to the fact that we consider price expectations as well as income expectations.

11/ See Hicks (18, p. 205), Arrow and Nerlove (3), and Nerlove, *op. cit.*, pp. 50-51.

If the prices in equation (31) are expressed, not in logarithms, but in terms of actual values, it seems appropriate to call β, not the elasticity of expectations, but the coefficient of expectations. Thus,

$$p_t^* - p_{t-1}^* = \beta\,[p_t - p_{t-1}^*] \tag{32}$$

where β is now defined as the coefficient of expectations.

Equation (32) exactly expresses the distinction between the permanent and transitory components of a change in current price. If the change in current price is expressed as a deviation from the previous expected normal price, the coefficient of expectation, β, is just the proportion of the change which is regarded as the permanent component of the change; $1-\beta$ is the proportion regarded as the transitory component.

Variables that affect the coefficient of expectations.--The elasticity or coefficient of expectations may be a function of a variety of variables. As indicated on page 21, it is related to the length of a consumer's economic horizon. In times of great social upheaval, war, or revolution, the representative economic horizon may grow very short. Other institutional factors may influence the length of the economic horizon as well, but if they do, these factors are likely to affect equally the elasticities or coefficients of expectation for different commodities.

More important are those factors which may influence the coefficients of expectation for particular commodities while not affecting, or affecting differently, the coefficients for other commodities. For example, the elasticity or coefficient of expectation may be a decreasing function of the typical variance of prices. Consider two markets, one in which fluctuations about some "norm" are of small magnitude and one in which frequent fluctuations of large magnitude are common. In which market will consumers pay most attention to a given price change in forming their expectations of "normal" price? It seems intuitively clear that β will be larger for the first market and smaller for the second. 12/ If a market for a particular commodity changes in character and becomes, say, more like the second market and less like the first, then we might reasonably expect the coefficient of expectation for the price of that commodity to fall.

A model which applies when β is a constant between zero and one.--In spite of the difficulties discussed in the preceding paragraphs, it seems useful, as a first approximation, to assume that β is a constant. We further assume, in light of the relation between β and the concepts of permanent and transitory components, that β lies between zero and one. The model based upon equation (32) and the assumption that β is a constant has been used in the study of supply functions for agricultural commodities; the differential equation analogue of the model has been used by Cagan (5) in his study of hyper-inflations, and by Friedman (16) in his study of the aggregate consumption function.

12/ See Nerlove, *op. cit.*, pp. 54-59.

The model based on equation (32) may also be applied to income. Let y_t^* be expected normal income during period t, let y_t be actual income during t, and let α be the coefficient of income expectation; then, analogously to (32)

$$y_t^* - y_{t-1}^* = \alpha[y_t - y_{t-1}^*], \; 0 \leq \alpha < 1 \tag{33}$$

Equations (32) and (33) are first-order difference equations in expected normal price and expected normal income, respectively. Hence, they may be solved for p_t^* and y_t^* as functions of the functions p_t and y_t, respectively. The solution to (32) is

$$p_t^* = \sum_{\lambda=0}^{t} \beta(1-\beta)^{t-\lambda} \; p_\lambda \tag{34}$$

and the solution to (33) is

$$y_t^* = \sum_{\lambda=0}^{t} \alpha(1-\alpha)^{t-\lambda} \; y_\lambda \quad \underline{13}/ \tag{35}$$

We assume that period 0 is in the distant past.

The current demand for a commodity, x_t, generally depends on both the current and expected normal values of price and income. If all prices other than the price of the commodity in question are held constant, we may write a typical demand function, in linear approximation, as

$$x_t = a + b_0 \, p_t + b_1 \, p_t^* + c_0 \, y_t + c_1 \, y_t^* \tag{36}$$

Substitution of (34) and (35) into (36) yields

$$x_t = a + b_0 \, p_t + b_1 \sum_{\lambda=0}^{t} \beta(1-\beta)^{t-\lambda} \; p_\lambda \; +$$

$$c_0 \, y_t + c_1 \sum_{\lambda=0}^{t} \alpha(1-\alpha)^{t-\lambda} \; y_\lambda$$

$$= a + (b_0 + b_1 \beta)p_t + b_1 \sum_{\lambda=0}^{t-1} \beta(1-\beta)^{t-\lambda} \; p_\lambda \; +$$

$$(c_0 + c_1 \alpha)y_t + c_1 \sum_{\lambda=0}^{t-1} \alpha(1-\alpha)^{t-\lambda} \; y_\lambda \tag{37}$$

13/ See Nerlove, op. cit., pp. 52-53.

Hence, the quantity demanded, x_t, is a function of price and income each taken with a distributed lag. If we neglect the terms in unlagged price and income, each distribution of lag, for price and for income, taken by itself, is similar to Koyck's distribution of lag, discussed on pages 12-13. Each taken by itself is also similar to the distribution of lag discussed beginning on page 18 when expectations were assumed to be static but rigidities of a technological or institutional nature existed. Note, however, that while, in the case of static expectations, we found the distribution of lag to be the _same_ for both price and income, we now have _different_ distributions of lag for price and income.

Reduced equations.--When expectations were assumed to be static and rigidities of a technological nature were introduced, a simple reduction of an equation involving a distributed lag was found to be possible. Thus, equation (29) (see page 20) was reduced to equation (30). The former, (29), involved a distributed lag; the latter, (30), did not. Equation (29) was not reduced directly to equation (30); instead the two equations from which (29) had been derived, equations (23) and (28), were used to derive equation (30). In what follows we call an equation such as (30), which is derived from or related to an equation which does involve a distributed lag, such as (29), a "reduced equation." This term should not be confused with the "reduced forms" used in the theory of estimation of simultaneous equations.

The study of reduced equations that involve distributed lags is important because such reduced equations may be estimated statistically more easily than the parent equations that involve distributed lags. Koyck (_20_) proceeds directly from an equation involving a distributed lag to its reduced version. He does this primarily because he does not formulate a model leading to a distributed lag but directly assumes the existence and form of such a lag. Under the assumption of static expectations, our indirect method of arriving at a reduced equation is only slightly easier than Koyck's direct method. When the distributed lags are of an expectational nature, however, our method is more direct than Koyck's and indeed makes reductions possible which would be impossible if Koyck's method were used. When we come to consider models involving both uncertainty and technological or institutional rigidity, its virtues become even more apparent.

An Indirect Method for Obtaining Reduced Equations

The general demand equation corresponding to (36) is

$$x_t = a + \sum_{i=1}^{n} [b_{0i}\, p_{it} + b_{1i}\, p_{it}^{*}] + c_0\, y_t + c_1\, y_t^{*} \tag{38}$$

where p_{it} is the actual price of the ith commodity and p_{it}^{*} is the expected normal price of the ith commodity. We assume that there are n-1 commodities related in demand to the commodity under consideration. If β_i is the coefficient of expectations for the ith commodity, if

$$p_{it}^{*} - p_{it-1}^{*} = \beta_i \, [p_{it} - p_{it-1}^{*}] \tag{39}$$

for all commodities, and if equation (33) holds, we have, for the general demand equation involving a distributed lag,

$$x_t = a + \sum_{i=1}^{n} \left\{ (b_{0i} + b_{1i} \;\; \beta_i) p_{it} + \sum_{\lambda=0}^{t-1} \beta_i (1-\beta_i)^{t-\lambda} \;\; p_{i\lambda} \right\} +$$

$$(c_0 + c_1 \alpha) y_t + \sum_{\lambda=0}^{t-1} \alpha (1-\alpha)^{t-\lambda} \;\; y_\lambda \tag{40}$$

An equation such as (40) is virtually impossible to reduce directly.

In order to illustrate our indirect method, we may assume, without loss of generality, that

$$a = 0$$

$$b_{0i} = 0 \text{ for all } i,$$

$$\text{and} \quad c_0 = 0 \tag{41}$$

for otherwise we might simply define the variable on the left hand side of (38), not as the quantity demanded, x_t, but as

$$(x_t - a - \sum_{i=1}^{n} b_{0i} \, p_{it} - c_0 \, y_t)$$

The method may best be illustrated by considering only simple cases at first.

<u>When only one expected price varies</u>.--Let all expected prices but one be held constant and let expected real income be held constant, then our demand equation becomes

$$x_t = b_{1i} \, p_{it}^{*} \tag{42}$$

and our expectation function may be written

$$- \beta_i \, p_{it} = - p_{it}^{*} + (1-\beta_i) \, p_{it-1}^{*} \quad \underline{14}/ \tag{43}$$

Lagging both equations (42) and (43) one period we have

$$\left.\begin{aligned} x_{t-1} &= b_{1i} \, p_{it-1}^{*} + 0 \\ - \beta_i \, p_{it-1} &= - p_{it-1}^{*} + (1-\beta_i) \, p_{it-2}^{*} \end{aligned}\right\} \tag{44}$$

<u>14</u>/ Equation (43) may be derived from equation (39) above.

These two equations may be solved for a unique value of p^*_{it-1}. Let

$$\Delta = \begin{vmatrix} b_{1i} & 0 \\ -1 & (1-\beta_i) \end{vmatrix} \quad (45)$$

$$\text{and } \Delta_{(i)} = \begin{vmatrix} x_{t-1} & 0 \\ -\beta_i\, p_{it-1} & (1-\beta_i) \end{vmatrix}$$

Then

$$p^*_{it-1} = \frac{\Delta(i)}{\Delta} \quad (46)$$

so that, substituting (46) in (43), we have

$$-\beta_i\, p_{it} = -p^*_{it} + (1-\beta_i)\frac{\Delta(i)}{\Delta} \quad (47)$$

or

$$p^*_{it} = \beta_i\, p_{it} + (1-\beta_i)\frac{\Delta(i)}{\Delta} \quad (48)$$

Substituting (48) into (42) we have

$$x_t = b_{1i}\,\beta_i\, p_{it} + (1-\beta_i) b_{1i}\frac{\Delta(i)}{\Delta} \quad (49)$$

On expanding the determinants $\Delta_{(i)}$ and Δ, we have a reduced equation

$$x_t = b_{1i}\,\beta_i\, p_{it} + (1-\beta_i)\, x_{t-1} \quad (50)$$

which is just the equation we should have found if we had employed a direct method of reduction. The original demand equation (42) involves <u>one</u> expectational variable and its reduced form, (50), involves <u>one</u> <u>more</u> variable than in the original equation.

<u>When one expected price and real income vary</u>.--Next, consider the case in which one expected price and expected real income are left free to vary. Our system is thus composed of one demand equation and two expectational equations:

$$\left.\begin{aligned} x_t &= b_{1i}\, p^*_{it} + c_1\, y^*_t\,, \\ -\beta_i\, p_{it} &= -p^*_{it} + (1-\beta_i)\, p^*_{it-1}\,, \\ -\alpha\, y_t &= -y^*_t + (1-\alpha)\, y^*_{t-1} \end{aligned}\right\} \quad (51)$$

If we lag the demand equation one and two periods and lag each of the expectational equations one period, we obtain a system of four equations

$$\left.\begin{array}{rcl} x_{t-1} & = & b_{1i}\, p^*_{it-1} + c_1\, y^*_{t-1} + 0 + 0 \\ x_{t-2} & = & 0 + 0 + b_{1i}\, p^*_{it-2} + c_1\, y^*_{t-2} \\ -\beta_i\, p_{it-1} & = & p^*_{it-1} + 0 + (1-\beta_i) p^*_{it-2} + 0 \\ -\alpha\, y_{t-1} & = & 0 + -y^*_{t-1} + 0 + (1-\alpha) y^*_{t-2} \end{array}\right\} \quad (52)$$

Let

$$\Delta = \begin{vmatrix} b_{1i} & c_1 & 0 & 0 \\ 0 & 0 & b_{1i} & c_1 \\ -1 & 0 & (1-\beta_i) & 0 \\ 0 & -1 & 0 & (1-\alpha) \end{vmatrix}$$

$$\Delta(i) = \begin{vmatrix} x_{t-1} & c_1 & 0 & 0 \\ x_{t-2} & 0 & b_{1i} & c_1 \\ -\beta_i\, p_{it-1} & 0 & (1-\beta_i) & 0 \\ -\alpha\, y_{t-1} & -1 & 0 & (1-\alpha) \end{vmatrix} \quad (53)$$

and

$$\Delta' = \begin{vmatrix} b_{1i} & x_{t-1} & 0 & 0 \\ 0 & x_{t-2} & b_{1i} & c_1 \\ -1 & -\beta_i\, p_{it-1} & (1-\beta_i) & 0 \\ 0 & -\alpha\, y_{t-1} & 0 & (1-\alpha) \end{vmatrix}$$

Clearly,

$$p^*_{it-1} = \frac{\Delta(i)}{\Delta} \text{ and } y^*_{t-1} = \frac{\Delta'}{\Delta} \quad (54)$$

Hence, substituting (54) into (51) and solving for x_t, we have

$$x_t = b_{1i}\, \beta_i\, p_{it} + c_1\, \alpha y_t + b_{1i}\, (1-\beta_i) \frac{\Delta(i)}{\Delta} + c_1\, (1-\alpha) \frac{\Delta'}{\Delta} \quad (55)$$

Equation (55) is the reduced equation of the system defined by equations (51); it involves the variables x_t, p_{it}, y_t, x_{t-1}, x_{t-2}, p_{it-1}, and y_{t-1}. The original demand equation involved two expectational variables and its reduced equation involves four more non-expectational variables than expectational variables in the original equation.

The determinants $\Delta_{(i)}$, Δ', and Δ may be expanded to yield an explicit equation in x_t, p_{it}, y_t, x_{t-1}, x_{t-2}, p_{it-1}, and y_{t-1}. If this is done, we find (55) gives

$$\begin{aligned} x_t = {} & b_{1i}\ \beta_i\ p_{it} + c_1\ \alpha\ y_t - \\ & b_{1i}\ (1-\alpha)\ \beta_i\ p_{it-1} - c_1\ \alpha\ (1-\beta_i)\ y_{t-1} + \\ & [(1-\alpha) + (1-\beta)]\ x_{t-1} - \\ & (1-\alpha)\ (1-\beta_i)\ x_{t-2} \end{aligned} \tag{56}$$

This same reduced equation may be found by using a variant of Koyck's method, but reduction is then quite involved. 15/

Application to a general case.--The method of solution in the general case follows quite readily from the simple illustrations discussed above. In our general system we have one demand equation and n+1 expectational equations. We assume that n-1 commodities are related, in demand, to the commodity under consideration. Our general system is

$$\begin{aligned} x_t &= \sum_{i=1}^{n} b_{1i}\ p^*_{it} + c_1\ y^*_t \\ -\ \beta_i\ p_{it} &= -\ p^*_{it} + (1-\beta_i)\ p^*_{it-1} \quad (i=1,\ 2,\ \ldots n) \\ -\ \alpha\ y_t &= -\ y^*_t + (1-\alpha)\ y^*_{t-1} \end{aligned} \tag{57}$$

We now lag the demand equation 1 period, 2 periods, etc., up to n+1 periods; and we lag each of the expectational equations 1, 2, ..., up to n periods. This gives us a system of $n+1 + n(n+1) = (n+1)^2$ equations involving $(n+1)^2$ expectational variables. This system may be written

---

15/ See Nerlove, op. cit., pp. 258-260.

$$x_{t-1} = \sum_{i=1}^{n} b_{1i}\, p^*_{it-1} + c_1 \cdot y^*_{t-1}$$

$$\vdots$$

$$x_{t-n-1} = \sum_{i=1}^{n} b_{1i}\, p^*_{it-n-1} + c_1\, y^*_{t-n-1}$$

$$\left.\begin{aligned} -\beta_i\, p_{it-1} &= -p^*_{it-1} + (1-\beta_i)\, p^*_{it-2} \\ &\vdots \\ -\beta_i\, p_{it-n} &= -p^*_{it-n} + (1-\beta_i)\, p^*_{it-n-1} \end{aligned}\right\} \quad i=1, \ldots, n \tag{58}$$

$$\alpha y_{t-1} = -y^*_{t-1} + (1-\alpha)\, y^*_{t-2}$$

$$\vdots$$

$$\alpha y_{t-n} = -y^*_{t-n} + (1-\alpha)\, y^*_{t-n-1}$$

The relevant determinant of these equations is

$$\Delta = \begin{vmatrix} b_{11} & b_{12} & \ldots & b_{1n} & \ldots\, 0 & \ldots & 0 & c_1 & 0 & \ldots & 0 \\ \vdots & \vdots & & & & & & & & & \\ 0 & 0 & \ldots & 0 & \ldots\, b_{11} & \ldots & b_{1n} & 0 & 0 & \ldots & c_1 \\ -1 & 0 & \ldots & 0 & \ldots (1-\beta_1) & \ldots & 0 & -1 & (1-\alpha) & \ldots & 0 \\ 0 & -1 & & 0 & \ldots\, 0 & \ldots & 0 & 0 & -1 & \ldots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots & \vdots & & \\ 0 & 0 & \ldots & -1 & 0 & \ldots & (1-\beta_n) & 0 & 0 & \ldots & (1-\alpha) \end{vmatrix} \tag{59}$$

The determinant $\Delta_{(i)}$ is Δ with the substitution of the vector

$$(x_{t-1}, \ldots, x_{t-n-1}, -\beta_1 p_{1t-1}, \ldots, -\beta_n p_{nt-1}, \alpha y_{t-1}, \ldots \alpha y_{t-n})'$$

for its ith column, i=1, 2, ..., n. The determinant Δ' is Δ with the substitution of the same vector for its (n^2+n+1)st column. The solution of equations (58) for p^*_{it-1} and y^*_{t-1} may be written

$$p^*_{it-1} = \frac{\Delta(i)}{\Delta} \qquad (i=1, \ldots, n)$$

$$\text{and } y^*_{t-1} = \frac{\Delta'}{\Delta} \tag{60}$$

Consequently, equations (57) reduce to

$$x_t = \sum_{i=1}^{n} b_{1i} \ \beta_i \ p_{it} + c_1 \ \alpha \ y_t +$$

$$\sum_{i=1}^{n} b_{1i} \ (1-\beta_i) \ \frac{\Delta(i)}{\Delta} + c_1 \ (1-\alpha) \ \frac{\Delta'}{\Delta} \qquad (61)$$

Equation (61) is the reduced equation of the demand equation involving distributed lags that could have been derived from equations (57). The original demand equation involved n+1 expectational variables; equation (61) involves $(n+1)^2$ more non-expectational variables than the original equation involved expectational variables. In principle the determinants Δ, $\Delta_{(i)}$, and Δ', could be expanded to yield an explicit equation for x_t.

The method of reduction described in the preceding paragraphs involves the reduction of only one equation, the particular demand equation in which we have an interest. For that reason we may call it the "single-equation" method of reduction. As noted above the reduced equation obtained by the single-equation method contains many variables; in fact, the number of additional variables added to the number already in the equation increases as the square of the number of expectational variables involved. For purposes of statistical analysis, therefore, the single-equation method is only efficient when a small number of expectational variables are involved. Two or three expectational variables are all that can be handled comfortably by the single-equation method of reduction: in the case of two expectational variables, the reduced equation contains six independent non-expectational variables; in the case of three expectational variables, it contains twelve. For cases that involve more than two or three expectational variables, the "multiple-equation" method of reduction is preferred.

The Multiple Equation Method of Reduction

If one commodity is a substitute for or a complement of another, that is, if the two commodities are related in demand, it follows that the second will also be a substitute for or a complement of the first; that is, if the price of one commodity enters the demand equation for another, the price of the latter enters the demand equation for the former. Furthermore, if the quantity demanded of individual commodities is related to income, so is the aggregate consumption. The multiple equation method of reduction of a demand equation involving distributed lags of an expectational nature takes advantage of the existence of all of the equations which involve the same expectational variables. 16/

16/ For an application of this method to supply equations, see Nerlove, *op. cit.*, pp. 260-262.

When only one expected price varies.--Consider first the simple case in which all expected normal prices but one are held constant. We retain the assumption that the quantity demanded depends only on expected normal price and expected normal income and not on current price or current income. This assumption is relaxed in the models discussed beginning on page 39.

In this simple case we must consider two equations, a demand equation and a consumption function, in order to apply the multiple equation method. Let C_t be aggregate real consumption expenditures in period t, and let c_{1n+1} be the marginal propensity to consume out of expected normal income. Neglecting constant terms, our system is

$$\left.\begin{aligned} x_t &= b_{1i}\, p^*_{it} + c_1\, y^*_t \\ C_t &= \quad 0 \quad + c_{1n+1}\, y^*_t \\ p^*_{it} &= \beta_i\, p_{it} + (1-\beta_i)\, p^*_{it-1} \\ y^*_t &= \alpha\, y_t + (1-\alpha)\, y^*_{t-1} \end{aligned}\right\} \qquad (62)$$

where the subscript i indicates that we are varying only the price of the ith commodity. Equations (62) may be written in the form of two matrix equations

$$\left.\begin{aligned} \begin{pmatrix} x_t \\ C_t \end{pmatrix} &= \begin{pmatrix} b_{1i} & c_1 \\ 0 & c_{1n+1} \end{pmatrix} \begin{matrix} p^*_{it} \\ y^*_t \end{matrix} \\ \begin{pmatrix} p^*_{it} \\ y^*_t \end{pmatrix} &= \begin{pmatrix} \beta_i & 0 \\ 0 & \alpha \end{pmatrix} \begin{pmatrix} p_{it} \\ y_t \end{pmatrix} + \left[I - \begin{pmatrix} \beta_i & 0 \\ 0 & \alpha \end{pmatrix} \right] \begin{pmatrix} p^*_{it-1} \\ y^*_{t-1} \end{pmatrix} \end{aligned}\right\} \qquad (63)$$

Lagging the first of these equations one period and multiplying both sides by the inverse of

$$\begin{pmatrix} b_{1i} & c_1 \\ 0 & c_{1n+1} \end{pmatrix}$$

we obtain an expression for (p^*_{it-1}, y^*_{t-1}) which may be substituted in the second of equations (63):

$$\begin{pmatrix} p^*_{it} \\ y^*_t \end{pmatrix} = \begin{pmatrix} \beta_i & 0 \\ 0 & \alpha \end{pmatrix} \begin{pmatrix} p_{it} \\ y_t \end{pmatrix} + \left[I - \begin{pmatrix} \beta_i & 0 \\ 0 & \alpha \end{pmatrix} \right] \begin{pmatrix} b_{1i} & c_1 \\ 0 & c_{1n+1} \end{pmatrix}^{-1} \begin{pmatrix} x_{t-1} \\ C_{t-1} \end{pmatrix} \qquad (64)$$

Equation (64) may be substituted in the first of equations (63) to yield a reduced matrix equation:

$$\begin{pmatrix} x_t \\ C_t \end{pmatrix} = \begin{pmatrix} b_{1i} & c_1 \\ 0 & c_{1n+1} \end{pmatrix} \begin{pmatrix} \beta_i & 0 \\ 0 & \alpha \end{pmatrix} \begin{pmatrix} p_{it} \\ y_t \end{pmatrix} +$$

$$\begin{pmatrix} b_{1i} & c_1 \\ 0 & c_{1n+1} \end{pmatrix} \left[I - \begin{pmatrix} \beta_i & 0 \\ 0 & \alpha \end{pmatrix} \right] \begin{pmatrix} b_{1i} & c_1 \\ 0 & c_{1n+1} \end{pmatrix}^{-1} \begin{pmatrix} x_{t-1} \\ C_{t-1} \end{pmatrix} \qquad (65)$$

This yields the following two equations when expanded

$$\left. \begin{aligned} x_t &= b_{1i} \ \beta_i \ p_{it} + c_1 \ \alpha y_t + (1-\beta_i) \ x_{t-1} + \\ &\qquad [(1-\alpha) - \frac{b_{1i}}{c_{1n+1}} (1-\beta_i)] \ C_{t-1} \\ C_t &= \qquad c_{1n+1} \ \alpha y_t + (1-\alpha) \ C_{t-1} \end{aligned} \right\} \qquad (66)$$

Thus, by the multiple equation method, a demand equation involving two expectational variables and a consumption function involving one expectational variable are reduced to a demand equation involving only two more non-expectational variables than the expectational variables involved in the original demand equation, and a consumption function involving only one more non-expectational variable than the expectational variables involved in original consumption function. This result should be contrasted with the result of the single-equation method when applied to a similar case. (See page 29.)

Application to a general case.--The multiple-equation method may easily be generalized to the case in which each demand equation contains the prices of n commodities. It is only necessary to consider the demand equations for all n commodities. Let x_{it} be the quantity demanded of the ith commodity, then the complete set of n demand equations may be written

$$x_{it} = \sum_{j=1}^{n} b_{1ij} \ p_{jt}^{*} + c_{1i} \ y_t^{*} \qquad (i=1, \ldots, n) \qquad (67)$$

To this we must add a consumption function

$$C_t = \qquad c_{1n+1} \ y_t^{*} \qquad (68)$$

and n+1 expectational equations

$$\left. \begin{aligned} p_{it}^{*} &= \beta_i \ p_{it} + (1-\beta_i) \ p_{it-1}^{*} \qquad (i=1, \ldots, n) \\ y_t^{*} &= \alpha \ y_t + (1-\alpha) \ y_{t-1}^{*} \end{aligned} \right\} \qquad (69)$$

Let

$$\left.\begin{aligned} x_t &= (x_{1t}, x_{2t}, \ldots, x_{nt})' \\ p_t^* &= (p_{1t}^*, p_{2t}^*, \ldots, p_{nt}^*)' \\ p_t &= (p_{1t}, p_{2t}, \ldots, p_{nt})' \\ c &= (c_{11}, c_{12}, \ldots, c_{1n})' \end{aligned}\right\} \tag{70}$$

and let $B = \| b_{1ij} \|$ be the matrix of the coefficients of expected normal prices in equations (67) and β be a matrix with the β_i along the diagonal and zeros elsewhere. Then equations (67) and (68) may be written in matrix form as

$$\begin{pmatrix} x_t \\ C_t \end{pmatrix} = \begin{pmatrix} B & c \\ 0 & c_{1n+1} \end{pmatrix} \begin{pmatrix} p_t^* \\ y_t^* \end{pmatrix} \quad \underline{17}/ \tag{71}$$

Equations (69) may be written in matrix form as

$$\begin{pmatrix} p_t^* \\ y_t^* \end{pmatrix} = \begin{pmatrix} \beta & 0 \\ 0 & \alpha \end{pmatrix} \begin{pmatrix} p_t \\ y_t \end{pmatrix} + \left[I - \begin{pmatrix} \beta & 0 \\ 0 & \alpha \end{pmatrix} \right] \begin{pmatrix} p_{t-1}^* \\ y_{t-1}^* \end{pmatrix} \tag{72}$$

Equations (71) and (72) are of exactly the same form as equations (63), and they may be solved for the reduced versions of the demand equations and the consumption function in exactly the same way. Doing so we have

$$\begin{pmatrix} x_t \\ C_t \end{pmatrix} = \begin{pmatrix} B & c \\ 0 & c_{1n+1} \end{pmatrix} \begin{pmatrix} \beta & 0 \\ 0 & \alpha \end{pmatrix} \begin{pmatrix} p_t \\ y_t \end{pmatrix} +$$

$$\begin{pmatrix} B & c \\ 0 & c_{1n+1} \end{pmatrix} \left[I - \begin{pmatrix} \beta & 0 \\ 0 & \alpha \end{pmatrix} \right] \begin{pmatrix} B & c \\ 0 & c_{1n+1} \end{pmatrix}^{-1} \begin{pmatrix} x_{t-1} \\ C_{t-1} \end{pmatrix} \tag{73}$$

Equation (73) may be reduced to still simpler form. We note that

$$\begin{pmatrix} B & c \\ 0 & c_{1n+1} \end{pmatrix}^{-1} = \begin{pmatrix} B^{-1} & -B^{-1}c\, c_{1n+1}^{-1} \\ 0 & c_{1n+1}^{-1} \end{pmatrix} \tag{74}$$

where $c_{1n+1}^{-1} = {}^1/c_{1n+1}$ in this particular case. Substituting (74) into (73) we have

---

$\underline{17}/$ The reader should note that B and β in this and the equations that follow are matrices and x_t, c, and p_t^* are column vectors. C_t, c_{1n+1}, and y_t^* in this equation are scalers.

$$\begin{pmatrix} x_t \\ c_t \end{pmatrix} = \begin{pmatrix} B\beta & c\alpha \\ 0 & c_{1n+1}\alpha \end{pmatrix} \begin{pmatrix} p_t \\ y_t \end{pmatrix}$$

$$+ \begin{pmatrix} B & c \\ 0 & c_{1n+1} \end{pmatrix} \begin{bmatrix} (I-\beta) & 0 \\ 0 & (1-\alpha) \end{bmatrix} \begin{pmatrix} B^{-1} & -B^{-1} c\, c_{1n+1}^{-1} \\ 0 & c_{1n+1}^{-1} \end{pmatrix} \begin{pmatrix} x_{t-1} \\ c_{t-1} \end{pmatrix}$$

$$= \begin{pmatrix} B & c\alpha \\ 0 & c_{1n+1}\alpha \end{pmatrix} \begin{pmatrix} p_t \\ y_t \end{pmatrix}$$

$$+ \begin{bmatrix} B(I-\beta)B^{-1} & -B(I-\beta)B^{-1} c\, c_{1n+1}^{-1} + c(1-\alpha)c_{1n+1}^{-1} \\ 0 & (1-\alpha) \end{bmatrix} \begin{pmatrix} x_{t-1} \\ c_{t-1} \end{pmatrix} \tag{75}$$

Since the matrix β is not an identity matrix times a scalar, that is, since the coefficients of expectations for the prices of different commodities generally differ, we cannot reduce the terms $B(I-\beta)B^{-1}$ any further.

Since the terms involving $B(I-\beta)B^{-1}$ cannot be further reduced, each reduced demand equation, in addition to n current prices and current income, also generally involves lagged values of the quantities demanded of the n commodities included in the system and the lagged value of aggregate consumption. Hence, each reduced demand equation contains twice as many non-expectational independent variables as the expectational independent variables contained in each original demand equation. Thus if each original demand equation contains n expected normal prices and expected normal income, each reduced demand equation contains, in addition to n+1 current values of prices and income, n+1 lagged values of the quantities demanded and aggregate consumption. This result should be contrasted with the result for a similar case obtained using the single equation method of reduction. (See page 31.)

The single equation method leads to one reduced demand equation containing $(n+1) + (n+1)^2$ independent variables; the multiple equation method leads to n reduced demand equations each containing $2(n+1)$ independent variables. For statistical purposes, the latter are clearly preferable: whenever the number of expectational variables to be included in the demand equation exceeds two or three, the multiple equation method of reduction should be employed.

Separability of Equations

Equations (75) show that the reduced equation of the consumption function is "separable" from the reduced equations of the demand equations in the sense that it involves only lagged consumption in addition to current income. The demand equations, however, are not separable in this sense. This property of separability may be generalized and, as we shall see, the case in which the original demand equations involve non-expectational variables is a special case of generalized separability. The case discussed immediately above is also a special case of generalized separability.

Consider the system of demand equations

$$\xi_t = \Gamma \ \eta_t^* \tag{76}$$

where ξ_t is a vector of quantities demanded, η_t^* is a vector of expectational variables, and Γ is an (n+m) x (n+m) matrix of coefficients. The expectational variables η_t^* are assumed to be generated by expectational equations of the form

$$\eta_t^* = \epsilon \eta_t + (I - \epsilon) \ \eta_{t-1}^* \tag{77}$$

where ϵ is a matrix with ϵ_i (i=1, 2, ... n+m) along the diagonal and zeros elsewhere. We call the system defined by (76) and (77) <u>separable</u> if and only if the matrix Γ can be partitioned in the following way:

$$\Gamma = \begin{pmatrix} B & D \\ 0 & A \end{pmatrix} \tag{78}$$

where B is an nxn matrix, A is an mxm matrix, and D is an nxm matrix.

Corresponding to any partition, such as (78), we may also partition ξ_t, η_t, η_t^*, and ϵ as follows:

$$\xi_t = \begin{pmatrix} x_t \\ z_t \end{pmatrix}$$

$$\eta_t = \begin{pmatrix} p_t \\ q_t \end{pmatrix}$$

$$\eta_t^* = \begin{pmatrix} p_t^* \\ q_t^* \end{pmatrix}$$

$$\text{and} \quad \epsilon = \begin{pmatrix} \beta & 0 \\ 0 & \alpha \end{pmatrix} \tag{79}$$

where z_t is a vector $(z_{1t}, \ldots, z_{mt})'$, q_t is a vector $(q_{1t}, \ldots, q_{mt})'$, q_t^* is a vector $(q_{1t}^*, \ldots, q_{mt}^*)$, and α is an mxm matrix with α_i (i=1, ..., m) along its diagonal and zeros elsewhere. Thus, a separable system may be written

$$\begin{pmatrix} x_t \\ z_t \end{pmatrix} = \begin{pmatrix} B & D \\ 0 & A \end{pmatrix} \begin{pmatrix} p_t^* \\ q_t^* \end{pmatrix} \tag{80}$$

and its expectational equations may be written

$$\begin{pmatrix} p_t^* \\ q_t^* \end{pmatrix} = \begin{pmatrix} \beta & 0 \\ 0 & \alpha \end{pmatrix} \begin{pmatrix} p_t \\ q_t \end{pmatrix} + \begin{bmatrix} (I-\beta) & 0 \\ 0 & (I-\alpha) \end{bmatrix} \begin{pmatrix} p_{t-1}^* \\ q_{t-1}^* \end{pmatrix} \tag{81}$$

Equations (80) and (81) may be reduced by the multiple equation method in exactly the same way as equations (63) were reduced; the result is

$$\begin{pmatrix} x_t \\ z_t \end{pmatrix} = \begin{pmatrix} B & D \\ 0 & A \end{pmatrix} \begin{pmatrix} \beta & 0 \\ 0 & \alpha \end{pmatrix} \begin{pmatrix} p_t \\ q_t \end{pmatrix} +$$

$$\begin{pmatrix} B & D \\ 0 & A \end{pmatrix} \begin{bmatrix} (I-\beta) & 0 \\ 0 & (I-\alpha) \end{bmatrix} \begin{pmatrix} B & D \\ 0 & A \end{pmatrix}^{-1} \begin{pmatrix} x_{t-1} \\ z_{t-1} \end{pmatrix} \tag{82}$$

Since

$$\begin{pmatrix} B & D \\ 0 & A \end{pmatrix}^{-1} = \begin{pmatrix} B^{-1} & -B^{-1}D\,A^{-1} \\ 0 & A \end{pmatrix} \tag{83}$$

equation (82) further reduces to

$$\begin{pmatrix} x_t \\ z_t \end{pmatrix} = \begin{pmatrix} B\beta & D\alpha \\ 0 & A\alpha \end{pmatrix} \begin{pmatrix} p_t \\ q_t \end{pmatrix} +$$

$$\begin{bmatrix} B(I-\beta)B^{-1} & -B(I-\beta)B^{-1}D\,A^{-1} + D(I-\alpha)A^{-1} \\ 0 & A(I-\alpha)A^{-1} \end{bmatrix} \begin{pmatrix} x_{t-1} \\ z_{t-1} \end{pmatrix} \tag{84}$$

Equation (84) clearly indicates that the reduced demand equations for the z_{it} (i=1, ..., m) do not involve the lagged values x_{it-1} (i=1, ..., n). These reduced demand equations are therefore separable in exactly the same sense as the reduced consumption function of the previous example.

A separable system of demand equations possesses an interesting property, as illustrated by the following. Equation (84) may be written as the two equations

$$x_t = (B\ \beta\ ,\ D\ \alpha\ ) \begin{pmatrix} p_t \\ q_t \end{pmatrix} +$$

$$(B(I-\beta)B^{-1},\ -B(I-\beta)B^{-1}DA^{-1} + D(I-\alpha)A^{-1}) \begin{pmatrix} x_{t-1} \\ z_{t-1} \end{pmatrix} \tag{85}$$

and

$$z_t = A\alpha\ q_t + A(I-\alpha)A^{-1}\ z_{t-1} \tag{86}$$

Equation (86) may be solved for q_t:

$$q_t = \alpha^{-1}A^{-1}\ z_t - \alpha^{-1}A^{-1}\ A(I-\alpha)A^{-1}\ z_{t-1}$$

$$= \alpha^{-1}A^{-1}\ z_t - (I-\alpha)A^{-1}\ z_{t-1} \tag{87}$$

Substituting (87) into (85), we have

$$x_t = (B\ \beta\ ,\ DA^{-1}) \begin{pmatrix} p_t \\ z_t \end{pmatrix} - D(I-\alpha)A^{-1}\ z_{t-1} +$$

$$(B(I-\beta)B^{-1},\ -B(I-\beta)B^{-1}\ DA^{-1}) \begin{pmatrix} x_{t-1} \\ z_{t-1} \end{pmatrix} +$$

$$D(I-\alpha)A^{-1}\ z_{t-1}$$

$$= (B\ \beta\ ,\ DA^{-1}) \begin{pmatrix} p_t \\ z_t \end{pmatrix} +$$

$$(B(I-\beta)B^{-1},\ -B(I-\beta)B^{-1}\ DA^{-1}) \begin{pmatrix} x_{t-1} \\ z_{t-1} \end{pmatrix} \tag{88}$$

On the other hand, we might start with equation (80) and solve part of the equations for q_t^* in terms of z_t. Since $z_t = Aq_t^*$, we obtain

$$x_t = (B,\ D) \begin{pmatrix} p_t^* \\ A^{-1}z_t \end{pmatrix} = (B,\ DA^{-1}) \begin{pmatrix} p_t^* \\ z_t \end{pmatrix} \tag{89}$$

Application of the multiple equation method of reduction to equation (89), with expectational equations

$$p_t^* = \beta\ p_t + (I-\beta)\ p_{t-1}^*$$

leads to equation (88).

If, therefore, a system of demand equations is separable, it does not matter whether we separate the original equations by the elimination of some expectational variables or whether we separate the reduced equations by elimination of some current variables. 18/ Thus, in the example involving the consumption function discussed beginning on page 32, we could proceed in two ways: (1) as we did, by reducing the whole system, including the consumption function; or (2) by solving the consumption function for expected income, substituting this solution into the demand equations, and then reducing only the demand equations without bothering about the consumption function.

Modifications When Current Price and Income Enter the Demand Equations

In most of the preceding discussion we have maintained the assumption that current values of price and income do not enter the demand equations; that is, we have assumed for each commodity that the coefficients b_{0i} and c_0, in equation (39) (see page 26), are zero for all i. If current prices and current income do enter the demand equations, then b_{0i} (i=1, ..., n) and c_0 cannot be assumed to equal zero.

A system of demand equations involving both current and expected normal values of prices and income may be written in matrix form as

$$x_t = B_0 p_t + B_1 p_t^* + c_0 y_t + c_1 y_t^* \tag{90}$$

where B_0 is a matrix of coefficients, b_{0ij}, of current prices; B_1 is a matrix of coefficients, b_{1ij}, of expected normal prices; c_0 is a vector of coefficients, c_{0i}, of current income; and c_1 is a vector of coefficients, c_{1i}, of expected normal income. The subscript i refers to the ith commodity, and, for simplicity, we again assume that all constant terms in the demand equations are zero. To make the system complete we must, of course, add a consumption function and a set of expectational equations such as (72). The complete system thus obtained is a special case of a system in which the demand equations involve non-expectational variables as well as expectational variables. We shall show that such a system is, in turn, a special case of a separable system (as defined by equations (80) and (81)).

Suppose we have a system of demand equations that express the quantity demanded of each commodity as a linear function of n expectational variables and m non-expectational variables. Let ξ_t be a vector of quantities demanded, η_t^* be a vector of expectational variables, ζ_t be a vector of non-expectational variables, Γ be a matrix of coefficients of the expectational variables, and Δ be a matrix of the coefficients of the non-expectational variables. Then the system of demand equations may be written

$$\xi_t = \Gamma \eta_t^* + \Delta \zeta_t \tag{91}$$

18/ A statistical difference, however, is involved in the two approaches. See the material beginning on page 64.

If β is a matrix with coefficients of expectation along its diagonal and zero elsewhere, we may write the n expectational equations

$$\eta_t^* = \beta \eta_t + (I-\beta) \eta_{t-1}^* \tag{92}$$

where η_t is the vector of current values of the expectational variables.

In order to show that (91) and (92) define a special sort of separable system, we introduce a dummy variable ζ_t^* defined by the equation

$$\zeta_t^* = \zeta_t \tag{93}$$

If we add equation (93) to equations (91) and (92), we have

$$\begin{pmatrix} \xi_t \\ \zeta_t \end{pmatrix} = \begin{pmatrix} \Gamma & \Delta \\ 0 & I \end{pmatrix} \begin{pmatrix} \eta_t^* \\ \zeta_t^* \end{pmatrix} \tag{94}$$

and

$$\begin{pmatrix} \eta_t^* \\ \zeta_t^* \end{pmatrix} = \begin{pmatrix} \beta & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} \eta_t \\ \zeta_t \end{pmatrix} + \begin{pmatrix} (I-\beta) & 0 \\ 0 & (I-I) \end{pmatrix} \begin{pmatrix} \eta_{t-1}^* \\ \zeta_{t-1}^* \end{pmatrix} \tag{95}$$

The system defined by (94) and (95) is clearly separable.

Comparison with equations (80) and (81) show that the following substitutions yield (94) and (95): $x_t = \xi_t$, $z_t = \zeta_t$, $p_t^* = \eta_t^*$, $q_t^* = \zeta_t^*$, $B = \Gamma$, $D = \Delta$, $A = I$, and $\alpha = I$. Consequently, from (84), we see that the reduced equation of (94) is

$$\begin{pmatrix} \xi_t \\ \zeta_t \end{pmatrix} = \begin{pmatrix} \Gamma\beta & \Delta \\ 0 & I \end{pmatrix} \begin{pmatrix} \eta_t \\ \zeta_t \end{pmatrix} + \begin{bmatrix} \Gamma(I-\beta)\Gamma^{-1} & -\Gamma(I-\beta)\Gamma^{-1}\Delta \\ 0 & 0 \end{bmatrix} \begin{pmatrix} \xi_{t-1} \\ \zeta_{t-1} \end{pmatrix} \tag{96}$$

that is,

$$\xi_t = \Gamma \beta \eta_t + \Delta \zeta_t + \Gamma (I-\beta)\Gamma^{-1} \xi_{t-1} - \Gamma(I-\beta)\Gamma^{-1}\Delta \zeta_{t-1} \tag{97}$$

The reduced version of the system defined by equations (90) and (72), that is, the system in which the demand equations include current values as well as expected normal values, follows as a special case of equation (97); it is only necessary to substitute (B_1, c_1) for Γ, (B_0, c_0) for Δ, η_t and ζ_t for $(p_t, y_t)'$, and to let the matrix of coefficients of expectations be

$$\begin{pmatrix} \beta & 0 \\ 0 & \alpha \end{pmatrix}$$

rather than β . From (97) we have, therefore, that the reduced equation of the system (90) and (72) is

$$x_t = (B_1 \; c_1) \begin{pmatrix} \beta & 0 \\ 0 & \alpha \end{pmatrix} \begin{pmatrix} p_t \\ y_t \end{pmatrix} + (B_0 \; c_0) \begin{pmatrix} p_t \\ y_t \end{pmatrix} +$$

$$(B_1 \; c_1) \begin{pmatrix} (I-\beta) & 0 \\ 0 & (1-\alpha) \end{pmatrix} (B_1 \; c_1)^{-1} \begin{pmatrix} p_{t-1} \\ y_{t-1} \end{pmatrix} -$$

$$(B_1 \; c_1) \begin{pmatrix} I-\beta & 0 \\ 0 & (1-\alpha) \end{pmatrix} (B_1 \; c_1)^{-1} \; (B_0 \; c_0) \begin{pmatrix} p_{t-1} \\ y_{t-1} \end{pmatrix}$$

$$= (B_1 \; \beta + B_0, \; c_1 \; \alpha + c_0) \begin{pmatrix} p_t \\ y_t \end{pmatrix} +$$

$$(B_1 \; c_1) \begin{pmatrix} (I-\beta) & 0 \\ 0 & 1-\alpha \end{pmatrix} (B_1 \; c_1)^{-1} \; [I-(B_0 \; c_0)] \begin{pmatrix} p_{t-1} \\ y_{t-1} \end{pmatrix} \qquad (98)$$

MODELS THAT INVOLVE BOTH RIGIDITIES AND UNCERTAINTY

Reduction of equations with distributed lags to equations without distributed lags is of importance in connection with statistical estimation. Our discussion of reduction illustrates two points that are crucial to the understanding of distributed lags: (1) Reduction of an equation with a distributed lag due to technological or institutional rigidities alone can be accomplished simply (see equation (30), page 20). (2) Reduction of equations involving distributed lags of an expectational nature is much more complicated, but a considerable degree of simplification may be obtained through the multiple-equation method of reduction. The reason for the great simplicity of reduction in the case of distributed lags of a non-expectational nature is the fact that <u>the distribution of lag for every variable entering the equation with a distributed lag is the same</u>. The simplicity of reduction in the case of distributed lags of an expectational nature through the use of the multiple-equation method is possible only because <u>the distributions of lag for the same variable in different equations are the same, even though the distributions for different variables in the same equation are not</u>. When distributed lags are due both to uncertainty and rigidities of technological or institutional nature, the distribution of lag is both different for different variables in the same equation and different for the same variable in different equations; hence, no simplicity of reduction is possible in general. If, however, the distributions of lag are of a special form, simple reduction is still possible.

When Current Price Does Not Affect the Long-Run Equilibrium Demand

First model.--Suppose that real income and other prices, current and expected, are held constant. If we suppose that the current price of the commodity in question does not affect the long-run equilibrium demand for it, we have, in this simple case

$$x_t^* = b\ p_t^* \tag{99}$$

where x_t^* is the long-run equilibrium quantity demanded, and p_t^* is the expected normal price of the commodity in question. We suppose, without loss of generality, that the constant term in the demand equation is zero. We may also suppose that the expected normal price is related to past actual prices by an equation of the same form as equation (32) (see page 23).

Although current price does not affect the long-run equilibrium quantity demanded, we cannot assume that it does not effect the current demand. Consequently, we cannot use an equation of the form of equation (23) (see page 18) to express the relation between the current quantity demanded and the long-run equilibrium quantity demanded. Instead, we suppose that the change in current demand depends not only upon the difference between the actual and long-run equilibrium quantities demanded but also upon the difference between current price and expected normal price, that is,

$$x_t - x_{t-1} = \delta_0[p_t - p_t^*] + \delta_1[x_t^* - x_{t-1}] \tag{100}$$

If current price is greater than expected normal price, we might expect the adjustment of current demand to long-run equilibrium demand to be slower than if current price were less than expected normal price, with a given difference between long-run equilibrium demand and current demand; hence, we assume that δ_0 is negative or zero, but not positive.

Equations (32), (99), and (100) lead to a relationship between the current quantity demanded and price taken with a distributed lag. The distribution of lag in this case is complicated. Equations (32) and (100) each are first-order difference equations: the solution to (32) is given in equation (34) (see page 24); the solution to (100) is

$$x_t = \sum_{\lambda=0}^{t} (1-\delta_1)^{t-\lambda}\ [\delta_1 x_\lambda^* + \delta_0 (p_\lambda - p_\lambda^*)] \tag{101}$$

Substituting (34) into (99) and (101), and substituting the result from (101) into the result from (99), we have

$$x_t = \sum_{\lambda=0}^{t} \left\{ (1-\delta_1)^{t-\lambda} \left[\delta_1 b \sum_{\mu=0}^{\lambda} \beta(1-\beta)^{\lambda-\mu} p_\mu + \delta_0 p_\lambda - \delta_0 \sum_{\mu=0}^{\lambda} \beta(1-\beta)^{\lambda-\mu} p_\mu \right] \right\}$$

$$= \sum_{\lambda=0}^{t} (1-\delta_1)^{t-\lambda} \delta_0 p_\lambda + \sum_{\lambda=0}^{t} \left\{ (1-\delta_1)^{t-\lambda} (\delta_1 b - \delta_0) \sum_{\mu=0}^{\lambda} \beta(1-\beta)^{\lambda-\mu} p_\mu \right\} \qquad (102)$$

Clearly, the distribution of lag is complicated. The reduced form of (102) is not, however, particularly complicated. Substituting (100) into (99) we have

$$x_t = (b\ \delta_1 - \delta_0)\ p_t^* + \delta_0\ p_t + (1-\delta_1)\ x_{t-1} \qquad (103)$$

Applying equation (97) (see page 40), with the appropriate substitutions, we have

$$x_t = [b\ \beta\ \delta_1 + (1-\beta)\ \delta_0]\ p_t - (1-\beta)\ \delta_0\ p_{t-1} + [(1-\beta) + (1-\delta_1)]\ x_{t-1} - (1-\beta)\ (1-\delta_1)\ x_{t-2} \qquad (104)$$

Demand equations which contain more than one expected normal price and/or expected normal income offer no further complications. Substitution of the long-run demand equations into the relations between the change in current demand and the difference between long-run equilibrium and current demand yields a system of equations in which expectational and non-expectational variables enter. These equations, along with the corresponding expectational equations, may be reduced without difficulty by the multiple equation method.

Second model.--A formulation, alternative to equation (100), of the relation between the rate of change in current demand and the differences between current and expected normal price and between current and long-run equilibrium demand may be stated: In our discussion of rigidities due to institutional or technological causes, we expressed the derivative of the time-path of the movement of the current quantity demanded toward the long-run equilibrium quantity as

$$\frac{dx}{dt} = \delta(t)\ [x^* - x]$$

or in first-difference form as

$$x_t - x_{t-1} = \delta (t) [x_t^* - x_{t-1}]$$

Equation (100) arose out of a specific choice of the function $\delta (t)$, namely

$$\delta (t) = \frac{\delta_1 + \delta_0 [p_t - p_t^*]}{[x_t^* - x_{t-1}]} \tag{105}$$

Another reasonable choice for $\delta (t)$ might be

$$\delta (t) = \delta_1 + \delta_0 [p_t - p_t^*] \tag{106}$$

Equation (106) thus leads to

$$x_t - x_{t-1} = (\delta_1 + \delta_0 [p_t - p_t^*]) [x_t^* - x_{t-1}] \tag{107}$$

Equation (107) suggests that if the difference between current and long-run equilibrium demand is zero, no change in current demand is to be expected even if current price is not equal to expected normal price. On the other hand, equation (100) suggests that even if current demand is equal to long-run equilibrium demand, some change may be expected in current demand as long as current price differs from expected normal price. If we suppose that current prices do influence demand, even though that influence may be slight, the formulation expressed by equation (100) seems somewhat more reasonable than that expressed by equation (107). In any case, models which incorporate equations of the form of (107) are exceedingly difficult to reduce by the methods outlined thus far.

<u>An identification problem</u>.--If the influence of current price or income on current demand is negligible, an interesting "identification" problem arises. Suppose that expected real income and the expected normal prices of all commodities except the one in which we are interested are held constant; then we have a system consisting of equations (99), (32), and (23), that is, we have:

$$x_t^* = b\ p_t^* \tag{99}$$

$$p_t^* - p_{t-1}^* = \beta [p_t - p_{t-1}^*] \tag{32}$$

$$x_t - x_{t-1} = \delta [x_t^* - x_{t-1}] \tag{23}$$

Comparison of (23) with either (100) or (107) shows that we assume that $\delta_0 = 0$ and $\delta_1 = \delta$. Solving the difference equations (32) and (23) and making the appropriate substitutions we have

$$x_t = b\,\beta\,\delta \sum_{\lambda=0}^{t} (1-\delta)^{t-\lambda} \sum_{\mu=0}^{\lambda} (1-\beta)^{\lambda-\mu}\, p_\mu$$

$$= b\,\beta\,\delta \Big\{ p_t + [(1-\beta) + (1-\delta)]\, p_{t-1} + [(1-\beta)^2 + (1-\beta)(1-\delta) + (1-\delta)^2]\, p_{t-2} + [(1-\beta)^3 + (1-\beta)^2(1-\delta) + (1-\beta)(1-\delta)^2 + (1-\delta)^3]\, p_{t-3} + \ldots \Big\} \quad (108)$$

Comparison of equation (108) with equation (102) (see page 43) indicates that a much simpler distribution of lag arises when the effect of current price is assumed to be negligible. A difficulty also arises, however.

The important thing to note about equation (108) is that β and δ enter into the expression for x_t in terms of past prices in a way that is exactly symmetric. The particular value of β which we observe is the result of the behavior of consumers when confronted by uncertainty about the future. If the conditions under which that uncertainty exists change, then so will the value of the parameter β. Similarly, if technological or institutional factors change, the parameter δ may be expected to change. The fact that β and δ enter the relationship between current quantity demanded and past prices in a symmetrical way means that we cannot distinguish the effects of changes in β from the effects of changes in δ.

The two need to be separated, however, inasmuch as the relationship of current demand to past prices differs for different products and changes in different ways depending on which type of lag predominates, the one due to uncertainty or the one due to technological or institutional rigidities. If the first type of lag were more important, we might expect differences among products in the relationship between current demand and past prices to depend in large part upon the characteristics of the market which consumers face and upon the strength of their habits with regard to the individual commodities. In this case, changes in these factors produce changes in the relationship, but changes in technological or institutional factors have little effect. On the other hand, if the second type of lag were more important, differences in the relationships among products depend on the ease with which one commodity may be substituted for another and other technological or institutional factors.

Take, for example, an agricultural product like apples in comparison with an industrial product like refrigerators. Presumably, the typical variance of apple prices about some normal level is greater than the variance of refrigerator prices, but apples can be more easily substituted for other commodities than can refrigerators. If the distributed lag is due primarily to uncertainty, we might expect the distribution of lag to have a greater variance for apples than for refrigerators; on the other hand, if the distributed lag is due primarily to technological rigidities, we might expect the variance of the distribution of lag to be greater for refrigerators than for apples.

The fact that β and δ enter (108) symmetrically thus leads to a problem of identification which is similar to the type of identification problem considered in the theory of estimation of simultaneous equation systems.

Exactly the same identification problem arises when we consider the reduced version of equation (108), that is, of the system (99)-(32)-(23). If (99) is substituted into (23), the result may be reduced by the multiple equation method. The general form of the result is given by equation (97) (see page 40); the specific reduced equation turns out to be

$$x_t = b\beta\delta\, p_t + [(1-\beta) + (1-\delta)]\, x_{t-1} - (1-\beta)(1-\delta)\, x_{t-2} \qquad (109)$$

Again, β and δ enter this expression symmetrically. 19/

When Current Price Does Affect the Long-Run Equilibrium Demand

Suppose we have a system of demand equations

$$\xi_t^* = \Gamma\, \eta_t^* \qquad (110)$$

where ξ_t^* is a vector of long-run equilibrium quantities demanded or other such variables, η_t^* is a vector of expectational variables, and Γ is a matrix of coefficients. The expectational equations correspond to equations (77) (see page 36); they are

$$\eta_t^* = \epsilon\, \eta_t + (I-\epsilon)\, \eta_{t-1}^* \qquad (77)$$

where ϵ is a diagonal matrix of coefficients of expectations. The relations between current demand and long-run equilibrium demand for the ith commodity may be written

$$\xi_{it} = \rho_i\, \xi_{it}^* + (1-\rho_i)\, \xi_{it-1}$$

Hence, the complete set may be written in matrix form as

$$\xi_t = R\, \xi_t^* + (I-R)\, \xi_{t-1} \qquad (111)$$

where ξ_t is a vector of the current quantities demanded. Substituting (110) into (111), we have

$$\xi_t = R\Gamma\, \eta^* + (I-R)\, \xi_{t-1} \qquad (112)$$

19/ Equation (104) (see page 43) presents much the same problem, only it is not so apparent. Further discussion of the type of identification problem posed by (104) is left to the section on methods of statistical estimation.

Applying equation (97) to the system consisting of equations (112) and (77) we have

$$\begin{aligned}\xi_t &= R\,\Gamma\,\epsilon\,\eta_t + (I-R)\,\xi_{t-1} + \\ &\quad R\,\Gamma\,(I-\epsilon)\Gamma^{-1}R^{-1}\,\xi_{t-1} - \\ &\quad R\,\Gamma\,(I-\epsilon)\Gamma^{-1}R^{-1}\,(I-R)\,\xi_{t-2} \\ &= R\,\Gamma\,\epsilon\,\eta_t + [(I-R) + R\,\Gamma\,(I-\epsilon)\Gamma^{-1}R^{-1}]\,\xi_{t-1} \\ &\quad - R\,\Gamma\,(I-\epsilon)\Gamma^{-1}R^{-1}\,(I-R)\,\xi_{t-2} \qquad (113)\end{aligned}$$

In this case, the matrices R and ϵ <u>do</u> <u>not</u> enter equation (113) symmetrically; hence, an identification problem of the sort described previously does not arise. Other difficulties do arise, however, and are discussed in the next section.

It is clear that the method of reduction used in the preceding example applies to separable systems in general, and in particular to systems of demand equations which contain non-expectational variables; hence, the method also applies when the difference between current and expected normal price has an effect on the rate of change of current demand (see equation (100) on page 42).

METHODS OF STATISTICAL ESTIMATION

In this section we discuss two methods of estimating the parameters of demand equations involving distributed lags: the first deals directly with an equation involving a distributed lag; the second, with the reduced equation of such an equation or system of equations.

Consider the simple model of demand based on equations (21) and (23):

$$x_t^* = a + b\,p_t + u_t \qquad (21)$$

$$x_t - x_{t-1} = \delta\,[x_t^* - x_{t-1}],\ 0 \leq \delta < 1 \qquad (23)$$

Note that we have added a residual term u_t to equation (21) to represent the factors unaccounted for; the result is to make the system (21) and (23) a <u>statistical</u> model as well as a mathematical one. 20/ As shown on pages 16-18 equations (21) and (23) lead to the following demand equation containing a distributed lag:

20/ In private correspondence, Professor L. R. Klein of the Oxford Institute of Statistics has suggested that a residual term be added to the adjustment equation (23) as well as to the basic demand equation. His suggestion applies equally to the expectational models discussed beginning on page 20. Such an addition may introduce statistical problems of considerable complexity. It is hoped that subsequent research will be devoted to these problems.

$$x_t = a + b \sum_{\lambda=0}^{t} \delta(1-\delta)^{t-\lambda} p_\lambda + w_t \tag{25}$$

where

$$w_t = \sum_{\lambda=0}^{t} \delta(1-\delta) u_\lambda$$

Thus x_t is obtained as a function of past prices. This is a linear function, but the parameter δ does not enter it linearly, that is, the coefficients of past prices are not themselves linear functions of the parameter δ.

A Method That Makes Direct Use of Equations That Involve Distributed Lags

The first of the two methods discussed here attempts to estimate the parameters a, b, and δ directly in the context of an equation such as (25); the method is similar to the short-cut method described by Fisher (15) for calculating a distributed lag. 21/

If δ were known, a and b could be estimated easily. Since $0 < \delta \leq 1$, the size of the weights for prices in the very distant past is negligible. In fact the sum of the weights for the first N past observed prices is

$$1 - (1-\delta)^{N+1} = S_N$$

and this can be made as close to one as we please by taking N large enough. S_N clearly depends on the size of δ : for very large δ, that is, δ close to one, only a few past prices need be included in order that S_N differ from one by an arbitrarily small amount. Thus, depending on the size of δ, all prices after a certain point N years back receive negligible weight in the aggregate.

Hence, if δ is known the expression $\sum_{\lambda=0}^{t} \delta(1-\delta)^{t-\lambda} p_\lambda$ may be computed from past prices with any desired degree of accuracy.

If δ were known, the parameters a and b of equation (25) could be estimated by ordinary least squares provided the w_t were independently distributed. Since δ is not known, another method must be employed.

Let us denote the weighted average of current and past prices at time t, $\sum_{\lambda=0}^{t} \delta(1-\delta)^{t-\lambda} p_\lambda$, by $\bar{p}(\delta)$. Let b* be a vector (a, b), and ξ be a

---

21/ See page 10. The method discussed here is also used by Cagan (5) in his study of hyper-inflations.

vector of observations on the observed quantity demanded and current and past prices. Provided the w_t are normally and independently distributed, say with mean zero and variance σ^2, the parameters b* and β can be estimated by maximum-likelihood methods. 22/

Derivation of maximum likelihood coefficients.--The likelihood function to be maximized may be written

$$L(b^*, \delta ; \xi) = \prod_{t=1}^{T} \frac{1}{\sqrt{2\pi\sigma}} \exp\left\{-\frac{1}{2\sigma^2}[x_t - a - b\,\bar{p}_t(\delta)]^2\right\} \tag{114}$$

where T is the total number of observations. L is a complicated function of b* and δ, and it is not possible to solve analytically the equations obtained by setting the derivatives of log L with respect to b* and δ equal to zero to obtain $\hat{b}^*$ and $\hat{\delta}$. Resort must be had to so-called stepwise maximization of the function L. 23/

Let us assume that L has a unique maximum with respect to b* and δ for every value of ξ. Suppose now that we maximize L with respect to b* for each δ given ξ. We thus obtain $\hat{b}^*$ as a function of δ and ξ:

$$\hat{b}^* = \hat{b}^*(\delta ; \xi) \tag{115}$$

where $\hat{b}^*$ is the value of b* that maximizes L given δ and ξ. Substituting (115) into (114) we obtain

$$L(\hat{b}^*(\delta ; \xi), \delta ; \xi) \equiv M(\delta ; \xi) \tag{116}$$

where M is a function of δ and ξ only. The definition of M implies that

$$M(\delta ; \xi) \geq L(b^*, \delta ; \xi), \text{ for all } b^*, \delta, \text{ and } \xi. \tag{117}$$

Now let us maximize M with respect to δ, obtaining a value of δ, $\hat{\delta}(\xi)$, such that

$$M(\hat{\delta}(\xi); \xi); \geq M(\delta ; \xi), \text{ for all } \delta \text{ and } \xi. \tag{118}$$

Comparison of (118) and (117) indicates that if we define

$$\hat{\hat{b}}^*(\xi) \equiv \hat{b}^*(\hat{\delta}(\xi); \xi) \tag{119}$$

we have

$$L(\hat{\hat{b}}^*(\xi), \hat{\delta}(\xi); \xi) \equiv M(\hat{\delta}(\xi); \xi) \geq L(b^*, \delta ; \xi) \tag{120}$$

22/ The discussion that follows is adapted from Nerlove, op. cit., pp. 248-256.

23/ This procedure is discussed in Koopmans and Hood (21, pp. 156-158).

for all b*, δ , and ξ . Clearly, then, the values $\hat{\hat{b}}$* and $\hat{\delta}$ obtained by stepwise maximization are the same as those obtained by simultaneous maximization of L with respect to b* and δ .

The value of b* which maximizes L given δ is simply the vector of least squares estimates of a and b in equation (25). For each value of δ we may compute these estimates and obtain a value of the function $\hat{b}*(\delta ; \xi)$. Substitution of these values into L yields the function (M(δ ; ξ) which we seek to maximize with respect to δ . But this must be done numerically. Ideally, one should select successive values of δ in such a way that, starting from any $\delta^{(0)}$, the successive values of δ chosen converge to the value $\hat{\delta}$ for which the function M reaches a maximum. Gradient methods of maximization, as described by Chernoff and Divinsky (6, pp. 246-49), do just this, but they involve evaluating the first and second derivatives of the function M with respect to δ at the value $\delta^{(i)}$ for the ith step of the successive approximations to $\hat{\delta}$. When the great complexity of M as a function of δ is recalled, evaluation of derivatives of M for particular values of δ does not appear feasible.

Two alternatives are open: (1) We might compute values of M for particular values of δ each differing from the preceding value by some small amount and find $\hat{\delta}$ by graphical interpolation. This procedure is feasible in the case under consideration, since δ is not a vector. Later we consider cases, however, where a parameter such as δ may conveniently be considered as a vector consisting of different components. In this case, graphical interpolation is not practical. (2) Instead, we might choose a small positive number e. Then at each stage of the approximation we might increase or decrease each component of δ by an amount depending on the product of e and the preceding increase of M divided by the increase or decrease of the component of δ in question. The vector of ratios is actually an approximation to the derivative of M with respect to δ , so that this procedure is an approximation to a strict gradient maximization.

In order to maximize M with respect to δ , we need not compute values of M for different values of δ . The function M is simply the likelihood function L in which the least squares estimates of b*, for the given value of δ , is substituted. Now the logarithm of M is a monotonic increasing function of M, so that log M reaches a maximum at the same value of δ for which M reaches a maximum. But

$$\log M = -\frac{T}{2}\log 2\pi\sigma^2 - \frac{1}{2\sigma^2}\sum_{t=1}^{T}[x_t - \hat{a} - \hat{b}\,\overline{p}_t(\delta)]^2$$

Hence, M reaches a maximum when

$$\sum_{t=1}^{T}[x_t - \hat{a} - \hat{b}\,\overline{p}_t(\delta)]^2 \tag{121}$$

reaches a minimum. But (121) is the expression for the residual sum of squares calculated by means of the least squares regression between x_t and $\bar{p}_t(\delta)$ for a given δ. Hence M reaches a maximum for that value of δ which maximizes the correlation between x_t and $\bar{p}_t(\delta)$.

Use of the likelihood ratio test.--Strictly speaking, it is impossible in general to obtain confidence intervals for the estimates of the parameters of a population distribution of known or assumed mathematical form, unless the sampling distributions of those estimates are known. 24/ The estimates of a, b, and δ are complicated functions of the observations on the quantities demanded and prices; hence their sampling distributions cannot be obtained with ease. Any point hypothesis, however, can be tested by means of the likelihood ratio test.

The likelihood ratio is defined by

$$\lambda = \frac{L(\hat{w})}{L(\hat{\Omega})} \tag{122}$$

where $L(\hat{w})$ is the maximum of L over the region of null hypotheses, and $L(\hat{\Omega})$ is the unrestricted maximum of L over the region of all alternative hypotheses. Under certain regularity conditions, the asymptotic distribution of $-2 \log \lambda$ is the χ^2 distribution with as many degrees of freedom as there are restrictions on the maximum of L contained in the set of null hypotheses. 25/

The use of the likelihood ratio may be clarified by an example: Suppose we wish to test the hypothesis that no more time than our unit period is required for consumers to adjust to their long-run equilibrium position, that is, the hypothesis that $\delta=1$. The maximum-likelihood estimate of the variance of the residual term w_t in equation (25) may be derived by stepwise maximization of the likelihood function. It is

$$\hat{\sigma}^2 = \frac{1}{T} \sum_{t=1}^{T} [x_t - \hat{a} - \hat{b}\, \bar{p}_t(\hat{\delta})]^2 \tag{123}$$

By the definition of R^2

$$\hat{\sigma}^2 = [1-R^2(\hat{\delta})]\ \hat{\sigma}_x^2 \tag{124}$$

where R^2 is the coefficient of correlation reached at $\hat{\delta}$, and $\hat{\sigma}_x^2$ is the variance of x. Similarly $\hat{\sigma}_1^2$ may be defined as the estimate of the residual variance obtained by maximizing L for $\delta=1$; it is

$$\hat{\sigma}_1^2 = [1-R^2(1)]\quad \sigma_x^2 \tag{125}$$

24/ See Cramer (9, p. 507).
25/ See Wilks (34, pp. 151-152).

From equation (114) we have

$$\log (w) = -\frac{T}{2} \log 2\pi \hat{\sigma}_1^2 - \frac{1}{2\hat{\sigma}^2} \sum_{t=1}^{T} [x_t - \hat{a} - \hat{b}\,\bar{p}_t(1)]^2 \tag{126}$$

Similarly

$$\log L(\hat{\Omega}) = -\frac{T}{2} \log 2\pi \hat{\sigma}^2 - \frac{1}{2\hat{\sigma}^2} \sum_{t=1}^{T} [x_t - \hat{\hat{a}} - \hat{\hat{b}}\,\bar{p}_t(\hat{\delta})]^2 \tag{127}$$

It follows from equations (126)-(127) that

$$\begin{aligned} -2 \log \lambda &= T \log \hat{\sigma}_1^2 - T \log \hat{\sigma}^2 \\ &= T \log [1-R^2(1)] - T \log [1-R^2(\hat{\delta})] \end{aligned} \tag{128}$$

Using a one-tail test we find the value of χ^2 for which the cumulative for one degree of freedom equals 0.95; call this value $\chi^2(1)_{0.05}$. Then if

$$T \log [1-R^2(\hat{\delta})] \leq T \log [1-R^2(1)] + \chi^2(1)_{0.05} \tag{129}$$

we reject the null hypothesis $\delta = 1$; otherwise we accept the hypothesis that consumers are always in long-run equilibrium.

Hypotheses concerning particular values which b^* may take also may be tested by means of the likelihood ratio. In addition, what might be called conditional confidence intervals can be obtained for a or b given the particular estimated value of δ. This conditional confidence interval is obtained in the usual way from the least squares regression of x_t and $\bar{p}_t(\delta)$. The problem is to determine the appropriate number of degrees of freedom to use in computing the interval. Since $\hat{\delta}$ is not given but has in fact been estimated, the usual confidence interval for a or b at a level of significance α is too small, or, what is the same, the interval obtained involves a level of significance less than α. Bearing these facts in mind, however, these conditional confidence intervals may prove a useful tool in the evaluation of results.

<u>Prevalence of serial correlation in the residuals</u>.--The method just described requires a large number of repeated steps; for this reason we call it the "iterative" method of estimation. In the case under discussion (in which the distributed lag is assumed to be due solely to rigidities of a technological or institutional nature), it is necessary to assume that the residuals, w_t, of equation (25) are normally and <u>independently</u> distributed in order to derive the iterative method. Thus the residuals of the equation containing a distributed lag are assumed not serially correlated. This will be true, however, only under special conditions: The w_t are uncorrelated serially only if the residuals, u_t, of the original demand equation follow a complicated autoregressive scheme:

$$u_t = - \sum_{\lambda=0}^{t-1} (1-\delta)^{t-\lambda} u_\lambda + \epsilon_t \tag{130}$$

where ϵ_t is normally and independently distributed. If the u_t do not follow (130), but rather are serially uncorrelated as is usually assumed, it can be shown that the w_t are positively serially correlated. 26/ Thus, use of the iterative procedure in the case under consideration may introduce serial correlation. This point is further discussed beginning on page 75.

Application to a general case.--Application of the iterative method to more general cases offers no further difficulty beyond that connected with serial correlation. Equation (29) gives the general form of a demand equation with distributed lags due to technological or institutional rigidities. All prices and income enter with the same distributed lag. Consequently, in the derivation of the iterative method, δ may be still considered as a scalar, but b* should now be considered as a vector $(a, b_1, b_2, \ldots, b_n, c)$. Because the iterative method rests on variation of the parameter δ, the fact that δ is a scalar is of utmost importance in determining the feasibility of using the iterative method. In the case in which distributed lags are due to technological or institutional rigidities, the iterative procedure is computationally feasible, but, as is shown below, when the distributed lags are of an expectational nature the iterative procedure is computationally feasible only in the simplest cases.

26/ The proof of this assertion is as follows:

$$w_t = \sum_{\lambda=0}^{t} \delta(1-\delta)^{t-\lambda} u_\lambda$$

hence

$$w_{t-1} = \sum_{\lambda=0}^{t-1} \delta(1-\delta)^{t-\lambda} u_\lambda$$

If the u_t are independently distributed, then for $i, j, = 0, \ldots, t, \ldots, \infty$

$$E\, u_{t-i}\, u_{t-j} \begin{cases} = 0, & \text{for } i \neq j \\ = \sigma_u^2, & \text{for } i = j \end{cases}$$

where $\sigma_u^2 = E\, u_t^2$; hence, provided t is large enough,

$$E\, w_t\, w_{t-1} = \frac{\delta^2(1-\delta)^2}{1-(1-\delta)^2} \sigma_u^2$$

since $0 \leq \delta < 1$. Clearly, then $E\, w_t\, w_{t-1} > 0$. By a similar proof we can show that $E\, w_t\, w_{t-i} > 0$ for all i. For further details, see page 78.

<u>Application under conditions of uncertainty when current price and income do not affect current demand</u>.--Consider the case in which the distributed lag is of an expectational nature and current price or current income does not affect the current quantity demanded. In this case the original demand equation may be written

$$x_t = a + b_1 p_t^* + c_1 y_t^* + u_t \tag{131}$$

where u_t is a random residual. 27/ The expectational equations necessary to complete the system are (32) and (33):

$$p_t^* - p_{t-1}^* = \beta [p_t - p_{t-1}^*] \tag{32}$$

$$y_t^* - y_{t-1}^* = \alpha [y_t - y_{t-1}^*] \tag{33}$$

Equations (131), (32), (33) lead to the following demand equation with distributed lags:

$$x_t = a + b_1 \sum_{\lambda=0}^{t} \beta(1-\beta)^{t-\lambda} p_\lambda$$

$$+ c_1 \sum_{\lambda=0}^{t} \alpha(1-\alpha)^{t-\lambda} y_\lambda + u_t \tag{132}$$

Two points about (132) are noteworthy: (1) the distributed lags for the two independent variables, price and income, are different, and (2) the residual term in (132) is the same as the residual term in equation (131). In order to use the iterative method, therefore, we must iterate on variations in a vector $\beta^* = (\beta, \alpha)$ rather than on variations in a scalar parameter.

On the other hand, no additional difficulties due to serial correlation are introduced by the use of the iterative method *per se*: If serial correlation is present in the residuals of equation (131), it will also be present in (132); but if it is absent in the residuals of (131), it will not be introduced by the iterative method into the residuals of the equation estimated, that is, (132). The addition of additional expectational variables to the system defined by equations (131), (32), and (33), poses no further conceptual problems. It is clear, however, that the computational difficulties are greatly increased.

<u>Application under conditions of uncertainty when current price or income affects current demand</u>.--An interesting case arises when current prices or current income is assumed to affect the current quantity demanded. The resulting demand equation with distributed lags is equation (37) plus a random residual term when only one price is assumed relevant:

27/ Equation (131) should be compared with equation (36) (see page 24). Note that b_0 and c_0 of (36) are assumed to be zero.

$$x_t = a + (b_0 + b_1\beta)\, p_t + b_1 \sum_{\lambda=0}^{t-1} \beta\,(1-\beta)^{t-\lambda}\, p_\lambda\ +$$

$$(c_0 + c_1\alpha)\, y_t + c_1 \sum_{\lambda=0}^{t-1} \alpha\,(1-\alpha)^{t-\lambda}\, y_\lambda\ + u_t \tag{37}$$

If $p^*_{t-1}(\beta) = \sum_{\lambda=0}^{t-1} \beta\,(1-\beta)^{t-\lambda}\, p_\lambda$ and $y^*_{t-1}(\alpha) = \sum_{\lambda=0}^{t-1} \alpha\,(1-\alpha)^{t-\lambda}$, then (37) may be written

$$x_t = a + (b_0 + b_1\beta)\, p_t + b_1\, p^*_{t-1}(\beta)$$

$$+ (c_0 + c_1\alpha)\, y_t + c_1\, y^*_{t-1}(\alpha) + u_t \tag{133}$$

If the iterative method is applied to equation (133), we obtain estimates of a, b_0, b_1, c_0, c_1, β, and α. Estimates of b_0 and c_0, however, must be obtained indirectly from the estimates of the coefficients of current price and income and estimates of b_1, c_1, β, and α. This case shows that it is possible to include non-expectational variables in the demand equation when the iterative method is used, and that doing so involves no great increase in computational difficulty. 28/ If the non-expectational variables are not current values of the expectational variables, all coefficients can be estimated directly by the iterative method.

Application to models that involve both rigidities and uncertainty.--Application of the iterative method to equations containing distributed lags of both an expectational and technological or institutional nature poses grave difficulties. These difficulties are due to the identification problem discussed beginning on page 44.

Consider the statistical model implied by equations (99), (32), and (33):

$$x^*_t = b\, p^*_t + u_t \tag{99}$$

where u_t is a random residual;

$$p^*_t - p^*_{t-1} = \beta\,[p_t - p^*_{t-1}] \tag{32}$$

$$x_t - x_{t-1} = \delta\,[x^*_t - x_{t-1}] \tag{33}$$

28/ The iterative method in this case is no longer maximum-likelihood, however. In order to obtain maximum-likelihood estimates, current price and current income should be included in $p^*_t(\beta)$ and $y^*_t(\alpha)$, respectively.

Solution of these equations by the approach outlined in the preceding section gives:

$$x_t = b\,\beta\,\delta\,[p_t + [(1-\beta) + (1-\delta)]\,p_{t-1} +$$

$$[(1-\beta)^2 + (1-\beta)(1-\delta) + (1-\delta)^2]\,p_{t-2} +$$

$$[(1-\beta)^3 + (1-\beta)^2(1-\delta) + (1-\beta)(1-\delta)^2 + (1-\delta)^3]\,p_{t-3} +$$

$$\ldots + \sum_{\lambda=0}^{t} \delta(1-\delta)^{t-\lambda}\,u_\lambda \tag{134}$$

Because of the introduction of technological or institutional rigidity, problems of serial correlation may arise if (134) is estimated by the iterative procedure as it may not be reasonable to assume that the residual term in (134), $\sum_{\lambda=0}^{t} \delta(1-\delta)^{t-\lambda}\,u_\lambda$, is serially uncorrelated. This difficulty is not, however, the most important.

By choosing pairs of β and δ in such a way that the series of weights for past prices entering (134) converges to zero as we go progressively farther back in time, we could, in principle, estimate equation (134) by the iterative method. 29/ As noted above, however, β and δ enter the weights of prices symmetrically. This means that the maximum correlation between the current quantity demanded and prices with respect to variations in β and δ is achieved at two points in the (β, δ) plane. Thus if one maximum is achieved at the point $\beta = \hat{\beta}$ and $\delta = \hat{\delta}$, the other maximum, which equals the first, is achieved at $\beta = \hat{\delta}$ and $\delta = \hat{\beta}$. Hence, having reached a maximum

---

29/ A sufficient, but not necessary, condition that the weights in (134) converge is that

$$|1-\beta + 1-\delta| < 1$$

since, after the weight of price lagged one year, the terms of the weight series are all less than the corresponding terms of the series

$$[1-\beta + 1-\delta]^i \qquad i=0, 1, \ldots, \infty$$

It is plausible, a priori, that this should be the case. Since $0 \leq \beta$, $\delta < 1$, the condition implies that

$$\beta + \delta > 1$$

Unless rigidity is extraordinary, this is reasonable.

with respect to the variation of two parameters, it is impossible to tell which parameter belongs to equation (32) and which to equation (33). It is clear, however, that the coefficient b can be estimated without difficulty by the iterative method.

When current price, as well as the long-run equilibrium quantity demanded, affects the rate of change of current demand, identification problems of this sort do not arise. But, as suggested by equation (102), the distribution of lag is extraordinarily complicated, so that it is difficult, if not impossible, to use the iterative method.

If it is possible to find a relevant non-expectational variable which enters the demand equation, no identification problem arises: Consider the system consisting of equations (32), (33) and the demand equation

$$x_t^* = b\, p_t^* + c\, z_t + u_t \tag{135}$$

where z_t is the non-expectational variable and u_t is a random residual. Equations (32), (33) and (135) lead to the following demand equation containing distributed lags:

$$\begin{aligned} x_t = {} & b\, \beta\, \delta \sum_{\mu=0}^{t} (1-\delta)^{t-\mu} \sum_{\lambda=0}^{\mu} (1-\beta)^{\mu-\lambda}\, p_\lambda \; + \\ & c\, \delta \sum_{\mu=0}^{t} (1-\delta)^{t-\mu}\, z_\mu + \sum_{\mu=0}^{t} \delta\, (1-\delta)^{t-\mu}\, u_\mu \end{aligned} \tag{136}$$

Provided the residual term in (136) can be reasonably assumed normally and independently distributed, the parameters of (136) may be estimated by the iterative method. The distribution of lag in z contains only the parameter δ; hence, no problem arises in distinguishing between the two parameters β and δ. The computations involved in using the iterative procedure, however, are laborious, since iteration must be performed on a pair of parameters.

A Method That Makes Use of Single Reduced Equations

We have seen that, except in the simplest cases, extraordinary computational difficulties are encountered if the iterative method is used. Consequently we turn now to a non-iterative method which results in fewer difficulties of a computational nature. This method rests on estimation of the reduced equations discussed extensively beginning on page 17.

Application when only rigidities are involved.--If a distributed lag is due only to rigidities of a technological or institutional nature, estimation is simple. Consider the model based on equations (23) and (28):

$$x_t^* = a + \sum_{i=1}^{n} b_i \, p_{it} + c \, y_t + u_t \tag{28}$$

where u_t is a random residual;

$$x_t - x_{t-1} = \delta \, [x_t^* - x_{t-1}] \tag{23}$$

Substitution of (28) into (23) leads to

$$x_t = a\,\delta + \sum_{i=1}^{n} b_i \, \delta \, p_{it} + c\,\delta\, y_t$$

$$+ (1-\delta)\, x_{t-1} + \delta\, u_t \tag{137}$$

The residual term of (137), $\delta\, u_t$, is much simpler than the residual term which would appear in the equation corresponding to (137), (25) page 19, if the former contains a distributed lag, namely

$$w_t = \sum_{\lambda=0}^{t} \delta\, (1-\delta)^{t-\lambda}\, u_\lambda$$

As remarked on page 53, if the u_t are serially uncorrelated, w_t tends to be positively serially correlated. It is clear, however, that if u_t is serially uncorrelated, so will $\delta\, u_t$ be; hence no <u>additional</u> difficulties with serial correlation are introduced if the parameters of the reduced equation are estimated rather than the parameters of the demand equation containing a distributed lag.

Equation (137) suggests that we run a regression of the following form:

$$x_t = \pi_0 + \sum_{i=1}^{n} \pi_1^{(i)}\, p_{it} + \pi_2\, y_t + \pi_3\, x_{t-1} + v_t \tag{138}$$

If the residuals v_t are uncorrelated with the p_{it}, y_t, and x_{t-1}, as well as serially uncorrelated, the parameters of (138) may be estimated by ordinary least squares. Let the estimated values of π_0, $\pi_1^{(i)}$ ($i=1, 2, \ldots, n$), π_2 and π_3 be denoted by $\hat{\pi}_0$, $\hat{\pi}_1^{(i)}$ ($i=1, 2, \ldots, n$), $\hat{\pi}_2$ and $\hat{\pi}_3$, respectively; then for estimates of a, b_i ($i=1, \ldots, n$) c, and δ we have:

$$\left.\begin{aligned}\hat{a} &= \frac{\hat{\pi}_0}{1-\hat{\pi}_3} \\ \hat{b}_i &= \frac{\hat{\pi}_{1(i)}}{1-\hat{\pi}_3} \qquad (i=1, \ldots, n) \\ \hat{c} &= \frac{\hat{\pi}_2}{1-\hat{\pi}_3} \\ \hat{\delta} &= 1-\hat{\pi}_3\end{aligned}\right\} \tag{139}$$

An ordinary confidence interval may be constructed for the estimate of δ, $\hat{\delta}$, but special methods must be used for the other estimates, $\hat{a}$, $\hat{b}_i$ $(i=1, \ldots, n)$, and $\hat{c}$. A method by which the variances and covariances of the estimates listed in (139) may be obtained is discussed by Klein (20, pp. 258-261).

Application to models based on uncertainty.--The situation becomes somewhat more complicated if distributed lags of an expectational nature are considered. Reduced demand equations containing distributed lags of an expectational nature may be obtained either by the single equation method or by the multiple equation method. As noted on pages 41-42, for statistical purposes the important distinction between distributed lags of an expectational nature and those due to technological or institutional rigidities is that the latter distributions of lag are the same for every variable in a given equation and different for the same variable in different demand equations, while the former are different for every variable in the same equation and the same for the same variable in different equations. The peculiar characteristics of distributed lags of an expectational nature lead to statistical complexities.

Consider the model defined by equations (131), (32), and (33):

$$x_t = a + b_1 p_t^* + c_1 y_t^* + u_t \tag{131}$$

$$p_t^* - p_{t-1}^* = \beta [p_t - p_{t-1}^*] \tag{32}$$

$$y_t^* - y_{t-1}^* = \alpha [y_t - y_{t-1}^*] \tag{33}$$

If these equations are reduced by the single equation method, a consumption function is not needed to complete the system. When we make use of equations (51)-(56), we have as the reduced form of the above system

$$x_t = a\alpha\beta + b_1\beta p_t + c_1\alpha y_t -$$

$$b_1(1-\alpha)\beta p_{t-1} - c_1\alpha(1-\beta)y_{t-1} +$$

$$[(1-\alpha)+(1-\beta)]x_{t-1} - (1-\alpha)(1-\beta)x_{t-2} +$$

$$u_t - [(1-\alpha)+(1-\beta)]u_{t-1} + (1-\alpha)(1-\beta)u_{t-2} \quad \underline{30}/ \qquad (140)$$

An important point to note about (140) is its residual term. If v_t is defined as

$$u_t - [(1-\alpha)+(1-\beta)]u_{t-1} + (1-\alpha)(1-\beta)u_{t-2} \qquad (141)$$

then the v_t are serially uncorrelated only if the u_t follow the following autoregressive scheme:

$$u_t = [(1-\alpha)+(1-\beta)]u_{t-1} - (1-\alpha)(1-\beta)u_{t-2} + \epsilon_t \qquad (142)$$

where ϵ_t is independently distributed. Second-order serial correlation as well as first-order must be assumed in the residuals u_t of the original demand equation in order to make the estimates of the parameters in the single-equation reduced equation unbiased and statistically efficient. (See page 80.) Comparison of the iterative method in this case shows that no additional difficulties with serial correlation are introduced by its use, whereas such difficulties may be introduced by the non-iterative method under discussion.

When the single-equation method of reduction is used on demand equations containing three or more expectational variables, estimation by the non-iterative method becomes extremely difficult. As indicated on page 31, if n expectational variables are included in the original demand equation, the single equation reduced equation contains n^2 more variables than the original equation. If non-expectational variables as well as expectational variables are included in the original demand equation, the situation becomes even more complex: If the original equation contains n expectational variables and m

---

30/ Comparison of equation (140) with equation (56) (see page 29) shows the effect of the addition of a constant term and a random residual term to the basic equation when the single equation method of reduction is used: The effect is to define new variables $y_t = x_t - a - u_t$, $y_{t-1} = x_{t-1} - a - u_{t-1}$, $y_{t-2} = x_{t-2} - a - u_{t-2}$. Substitution of these for x_t, x_{t-1}, and x_{t-2}, respectively, in equation (56) yields equation (140).

non-expectational variables, the single equation reduced form equation contains $n^2 + m$ more variables than the original equation. 31/

As a special case of this proposition we have the result that when the current quantity demanded depends on both the current and expected normal values of n expectational variables, the reduced demand equation contains $n(n+1)$ more variables than the original equation. In addition to the great difficulties with the number of independent variables, the single-equation method also introduces increasing difficulties with possible serial correlation as the number of variables in the original equation becomes larger. Only in the simplest of cases should the non-iterative method be based on the single equation reduced equation.

The case discussed beginning on page 59 is simple enough to warrant further discussion. Equation (140) suggests a regression of the form

$$x_t = \pi_0 + \pi_{11} p_t + \pi_{12} y_t + \pi_{21} p_{t-1} + \pi_{22} y_{t-1} + \pi_{31} x_{t-1} + \pi_{32} x_{t-2} + v_t \tag{143}$$

Estimates of the parameters a, b_1, c_1, β, and α may be derived from estimates of the πs in (143). Let the estimated value of a parameter be denoted by placing a hat over it, "$\wedge$", then one such set of estimates is

$$\hat{\beta} = 1 + \frac{\hat{\pi}_{22}}{\hat{\pi}_{12}}$$

$$\hat{\alpha} = 1 + \frac{\hat{\pi}_{21}}{\hat{\pi}_{11}}$$

$$\hat{a} = \frac{\hat{\pi}_0}{\left(1 + \frac{\hat{\pi}_{22}}{\hat{\pi}_{12}}\right)\left(1 + \frac{\hat{\pi}_{21}}{\hat{\pi}_{11}}\right)}$$

31/ This fact may be derived from equation (61) by setting $y_t = x_t - \sum_{i=1}^{m} a_i q_{it}$, where the q_{it} are non-expectational variables, and substituting y_t, y_{t-1}, etc., everywhere for x_t, x_{t-1}, etc.

$$\hat{b}_1 = \frac{\pi_{11}}{1 + \dfrac{\hat{\pi}_{22}}{\hat{\pi}_{12}}}$$

$$\hat{c}_1 = \frac{\hat{\pi}_{12}}{1 + \dfrac{\hat{\pi}_{21}}{\hat{\pi}_{11}}} \tag{144}$$

The estimates indicated by (144) represent only one possible set of such estimates, however. If β is estimated by $1 + \hat{\pi}_{22}/\hat{\pi}_{12}$ then two alternative estimates of α may be obtained in addition to the estimate in (144), one from $\hat{\pi}_{31}$ and one from $\hat{\pi}_{32}$. Consequently, two additional estimates of a, b_1, and c_1 may be obtained using these two estimates of α. Similarly, if α is estimated by $1 + \hat{\pi}_{21}/\hat{\pi}_{11}$, two additional estimates of β may be obtained. In addition, a quadratic in $\hat{\pi}_{31}$ and $\hat{\pi}_{32}$ may be solved to yield two pairs of estimates of β and α. Clearly, the set of estimates exhibited is only one among several possible sets; it is, however, the simplest of these sets and the least open to statistical difficulties. The variances and covariances of the estimates may be obtained by a method outlined by Klein (20, pp. 258-261).

A Method That Makes Use of Multiple Reduced Equations

Because of the difficulties involved in using the single reduced equation approach, it is usually advisable to use the multiple reduced equation method to make estimates. If, however, the distribution of lag for the same variable in different equations cannot be assumed the same, the multiple-equation method of reduction cannot be used. If the lags are not solely of an expectational nature, they tend to differ in different equations, so that the single equation method must be used. The value of the single equation method lies primarily in its wider range of applicability. If the multiple equation method can be used, use of reduced equations derived by it renders the non-iterative method of estimation much simpler computationally.

A simple example.--Consider the system defined by the following equations:

$$\left.\begin{aligned} x_{1t} &= b_{11}\, p_{1t}^{*} + b_{12}\, p_{2t}^{*} + u_{1t} \\ x_{2t} &= b_{21}\, p_{1t}^{*} + b_{22}\, p_{2t}^{*} + u_{2t} \\ p_{1t}^{*} &= \beta_1\, p_{1t}^{*} + (1-\beta_1)\, p_{1t-1}^{*} \\ p_{2t}^{*} &= \beta_2\, p_{2t} + (1-\beta_2)\, p_{2t-1}^{*} \end{aligned}\right\} \tag{145}$$

In matrix form, this system of equations may be written

$$\left.\begin{aligned} x_t &= B\, p_t^* + u_t \\ p_t^* &= \beta\, p_t + (I-\beta)\, p_{t-1}^* \end{aligned}\right\} \tag{146}$$

where x_t is a vector $(x_{1t}, x_{2t})'$, p_t^* is a vector $(p_{1t}^*, p_{2t}^*)'$, u_t is a vector (u_{1t}, u_{2t}), B is a matrix

$$\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$$

β is a matrix

$$\begin{pmatrix} \beta_1 & 0 \\ 0 & \beta_2 \end{pmatrix}$$

and p_t is a vector $(p_{1t}, p_{2t})'$. The reduced version of (146), obtained by the multiple equation method, is

$$\begin{aligned} x_t = B\ \beta\ p_t + B(I-\beta)B^{-1}\, x_{t-1} + \\ u_t - B(I-\beta)B^{-1}\, u_{t-1} \end{aligned} \tag{147}$$

Equations (147) suggest two regressions of the form

$$\begin{aligned} x_{1t} &= \pi_{11}^{(1)}\, p_{1t} + \pi_{12}^{(1)}\, p_{2t} + \pi_{11}^{(2)}\, x_{1t-1} + \pi_{12}^{(2)}\, x_{2t-1} + v_{1t} \\ x_{2t} &= \pi_{21}^{(1)}\, p_{1t} + \pi_{22}^{(1)}\, p_{2t} + \pi_{21}^{(2)}\, x_{1t-1} + \pi_{22}^{(2)}\, x_{2t-1} + v_{2t} \end{aligned} \tag{148}$$

or, in matrix form,

$$x_t = \Pi^{(1)}\, p_t + \Pi^{(2)}\, x_{t-1} + v_t \tag{149}$$

A complicated problem of serial correlation in the residuals of (148) or (149) may arise: Only if the residuals u_t follow a very complicated autoregressive scheme,

$$u_t = B(I-\beta)B^{-1}\, u_{t-1} + \epsilon_t \tag{150}$$

are the residuals v_t serially uncorrelated. Equation (150) implies only first order serial correlation, as contrasted with equation (142) above, but the current residuals from one demand equation must be correlated not only with their own lagged values but also with the lagged residuals of <u>other</u> demand equations. When the single equation method of reduction is used, successively higher order serial correlation must be assumed in the residuals of the original demand equation as more expectational variables are included. On the other hand, the multiple equation method never results in having to assume more than first order serial correlation. Furthermore, if the residuals of the original demand equations are in fact serially uncorrelated, the single equation method of reduction introduces serial correlation in the residuals of the reduced equation of successively higher orders; whereas, the multiple equation method introduces it only to the first order.

Estimates of the parameters of equations (145) or (146) may be derived from estimates of the parameters of equations (148) or (149). As before, let the sign "∧" over a parameter denote its estimated value, then, since

$$\left.\begin{aligned} \Pi^{(1)} &= B\beta \\ \Pi^{(2)} &= B(I-\beta)B^{-1} \end{aligned}\right\} \qquad (151)$$

we have

$$\left.\begin{aligned} \hat{B} &= (I-\hat{\Pi}^{(2)})^{-1}\,\hat{\Pi}^{(1)} \\ \hat{\beta} &= \hat{\Pi}^{(1)-1}\,(I-\hat{\Pi}^{(2)})\,\hat{\Pi}^{(1)} \end{aligned}\right\} \qquad (152)$$

Note, however, that unless the estimates of $\Pi^{(1)}$ and $\Pi^{(2)}$ are restricted in some way, the estimated matrix $\hat{\beta}$ is not necessarily diagonal, even though the matrix β is diagonal. Under the type of assumptions normally made, the off-diagonal elements of $\hat{\beta}$ might be taken to indicate the effects of current and past prices of other commodities upon the expected normal price of a given commodity. 32/

As in other examples, the standard errors of the estimates $\hat{B}$ and $\hat{\beta}$ might be derived from the variances and covariances of the estimates $\hat{\Pi}^{(1)}$ and $\hat{\Pi}^{(2)}$ by a method suggested by Klein (20, p. 258).

The general case when the original equations are separable.--The general case of a system such as that represented in (145) offers no difficulties beyond those previously encountered. Some interesting results are obtained, however, when the system of original equations is separable (see page 36). A separable system may be written

$$\begin{pmatrix} x_t \\ z_t \end{pmatrix} = \begin{pmatrix} B & D \\ 0 & A \end{pmatrix} \begin{pmatrix} p_t^* \\ q_t^* \end{pmatrix} + \begin{pmatrix} u_t \\ v_t \end{pmatrix} \qquad (153)$$

32/ In previous work on multiple equation reduction in connection with the study of agricultural supply functions, off-diagonal elements of $\hat{\beta}$ were typically close to zero as compared with the diagonal elements even though no restrictions were placed on the estimates of $\Pi^{(1)}$ or $\Pi^{(2)}$. See Nerlove, *op. cit.*, p. 285.

where x_t and z_t are vectors of quantities demanded or other non-expectational variables, B, D, A are matrices of coefficients, p_t^* and q_t^* are vectors of expectational variables, and u_t and v_t are vectors of random residual terms. The expectational equations needed to complete this system are

$$\begin{pmatrix} p_t^* \\ q_t^* \end{pmatrix} = \begin{pmatrix} \beta & 0 \\ 0 & \alpha \end{pmatrix} \begin{pmatrix} p_t \\ q_t \end{pmatrix} + \begin{bmatrix} (I-\beta) & 0 \\ 0 & (I-\alpha) \end{bmatrix} \begin{pmatrix} p_{t-1}^* \\ q_{t-1}^* \end{pmatrix} \tag{154}$$

As indicated by equation (84) (see page 37), the reduced form of (153) and (154) is

$$\begin{pmatrix} x_t \\ z_t \end{pmatrix} = \begin{pmatrix} B\beta & D\alpha \\ 0 & A\alpha \end{pmatrix} \begin{pmatrix} p_t \\ q_t \end{pmatrix} +$$

$$\begin{bmatrix} B(I-\beta)B^{-1} & -B(I-\beta)B^{-1}DA^{-1} + D(I-\alpha)A^{-1} \\ 0 & A(I-\alpha)A^{-1} \end{bmatrix} \begin{pmatrix} x_{t-1} \\ z_{t-1} \end{pmatrix}$$

$$+ \begin{pmatrix} u_t \\ v_t \end{pmatrix} -$$

$$\begin{bmatrix} B(I-\beta)B^{-1} & -B(I-\beta)B^{-1}DA^{-1} + D(I-\alpha)A^{-1} \\ 0 & A(I-\alpha)A^{-1} \end{bmatrix} \begin{pmatrix} u_{t-1} \\ v_{t-1} \end{pmatrix} \tag{155}$$

Equations (155) and (153) each represent sets of equations; the partitioning of the matrices represents a partitioning of the sets of equations into two sub-sets. The second of these sub-sets, obtained either from (153) or (154), may be used to eliminate some of the expectational variables, namely the q_t^*. Hence the first subset of equations of (155) may be written

$$x_t = (B\ \beta,\ DA^{-1}) \begin{pmatrix} p_t \\ z_t \end{pmatrix} +$$

$$(B(I-\beta)B^{-1},\ -B(I-\beta)B^{-1}\ DA^{-1}) \begin{pmatrix} x_{t-1} \\ z_{t-1} \end{pmatrix} +$$

$$u_t - B(I-\beta)B^{-1}\ u_{t-1} - B(I-\beta)B^{-1}\ DA^{-1}\ v_{t-1} + DA^{-1}\ v_t \tag{156}$$

Equation (156) may be obtained either by eliminating q_t^* from equation (153) and reducing the equation so obtained or by eliminating q_t from equations (155) (see page 39).

Equation (155) suggests a set of regressions which may be written in matrix form

$$\left.\begin{aligned} x_t &= \Pi_{11}\, p_t + \Pi_{12}\, q_t + \Pi_{13}\, x_{t-1} + \Pi_{14}\, z_{t-1} + w_{1t} \\ z &= \Pi_{22}\, q_t + \Pi_{24}\, z_{t-1} + w_{2t} \end{aligned}\right\} \tag{157}$$

whereas equation (156) suggests a set of regressions which may be written in the matrix form

$$x_t = \Pi_{11}^*\, p_t + \Pi_{12}^*\, z_t + \Pi_{13}^*\, x_{t-1} + \Pi_{14}^*\, z_{t-1} + w_{1t}^* \tag{158}$$

Although set (157) involves almost the same variables as set (158), the two sets of regressions cannot be interpreted in the same way. In order that the residual vector w_{1t} be serially uncorrelated, the vector of residuals u_t must follow the autoregressive scheme

$$\begin{aligned} u_t = {} & B(I-\beta)B^{-1}\, u_{t-1} - [B(I-\beta)B^{-1}\, DA^{-1} + \\ & D(I-\alpha)A^{-1}]\, v_{t-1} + \epsilon_t \end{aligned} \tag{159}$$

whereas, in order that the vector of residual terms w_{1t}^* be serially uncorrelated, the residuals u_t must be correlated not only with lagged values of themselves and v_t but also with current values of the residuals v_t, that is,

$$\begin{aligned} u_t = {} & -DA^{-1}\, v_t + B(I-\beta)B^{-1}\, u_{t-1} - \\ & B(I-\beta)B^{-1}\, DA^{-1}\, v_{t-1} + \epsilon_t \end{aligned} \tag{160}$$

Hence, if the regressions set forth in (158) are run and the u_t do not follow (160) to a first approximation, the true residuals of the fitted equations tend to be both serially correlated and correlated among themselves.

Estimates of <u>all</u> the parameters of the original equations, (153) and (154), may be derived only from estimates of the parameters in equations (157). Estimates of only <u>some</u> of the parameters of (153) and (154) may be derived from estimates of the parameters in equations (158). One of several possible sets of estimates of <u>all</u> the parameters of (153) and (154) is

$$\left.\begin{aligned}
\hat{\beta} &= \hat{\Pi}_{11}^{-1} (I - \hat{\Pi}_{13}) \hat{\Pi}_{11} \\
\hat{\alpha} &= \hat{\Pi}_{22}^{-1} (I - \hat{\Pi}_{24}) \hat{\Pi}_{22} \\
\hat{B} &= (I - \hat{\Pi}_{13})^{-1} \Pi_{11} \\
\hat{D} &= (I - \hat{\Pi}_{13})^{-1} [\hat{\Pi}_{12} + (I - \hat{\Pi}_{24})^{-1} \hat{\Pi}_{22} \hat{\Pi}_{14}] \\
\hat{A} &= (I - \Pi_{24})^{-1} \Pi_{22}
\end{aligned}\right\} \tag{161}$$

Only $\hat{B}$ and $\hat{\beta}$ may be obtained from the estimates of the parameters in (158). We may write the following matrix equations in these estimates:

$$\hat{\Pi}_{11}^{*} = \hat{B}\hat{\beta} \tag{162a}$$

$$\hat{\Pi}_{12}^{*} = \hat{D}\hat{A}^{-1} \tag{162b}$$

$$\hat{\Pi}_{13}^{*} = \hat{B}(I-\hat{\beta})\hat{B}^{-1} \tag{162c}$$

$$\hat{\Pi}_{14}^{*} = -\hat{B}(I-\hat{\beta})\hat{B}^{-1}\hat{D}\hat{A}^{-1} \tag{162d}$$

We have four equations to solve for the four unknown matrices $\hat{B}$, $\hat{\beta}$, $\hat{D}$, and $\hat{A}$, but these equations are either not independent or not consistent. In particular, (162d) may be derived from (162b-c), provided $\hat{\Pi}_{14}^{*} = \hat{\Pi}_{13}^{*} \hat{\Pi}_{12}^{*}$. Equations (162a) and (162c) may be solved in the usual way for $\hat{B}$ and $\hat{\beta}$. 33/ This particular difficulty does not occur if the separable system is the special case of non-expectational variables in the original equations. In general, though, unless we are interested only in $\hat{\beta}$ and $\hat{B}$, system (157) should be estimated in preference to system (158).

On page 36 we pointed out that a set of demand equations containing both expectational and non-expectational variables is actually a special case of a separable system. The same is true in the context of a statistical model. If the $q_t = z_t$ are non-expectational variables, the matrices A and α are identity matrices and the residuals v_t are all zero; hence the reduced equations are

---

33/ Under special circumstances, an assumption concerning either D or A can be made which surmounts this difficulty. See page 57 on the inclusion of non-expectational variables and page 109 on Friedman's permanent income hypothesis.

$$x_t = (B\ \beta\ ,\ D) \begin{pmatrix} p_t \\ z_t \end{pmatrix} +$$

$$(B(I-\beta)B^{-1},\ -\ B(I-\beta)B^{-1}D) \begin{pmatrix} x_{t-1} \\ z_{t-1} \end{pmatrix} +$$

$$u_t - B(I-\beta)B^{-1}\ u_{t-1} \tag{163}$$

Not only are the additional difficulties with intercorrelation removed in this special case of equation (156), 34/ but equations (162a-d) may be used to derive estimates of B, β , and D from $\hat{\Pi}_{11}^*$, $\hat{\Pi}_{12}^*$, $\hat{\Pi}_{13}^*$ and $\hat{\Pi}_{14}^*$.

Application to models that involve both rigidities and uncertainty.-- Separability of an equation system becomes especially useful when we deal with models that involve distributed lags due both to uncertainty about the future and to technological or institutional rigidities.

Consider the model consisting of equations (99), (32), and (23):

$$x_t^* = b\ p_t^* + u_t \tag{99}$$

$$p_t^* - p_{t-1}^* = \beta\ [p_t - p_{t-1}^*] \tag{32}$$

$$x_t - x_{t-1} = \delta\ [x_t^* - x_{t-1}] \tag{23}$$

Note that a random residual term has been added to (99). The reduced equation of the system (99)-(32)-(23) is

$$x_t = b\ \beta\ \delta\ p_t + [(1-\beta) + (1-\delta)]\ x_{t-1} -$$

$$(1-\beta)(1-\delta)\ x_{t-2} + \delta\ [u_t - (1-\beta)\ u_{t-1}] \tag{164}$$

For this equation, the residual term $\delta\ [u_t - (1-\beta)\ u_{t-1}]$ tends to be serially uncorrelated only if there is positive first order serial correlation among the residuals u_t. Second, as noted on page 46, the parameters β and δ enter (164) symmetrically (except on the residual term), and this produces a difficulty in estimation. Equation (164) suggests a regression of the following form

$$x_t = \pi_1\ p_t + \pi_2\ x_{t-1} + \pi_3\ x_{t-2} + v_t \tag{165}$$

Estimates of β and δ may be obtained from the estimates of the parameters of (165) in the following way: Let $\hat{\pi}_2$ and $\hat{\pi}_3$ denote the least squares estimates of π_2 and π_3. Comparison of (165) and (164) shows

34/ See equation (160) and the text beginning on page 64.

$$\hat{\beta} \ \underline{\text{or}} \ \hat{\delta} = 1 - \frac{\hat{\pi}_2 \pm \sqrt{\hat{\pi}_2 + 4 \hat{\pi}_3}}{2} \tag{166}$$

If $\hat{\beta}$ is given by the root taken with the plus sign, $\hat{\delta}$ is given by the root taken with the minus sign, and conversely. Hence, two pairs of estimates of β and δ may be obtained from a regression based on (165). A unique estimate of b may, however, be obtained from such a regression. The product of β and δ may be estimated by $1-\hat{\pi}_2-\hat{\pi}_3$, so that

$$\hat{b} = \frac{\hat{\pi}_1}{1-\hat{\pi}_2-\hat{\pi}_3} \tag{167}$$

This model may be generalized by making use of equations (110), (77) and (111):

$$\xi^*_t = \Gamma \ \eta^*_t + u_t \tag{110}$$

$$\eta^*_t = \epsilon \ \eta_t + (I-\epsilon) \ \eta^*_{t-1} \tag{77}$$

$$\xi_t = R \ \xi^*_t + (I-R) \ \xi_{t-1} \tag{111}$$

Note that a vector of random residuals has been added to (110). Following the method of reduction leading to equation (113) (see page 47), we have

$$\xi_t = R\,\Gamma\,\epsilon\,\eta_t + [(I-R) + R\Gamma(I-\epsilon)\Gamma^{-1}R^{-1}]\ \xi_{t-1} -$$

$$R\Gamma(I-\epsilon)\Gamma^{-1}R^{-1}\ (I-R)\ \ \xi_{t-2} + R[u_t - \Gamma(I-\epsilon)\Gamma^{-1}\ u_{t-1}] \tag{168}$$

Since R is a diagonal matrix, the residuals of (168) are serially uncorrelated only if the residuals u_t follow the autoregressive scheme

$$u_t = \Gamma\ (I-\epsilon)\Gamma^{-1}\ u_{t-1} + \ \epsilon_t \tag{169}$$

which is the same condition as for a model involving <u>only</u> distributed lags due to uncertainty.

Equation (168) suggests a set of regressions which may be written

$$\xi_t = \Pi_1\ \eta_t + \Pi_2\ \xi_{t-1} + \Pi_3\ \xi_{t-2} + v_t \tag{170}$$

where

$$v_t = R[u_t - \Gamma(I-\epsilon)\Gamma^{-1} u_{t-1}]$$

The following equations may be solved for the matrix of estimated values of each set of parameters $\hat{\Gamma}$, $\hat{R}$, and $\hat{\epsilon}$:

$$\hat{\Pi}_1 = \hat{R}\hat{\Gamma}\hat{\epsilon} \tag{171a}$$

$$\hat{\Pi}_2 = (I-\hat{R}) + \hat{R}\hat{\Gamma}(I-\hat{\epsilon})\hat{\Gamma}^{-1}\hat{R}^{-1} \tag{171b}$$

$$\hat{\Pi}_3 = \hat{R}\hat{\Gamma}(I-\hat{\epsilon})\hat{\Gamma}^{-1}\hat{R}^{-1}(I-\hat{R}) \tag{171c}$$

Substituting (171b) into (171c) we have

$$\hat{\Pi}_3 = (\hat{\Pi}_2 - (I-\hat{R}))(I-\hat{R}) \tag{172}$$

which is a quadratic in the matrix $(I-\hat{R})$. The individual equations from (172) are of the form

$$\hat{\pi}_{(3)ij} = \hat{\pi}_{(2)ij}(1-\hat{\rho}_j) - (1-\hat{\rho}_i)^2 \tag{173}$$

where ρ_i is the ith diagonal element of the diagonal matrix $\hat{R}$, and $\hat{\pi}_{(3)ij}$ and $\hat{\pi}_{(2)ij}$ are typical elements of the matrices $\hat{\Pi}_3$ and $\hat{\Pi}_2$, respectively. These equations cannot in general be solved if $\rho_i \neq \rho_j$ for all $i \neq j$; hence, (172) cannot be used to estimate the matrix R from estimates of the matrices Π_1, Π_2, and Π_3. 35/ If it were possible to estimate R, equations (171a-c) could be used to estimate Γ and ϵ. The reader will see, however, that each of the equations derived from (171a-c) other than (172) contains more than one of the matrices $\hat{R}$, $\hat{\Gamma}$, or $\hat{\epsilon}$, making the derivation of estimates difficult; hence, equation (170) does not represent a very fruitful set of regressions for estimation purposes.

If the system of original demand equations is a separable system, however, the situation is different. If a system of demand equations is separable it can be written

$$\begin{pmatrix} x_t^* \\ z_t^* \end{pmatrix} = \begin{pmatrix} B & D \\ 0 & A \end{pmatrix} \begin{pmatrix} p_t^* \\ q_t^* \end{pmatrix} + \begin{pmatrix} u_t \\ v_t \end{pmatrix} \tag{174}$$

where x_t^*, z_t^*, p_t^*, q_t^*, u_t, and v_t are vectors, and B, D, and A are matrices. The expectational equations may be written

$$\begin{pmatrix} p_t^* \\ q_t^* \end{pmatrix} = \begin{pmatrix} \beta & 0 \\ 0 & \alpha \end{pmatrix} \begin{pmatrix} p_t \\ q_t \end{pmatrix} + \begin{pmatrix} I-\beta & 0 \\ 0 & I-\alpha \end{pmatrix} \begin{pmatrix} p_{t-1}^* \\ q_{t-1}^* \end{pmatrix} \tag{175}$$

where β and α are diagonal matrices of coefficients of expectations.

35/ This was the difficulty alluded to in the text on page 46.

If the current values of the expectational variables are assumed to have no effect on the rate of change of the current quantities demanded, we may write the relationship between the current quantities and the long-run equilibrium quantities as

$$\begin{pmatrix} x_t \\ z_t \end{pmatrix} = \begin{pmatrix} R & 0 \\ 0 & S \end{pmatrix} \begin{pmatrix} x_t^* \\ z_t^* \end{pmatrix} + \begin{pmatrix} I-R & 0 \\ 0 & I-S \end{pmatrix} \begin{pmatrix} x_{t-1} \\ z_{t-1} \end{pmatrix} \tag{176}$$

where R and S are diagonal matrices whose elements lie between zero and one. Applying multiple equation reduction to the equation obtained by substituting (174) into (176) and equation (175), we have

$$\begin{pmatrix} x_t \\ z_t \end{pmatrix} = \begin{pmatrix} RB\beta & RD\alpha \\ 0 & SA\alpha \end{pmatrix} \begin{pmatrix} p_t \\ q_t \end{pmatrix} +$$

$$\left[\begin{pmatrix} I-R & 0 \\ 0 & I-S \end{pmatrix} + \begin{pmatrix} RB(I-\beta)B^{-1}R^{-1} & \begin{matrix} -RB(I-\beta)B^{-1}DA^{-1}S^{-1} \\ +RD(I-\alpha)A^{-1}S^{-1} \end{matrix} \\ 0 & SA(I-\alpha)A^{-1}S^{-1} \end{pmatrix}\right] \begin{pmatrix} x_{t-1} \\ z_{t-1} \end{pmatrix}$$

$$- \begin{pmatrix} RB(I-\beta)B^{-1}R^{-1} & \begin{matrix} -RB(I-\beta)B^{-1}DA^{-1}S^{-1} \\ +RD(I-\alpha)A^{-1}S^{-1} \end{matrix} \\ 0 & SA(I-\alpha)A^{-1}S^{-1} \end{pmatrix} \begin{pmatrix} I-R & 0 \\ 0 & I-S \end{pmatrix} \begin{pmatrix} x_{t-2} \\ z_{t-2} \end{pmatrix}$$

$$+ \begin{pmatrix} R & 0 \\ 0 & S \end{pmatrix} \left[\begin{pmatrix} u_t \\ v_t \end{pmatrix} - \begin{pmatrix} B(I-\beta)B^{-1} & \begin{matrix} -B(I-\beta)B^{-1}DA^{-1} \\ +D(I-\alpha)A^{-1} \end{matrix} \\ 0 & A(I-\alpha)A^{-1} \end{pmatrix} \begin{pmatrix} u_{t-1} \\ v_{t-1} \end{pmatrix}\right] \tag{177}$$

Equation (177) suggests two sets of regressions of the following form

$$x_t = \Pi_{11}^{(1)} p_t + \Pi_{12}^{(1)} q_t + \Pi_{21}^{(1)} x_{t-1} +$$

$$\Pi_{22}^{(1)} z_{t-1} + \Pi_{31}^{(1)} x_{t-2} + \Pi_{32}^{(1)} z_{t-2} + w_t^{(1)} \tag{178a}$$

and

$$z_t = \Pi_{12}^{(2)} q_t + \Pi_{22}^{(2)} z_{t-1} + \Pi_{32}^{(2)} z_{t-2} + w_t^{(2)} \tag{178b}$$

Provided the residuals u_t and v_t follow the autoregressive schemes

$$u_t = B(I-\beta)B^{-1} u_{t-1} - [B(I-\beta)B^{-1}DA^{-1}$$

$$+ D(I-\alpha)A^{-1}] v_{t-1} + \epsilon_{1t} \tag{179a}$$

and

$$v_t = A(I-\alpha)A^{-1} v_{t-1} + \epsilon_{2t} \tag{179b}$$

the matrices of coefficients in (178a-b) may be estimated by ordinary least squares. Estimates of the matrices of parameters in the equations (174), (175), and (176) may be derived from these estimates solely by the solution of linear equations.

From (177) and (178a-b) we have the following equations:

$$\hat{\Pi}_{11}^{(1)} = \hat{R}\hat{B}\hat{\beta} \tag{180a}$$

$$\hat{\Pi}_{12}^{(1)} = \hat{R}\hat{D}\hat{\alpha} \tag{180b}$$

$$\hat{\Pi}_{21}^{(1)} = I-\hat{R} + \hat{R}\hat{B}(I-\hat{\beta})\hat{B}^{-1}\hat{R}^{-1} \tag{180c}$$

$$\hat{\Pi}_{22}^{(1)} = -\hat{R}\hat{B}(I-\hat{\beta})\hat{B}^{-1}\hat{D}\hat{A}^{-1}S^{-1} + \hat{R}\hat{D}(I-\hat{\alpha})\hat{A}^{-1}\hat{S}^{-1} \tag{180d}$$

$$\hat{\Pi}_{31}^{(1)} = -\hat{R}\hat{B}(I-\hat{\beta})\hat{B}^{-1}\hat{R}^{-1}(I-\hat{R}) \tag{180e}$$

$$\hat{\Pi}_{32}^{(1)} = \hat{R}\hat{B}(I-\hat{\beta})\hat{B}^{-1}\hat{D}\hat{A}^{-1}\hat{S}^{-1}(I-\hat{S}) - \hat{R}\hat{D}(I-\hat{\alpha})\hat{A}^{-1}\hat{S}^{-1}(I-\hat{S}) \tag{180f}$$

$$\hat{\Pi}_{12}^{(2)} = \hat{S}\hat{A}\hat{\alpha} \tag{180g}$$

$$\hat{\Pi}_{22}^{(2)} = (I-\hat{S}) + \hat{S}\hat{A}(I-\hat{\alpha})\hat{A}^{-1}\hat{S}^{-1} \tag{180h}$$

$$\hat{\Pi}_{32}^{(2)} = -\hat{S}\hat{A}(I-\hat{\alpha})\hat{A}^{-1}\hat{S}^{-1}(I-\hat{S}) \tag{180i}$$

We have <u>nine</u> equations to solve for the <u>seven</u> unknown matrices, $\hat{B}$, $\hat{D}$, $\hat{A}$, $\hat{\beta}$, $\hat{\alpha}$, $\hat{R}$, and $\hat{S}$; hence, more than one solution may be possible.

One possible solution to (180a-i) may be obtained as follows: Equations (180d) and (180f) lead to

$$\hat{\Pi}_{32}^{(1)} = [\hat{R}\hat{B}(I-\hat{\beta})\hat{B}^{-1}\hat{D}\hat{A}^{-1}\hat{S}^{-1} - \hat{R}\hat{D}(I-\hat{\alpha})\hat{A}^{-1}\hat{S}^{-1}](I-\hat{S}) = -\hat{\Pi}_{22}^{(1)}(I-\hat{S})$$

So that

$$\hat{S} = I + \hat{\Pi}_{22}^{(1)-1}\hat{\Pi}_{32} \tag{181a}$$

Hence, S can be estimated and we can use $\hat{S}$ for S in subsequent calculations.

From equations (180g) and (180i) we have

$$\hat{\Pi}_{32}^{(2)} = -I + \hat{\Pi}_{12}^{(2)} \hat{A}^{-1} \hat{S}^{-1}$$

so that

$$\hat{A} = \hat{S}^{-1} (I + \hat{\Pi}_{32}^{(2)})^{-1} \hat{\Pi}_{12}^{(2)} \tag{181b}$$

S may be substituted in (181b) from (181a) to give an estimate of A; hence $\hat{A}$ can be used in place of A in subsequent calculations. Since $\hat{S}$ and $\hat{A}$ are known we have from (180g) that

$$\hat{\alpha} = \hat{A}^{-1}\hat{S}^{-1} \hat{\Pi}_{12}^{(2)} \tag{181c}$$

Hence, $\hat{\alpha}$ can be used for α in subsequent calculations. From equations (180d), (180b), and (180a), we have

$$\hat{S}\hat{A} \hat{\Pi}_{22}^{(1)} = \hat{R}\hat{D}(I-\hat{\alpha}) - \hat{R}\hat{B}(I-\hat{\beta})\hat{B}^{-1}\hat{D}$$

$$= \hat{R}\hat{D} - \hat{\Pi}_{12}^{(1)} - \hat{R}\hat{D} + \hat{\Pi}_{11}^{(1)} \hat{D}$$

$$= \hat{\Pi}_{11}^{(1)} \hat{D} - \hat{\Pi}_{12}^{(1)}$$

hence

$$\hat{D} = \hat{\Pi}_{11}^{(1)-1} (\hat{\Pi}_{12}^{(1)} + \hat{S}\hat{A} \hat{\Pi}_{22}^{(1)}) \tag{181d}$$

Since $\hat{S}$ and $\hat{A}$ are known from (181a) and (181b), D can be estimated from (181d). From equation (180b) we have

$$\hat{R} = \hat{\Pi}_{12}^{(1)} \hat{\alpha}^{-1}\hat{D}^{-1} \tag{181e}$$

Since $\hat{D}$ and $\hat{\alpha}$ are known from (181d) and (181c), (181e) may be used to find $\hat{R}$. From (180c) and (180a) we have

$$\hat{\Pi}_{21}^{(1)} = I-\hat{R} + I - \hat{\Pi}_{11}^{(1)} \hat{B}^{-1}\hat{R}^{-1}$$

hence

$$\hat{B} = \hat{R}^{-1} (\hat{\Pi}_{21}^{(1)} + \hat{R} - 2I)^{-1} \hat{\Pi}_{11}^{(1)} \tag{181f}$$

Finally, from (180a) we have

$$\hat{\beta} = \hat{\Pi}_{11}^{(1)} \hat{R}^{-1}\hat{B}^{-1} \tag{181g}$$

It follows from (181a-g) that estimates of the parameters of (174)-(176) may be derived from estimates of the parameters of (178a-b) by the solution of <u>linear</u> equations.

In the preceding examples, current values of expectational variables are assumed to have no effect on the rates of change of current quantities demanded. We now assume instead that the rate of change of the current quantity demanded depends not only on the difference between the long-run equilibrium and current quantities demanded but also on the difference between the current and expected normal prices of the commodity in question. Thus we use relations of the form of (100) rather than of the form of (23).

First, consider the model consisting of equations (99), (32), and (100)

$$x_t^* = b\ p_t^* + u_t \tag{99}$$

$$p_t^* - p_{t-1}^* = \beta\ [p_t - p_{t-1}^*] \tag{32}$$

$$x_t - x_{t-1} = \delta_0\ [p_t - p_t^*] + \delta_1\ [x_t^* - x_{t-1}] \tag{100}$$

The reduced version of the system (99), (32), (100) may be obtained by the same method that was used to obtain (104); it is

$$\begin{aligned} x_t = {} & [b\ \beta\ \delta_1 + (1-\beta)\ \delta_0]\ p_t - (1-\beta)\ \delta_0\ p_{t-1} + \\ & [(1-\beta) + (1-\delta_1)]\ x_{t-1} - \\ & (1-\beta)\ (1-\delta_1)\ x_{t-2} + \delta_1\ [u_t - (1-\beta)\ u_{t-1}] \end{aligned} \tag{182}$$

Here the residual term, $\delta_1\ [u_t - (1-\beta)\ u_{t-1}]$, has the same properties as the residual term of equation (164) (see page 68). β, δ_0, and δ_1 enter (182) asymmetrically, however.

Equation (182) suggests a regression of the form

$$x_t = \pi_{11}\ p_t + \pi_{12}\ p_{t-1} + \pi_{21}\ x_{t-1} +$$

$$\pi_{22}\ x_{t-2} + v_t \tag{183}$$

If the appropriate serial correlation exists, at least approximately, among the residuals u_t, equation (183) may be estimated by least squares.

Estimates of b, β, δ_0, and δ_1 may be obtained from the estimates of π_{11}, π_{12}, π_{21}, and π_{22}. Unfortunately, this requires the solution of a quadratic in the estimates $\hat{\pi}_{21}$ and $\hat{\pi}_{22}$ of π_{21} and π_{22}. As indicated on page 69, certain difficulties arise from this, particularly since β and δ_1 enter the coefficients π_{21} and π_{22} symmetrically. We therefore cannot tell whether a particular root is an estimate of β or of δ_1.

As in the preceding example, this case may be generalized. Substantially the same results emerge concerning the possibility of estimating the coefficients of the structural equations and the effects of separability.

EFFECTS OF SERIAL CORRELATION

As noted previously, the statistical estimation of the structural parameters of equations containing distributed lags by least squares frequently requires rather unusual assumptions concerning the nature of the residual terms in the original demand equations. When the distributed lags are due solely to technological or institutional rigidities, these complications do not arise in estimating parameters in the reduced equations but they do arise when we use the iterative procedure.

In our discussion of estimation techniques we have come across several types of problems that involve possible serial correlation. Outline 1 summarizes the possibilities in a number of cases: Column (1) shows the supposed basic cause of the distributed lag under consideration. Column (2) indicates the method of estimation to be employed; the iterative method is used to estimate the parameters of the demand equation and the distribution of lag directly from an equation involving a distributed lag, and the non-iterative method is used to estimate the parameters of the original demand equation and the distribution of lag indirectly by estimating the parameters of a reduced demand equation or set of such equations which involve only discrete lags. Column (3) shows the type of reduction used if appropriate. Single and multiple equation methods are indicated, and the multiple equation method for a separable system is designated by the term "System separated." In the last case we suppose that the equations which may be separated from the main body of the system (such as the consumption function) are solved for the expectational variables they involve; these solutions are then substituted in the main body of the system of equations before reduction. Column (4) shows the actual residual term in the equation or equations used for estimation in terms of the residual term(s) of the original demand equation(s). If the distributed lags are due to uncertainty and the non-iterative method is used, the methods of reduction are supposed to apply to a model such as

$$\left.\begin{aligned} x_{1t} &= b_{11}\, p_{1t}^{*} + b_{12}\, p_{2t}^{*} + u_{1t} \\ x_{2t} &= b_{21}\, p_{1t}^{*} + b_{22}\, p_{2t}^{*} + u_{2t} \end{aligned}\right\} \qquad (184a)$$

or

$$\left.\begin{aligned} x_{1t} &= b_{11}\, p_{1t}^{*} + c_1\, y_t^{*} + u_{1t} \\ C_t &= \quad 0 \quad + c_{n+1}\, y_t^{*} + u_{n+1\ t} \end{aligned}\right\} \qquad (184b)$$

In the case of single equation reduction, β and α refer to different coefficients of expectations, and in the last two cases the a's are complicated functions of the b's or c's in (184a-b) and the coefficients of expectations. Column (5) shows the type of serial correlation present in the residuals of the estimation equation if the residuals of the original demand equations, u_t

Outline 1.--Alternative methods of estimation: Effects on serial correlation in the residuals of the estimating equation

Case	Cause of distributed lag (1)	Method of estimation (2)	Type of reduction (3)	Residual term in estimating equation (4)	Nature of serial correlation introduced if u_t is independently distributed (5)
1	Technological or institutional rigidities	Iterative	None	$\sum_{\lambda=0}^{t} \delta(1-\delta)\, u_{\lambda}$ 1/	Moving summation (positive serial correlation)
2	Do.	Noniterative	Single equation	$\delta\, u_t$ 2/	None
3	Uncertainty	Iterative	None	u_t 3/	Do.
4	do.	Noniterative	Single equation 4/	$u_t - [(1-\alpha) + (1-\beta)]\, u_{t-1} + (1-\alpha)(1-\beta)\, u_{t-2}$ 5/	Moving average
5	do.	do.	Multiple equation 4/	$u_{1t} - (a_1^{(1)}u_{1t-1} + a_2^{(1)}u_{2t-1})$ 6/ $u_{2t} - (a_1^{(2)}u_{1t-1} + a_2^{(2)}u_{2t-1})$	Moving average Intercorrelation of residuals in different equations as well as serial correlation
6	do.	do.	System separated 4/ (multiple equation)	$u_{1t} + a_1^{(1)}u_{2t} + a_2^{(1)}u_{1t-1} + a_3^{(1)}u_{2t-1}$ 7/	Moving average

1/ See equation (130), page 53. 2/ See equation (137), page 58. 3/ See equation (132), page 54. 4/ Two expectational variables assumed to be involved in original equation or equations. 5/ See equation (140), page 60. 6/ See equation (147), page 63. 7/ See equation (156), page 65.

or u_{1t} and u_{2t}, are in fact independently distributed. 36/ Cases which result from a combination of technological rigidity and uncertainty are not discussed, inasmuch as the residual terms possess properties identical with those of a noncombination model.

A brief summary of the effects of nonindependence in the residual terms of an estimating equation when least squares estimating techniques are used is given here. If no lagged values of the dependent variable enter as independent variables, the estimated coefficients are statistically unbiased and consistent, but they are statistically inefficient and the ordinary t- and F-tests do not apply. The estimates of the residual variance are biased. When lagged values of the dependent variable enter as independent variables, the coefficients are biased for small samples. 37/ Mann and Wold (23, pp. 175-192) show, under fairly general assumptions about the nature of the residual serial correlation, that the least squares estimates of the coefficients are asymptotically unbiased and consistent. The term asymptotic indicates that these properties are realized only for large samples.

Alternative Cases

With the above comments in mind, we may analyze the effects of the alternative cases presented in outline 1.

Case 1.--Here the distributed lag or lags are due to technological or institutional rigidities and the iterative method of estimation is used. Unless the residuals of the original demand equation follow the complicated autoregressive scheme indicated in equation (130) (see page 53), the residuals

$$w_t = \sum_{\lambda=0}^{t} \delta (1-\delta)^{t-\lambda} u_\lambda \tag{185}$$

are not independently distributed. 38/ Provided we take our origin in the sufficiently distant past, it may be shown that the correlation of w_t with w_{t-n}, ρ_n, is

$$\rho_n = (1-\delta)^n \tag{186}$$

36/ The terms "moving summation" and "moving average" refer to a specified type of nonindependence in the residual terms. Their meaning should be clear in this context.

37/ See Hurwicz (19). Hurwicz shows, for a number of simple cases, that the bias of the coefficients of the lagged dependent variables tends toward zero. Some of his results are also applicable to maximum likelihood estimates of the coefficients in systems of simultaneous equations which contain lagged endogenous variables as well as predetermined variables.

38/ We partly proved this in footnote 26. The results which follow are based on the general formula for the correlogram of a stationary stochastic process of moving summation.

if the u_t are independently distributed. The variance of w_t can be shown to be

$$\sigma_w^2 = \frac{\delta^2}{1-(1-\delta)^2}\,\sigma_u^2 \tag{187}$$

where σ_u^2 is the variance of the residual term u_t of the original demand equation. Consequently the variance-covariance matrix, Ω, of the residuals in the estimating equation is

$$\Omega = \frac{\delta^2\,\sigma_u^2}{1-(1-\delta)^2}\begin{pmatrix} 1 & (1-\delta) & (1-\delta)^2 & \cdots \\ (1-\delta) & 1 & (1-\delta) & \cdots \\ (1-\delta)^2 & (1-\delta) & 1 & \cdots \\ \vdots & \vdots & \vdots & \end{pmatrix} \tag{188}$$

Equation (188) should be compared with the usual assumption that $\Omega = \sigma_w^2 I$, where I is the identity matrix. Because of this nonindependence, the estimates of the coefficients in the estimating equation are not statistically efficient <u>and the estimate of the residual variance is biased</u>. <u>39</u>/ Recalling that the iterative procedure involves maximizing the calculated multiple correlation between the observed quantity demanded and weighted moving averages of current and past prices with respect to the weights in those averages, we see that, since the estimate of the residual variance is biased, so may be our estimates of all the coefficients, δ included. In sum, independence of the u_t leads to biased estimates in the first case of outline 1, even though no lagged values of the dependent variable are included as independent variables in the estimating equation.

A relatively simple solution to this problem exists. Aitken (<u>1</u>) has shown that if the variance-covariance matrix of the residuals in the estimating equations is known, it is possible to specify a transformation of the variables which removes all nonindependence from the resulting residual terms. Furthermore, it is not actually necessary to specify the transformation, but only to compute the moments of the normal equations in a way which depends on the inverse of the original variance-covariance matrix, Ω. Since the matrix Ω depends only on the parameter δ (in the case under consideration), it is feasible to use Aitken's method at every iteration of the procedure: Just as the weighted moving averages of prices is different for each iteration, so is the exact way in which the moments are computed. All this, of course, rests on the assumption that the residuals u_t are independently distributed; if this assumption cannot be made, little can be done.

<u>39</u>/ The fact that the ordinary t- and F-tests do not apply need not concern us in this case, since they never apply when the iterative procedure is used regardless of whether the residuals are independently distributed or not.

Case 2.--Here the distributed lag is again due to rigidities of a technological or institutional nature but, in contrast to case 1, the noniterative method is used. In this case, the assumption that the residuals of the original demand equation are independently distributed does not lead to any complications: the residual δu_t also are independently distributed.

If, however, the u_t are not independently distributed, their lack of independence shows up in the residuals of the reduced equation. Since the lagged quantity consumed is treated as an independent variable, the estimates of the coefficients in the reduced equation are biased for small samples. In this instance we have no a priori basis on which to specify the form or content of the residual variance-covariance matrix; consequently Aitken's generalized least squares cannot be easily applied. Cochrane and Orcutt (7) suggest an iterative method of autoregressive transformation which might be applied in this case.

Case 3.--The next three cases presented in outline 1 deal with distributed lags due to uncertainty. When the iterative procedure is used, as in case 3, no serial correlation is introduced in the residuals of the estimating equation if there is none present in the residuals. This is in contrast to the situation when the iterative method is used but the distributed lags are due to technological or institutional rigidities.

If the residuals of the original demand are not distributed independently, the estimates of the parameters in the demand equation and the coefficient(s) of expectation may be biased, since the estimate of the residual variance is. Since we have no a priori basis on which to specify the residual variance-covariance matrix, Aitken's method cannot be employed. If we have reason to suspect that the u_t are not independently distributed, or if we find significant serial correlation in the calculated residuals of the final result, 40/ the best that can be done, at present, is to use the technique suggested by Cochrane and Orcutt (8).

Case 4.--Here the parameters of the original demand equation are to be estimated from a reduced equation obtained from the original by the single equation method. If the residuals of the reduced equation are a "moving average" of the residuals of the original demand equation, the residuals

$$w_t = u_t - [(1-\alpha) + (1-\beta)]\, u_{t-1} + (1-\alpha)(1-\beta)\, u_{t-2} \tag{189}$$

tend to be serially correlated if the u_t are not. The correlogram of the residuals w_t is as follows

40/ Using, for example, the Durbin-Watson statistic. See Durbin and Watson (10).

$$\rho_n = \begin{cases} 1 & n = 0 \\ -\dfrac{[(1-\alpha) + (1-\beta)]\,[1 + (1-\alpha)(1-\beta)]}{1 + [(1-\alpha) + (1-\beta)]^2 + [(1-\alpha)(1-\beta)]^2} & n = 1 \\ \dfrac{(1-\alpha)(1-\beta)}{1 + [(1-\alpha) + (1-\beta)]^2 + [(1-\alpha)(1-\beta)]^2} & n = 2 \\ 0 & n \geq 3 \end{cases} \tag{190}$$

Since $0 \leq \alpha$, $\beta < 1$, the first-order serial correlation is negative and the second-order serial correlation is positive. Because lagged values of the dependent variable are included as independent variables, the least squares estimates of the parameters in the reduced equation tend to be biased for small samples, if the residuals of the original equation are independently distributed. As indicated in our previous discussion of estimation based on single equation reductions, the more expectational variables included in the original equation, the greater the number of lagged residual terms, u_t, included in the residual term of the reduced form, w_t.

If all the coefficients of expectations (α and β in this case) were known a priori, Aitken's method of generalized least squares could be applied in case we had reason to believe the residuals of the original demand equation were independently distributed. Since, however, they are not known, other methods must be used. One possible procedure would be to apply the transformation procedure implied by Aitken's procedure, but with unspecified coefficients of expectations. We could then minimize the sum of squared residuals with respect to these parameters and with respect to the parameters in the original equation and the coefficients of expectations. 41/ Although this method leads to statistically efficient estimates, it is not certain that they are unbiased. Furthermore, the "normal" equations resulting from such minimization are not linear in the parameters and, hence, are difficult to solve. Alternatively we could use the method suggested by Cochrane and Orcutt (8).

In the case presented in the table, the residuals of the reduced equation are a linear function of the residuals of the original equation and two of their lagged values; thus we have a second-order difference equation in these original residuals. Solution of this difference equation yields an autoregressive scheme. The first few autoregressive coefficients in the scheme might then, as Cochrane and Orcutt suggest, be estimated from the calculated residuals. The results could then be used to transform the variables entering the reduced equation. Such a procedure would, however, be difficult in practice and likely to yield only crude approximations to the correct transformations.

41/ This method follows from one suggested by Klein (20, pp. 186-189).

Case 5.--Here the multiple equation method of reduction is used. Column (4) shows that the residuals of one of the reduced equations are linear functions of the residuals of one of the original demand equations and lagged values of the residuals of two of the original equations. If all the residuals of the original equations are independently distributed, the fact that the residual term of the reduced equation contains more than one lagged residual of the original equations causes no more difficulty than the case in which only one expectational variable is included. In the latter case single and multiple reduced equations are identical: the reduced residual is a linear function of the residual of the original equation and its lagged value. Because, in the case presented in outline 1, the residual u_{2t-1} is independently distributed, it may be taken as a random addition to the residual, w_{1t}, of the first reduced equation. As noted in the outline, however, the effect of the two random terms in w_{1t} and w_{2t} is to make them nonindependent. This, of itself, is not sufficient to cause bias in the estimates of the coefficients. The serial correlation of w_{1t} and w_{2t} does, however, cause bias in the estimates of the parameters of the reduced equation.

The correlogram of w_{it} is

$$\left.\begin{aligned} \rho_0^{(i)} &= 1 \\ \rho_1^{(i)} &= \frac{-a_1^{(i)}}{1+[a_1^{(i)}]^2} \\ \rho_n^{(i)} &= 0 \;, \qquad n > 1 \end{aligned}\right\} (i=1,\ 2) \qquad (191)$$

Since, as reference to equation (147) indicates, $a_1^{(i)}$ is a complicated function of the parameters in both original demand equations and both coefficients of expectations, it is impossible to apply Aitken's method. The best procedure probably is to use either of the two methods suggested in the previous case.

Case 6.--This deals with a separable system which is separated prior to reduction. When a system is separated, the dependent variables of some of the equations enter as independent variables in the equations used to estimate the parameters of part of the original system of demand equations. The residual terms of the eliminated equations also enter the residual terms of the reduced equations used for estimation; consequently, severe bias of the estimates of the coefficients may result. In view of this additional difficulty, it is usually wise not to separate a system of separable demand equations, but rather to estimate them as a group or any one of them from a single reduced equation. 42/

42/ This statement differs from a suggestion made in connection with a test of Friedman's permanent income hypothesis. See page 112.

Alternative Assumptions

As suggested in our discussion of estimation procedures, all the difficulties so far presented in this part may, in theory, be obviated by assuming that the residuals of the original demand equations follow more-or-less complicated autoregressive schemes. In view of the nature of most economic time series, to assume these autoregressive schemes is neither better nor worse than to assume the residuals in the original equations are independently distributed. On the other hand, false specification of any kind may lead to serious errors.

One procedure is as follows: Test the residuals of the estimating equations for evidence of serial correlation; if significant serial correlation is found, then worry about it.

INTERDEPENDENCY OF VARIABLES WITHIN A SYSTEM

Although the problems of serial correlation are serious, they are more easily managed by assumption than the group of problems to which we now turn. In our exposition of the multiple equation method of reduction, we made use of the fact that an individual demand equation generally is part of a system of equations. We restricted the other equations in this system, however, to be demand equations. In doing so we neglected a problem of crucial importance: both current prices and current quantities frequently are determined jointly by a system of demand and supply equations. If no distributed lags are involved in either set, the problem is that usually discussed in the theory of estimation of simultaneous equations. 43/ The introduction of distributed lags raises problems which are hitherto unsolved. As we show, these problems are closely connected with the problem of serial correlation.

Conditions Which Must Be Satisfied If the Method of Least Squares Is Used

In a recent article, Phillips (28) summarized the necessary conditions for the use of least squares on systems of simultaneous equations that involve distributed lags. Although Phillips deals only with a relatively simple case, his main results are useful. 44/

Consider the following demand equation involving a distributed lag:

$$x_t = b_0 p_t + b_1 p_{t-1} + b_2 p_{t-2} + \cdots + b_r p_{t-r} + u_t \tag{192}$$

43/ See Klein (20) or Koopmans and Hood (21).
44/ The model used here for expository purposes is slightly different from that given by Phillips.

where the constant term is neglected for simplicity. Note that this equation is almost identical to equation (4) (see page 8). We suppose that we also have a supply equation:

$$x_t = a_0 p_t + a_1 p_{t-1} + a_2 p_{t-2} + \cdots + a_s p_{t-s} + v_t \tag{193}$$

The quantity consumed and produced, x_t, and the current price, p_t, are assumed to be jointly determined by equations (192) and (193). Under what conditions may the coefficients of (192) be estimated without statistical bias by least squares?

Phillips gives four conditions, all of which must be satisfied if the least squares estimates are to be unbiased. If the number of past prices which we must include in (192), r, is known, Phillips' first condition is not applicable. In practice, however, r, is not known. Instead we might choose to introduce R past prices, so that we actually try to estimate

$$x_t = b_0 p_t + b_1 p_{t-1} + \cdots + b_R p_{t-R} + u_t' \tag{194}$$

Phillips' <u>first condition</u> is that

$$R \geq r \tag{195}$$

For if $R < r$, the lagged prices $p_{t-(R+1)}, p_{t-(R+2)}, \cdots p_{t-r}$, are omitted from (194) and the residual term is

$$u' = b_{R+1} p_{t-(R+1)} + b_{R+2} p_{t-(R+2)} + \cdots + b_r p_{t-r} + u_t \tag{196}$$

Equation (193) is an s-order difference equation in price as a function of time; it may be solved for p_t in terms of past values of x_t. Equation (192) and its lagged counterparts may be used to substitute for $x_t, x_{t-1}, \ldots$ in the result. Such substitution yields another difference equation in price as a function of time. In addition to price, this difference equation contains current and lagged values of u_t and v_t. The new difference equation may be regarded as an autoregressive process if u_t and v_t are independently distributed; consequently p_t is serially correlated and the R+1 current and past values of price included in (194) are correlated with the residual term u_t'. As is well known, this correlation leads to statistical bias in the least-squares estimates of $b_0, b_1, \ldots, b_R$.

As indicated (193) is an s-order difference equation; under the appropriate initial conditions it has a solution of the form

$$p_t = A_0 (x_t - v_t) + A_1 (x_{t-1} - v_{t-1}) + \cdots \tag{197}$$

Phillips' <u>second condition</u> is that

$$A_0 = 0 \tag{198}$$

that is, that the system (192)-(193) is recursive. 45/ Otherwise, p_t depends on x_t and, from (192), x_t depends on u_t; hence, by (196) p_t is correlated with u_t'.

Suppose that Phillips' first and second conditions are fulfilled and suppose further that

$$A_i = 0 \quad \text{for} \quad i = 0, 1, \ldots, j$$

so that price is not affected by $x_t, x_{t-1}, \ldots, x_{t-j}$. Phillips' third condition is that the residuals u_t must not be autocorrelated over intervals containing more than j periods; that is, the correlogram of u_t should be such that

$$\frac{Eu_t u_{t-n}}{Eu_t^2} = \rho_n = 0 \quad \text{for} \quad n > j \tag{199}$$

To prove this we solve (192) and (197) for p_t. As Phillips (26, p. 110) points out, the general solution may be written

$$p_t = (L_1 u_{t-1} + L_2 u_{t-2} + \ldots) + (w_t + M_1 w_{t-1} + M_2 w_{t-2} + \ldots) \tag{200}$$

where

$$w_t = -A_0 v_t - A_1 v_{t-1} - \ldots \tag{201}$$

and where the L's and M's are functions of the b's of (192) and the A's of (197). According to Phillips (28, p. 110), if $A_i = 0$ for all i from 1 to j, where j is any integer, then $L_i = 0$ for all i from 1 to j; and if $A_i \neq 0$ for $i = j+1$ then $B_i \neq 0$ for all $i \geq j+1$. Multiplying (200) by u_t and taking expected values we have

$$\begin{aligned} Ep_t u_t = {} & (L_1 Eu_t u_{t-1} + L_2 Eu_t u_{t-2} + \ldots) + \\ & (Eu_t w_t + M_1 Eu_t w_{t-1} + \ldots) \end{aligned} \tag{202}$$

It follows from (202) that $Ep_t u_t \neq 0$ unless $\rho_n = 0$ for all n such that $L_n \neq 0$, that is, for all $n > j$.

Phillips' fourth condition is that u_t must be uncorrelated with the w_t, $w_{t-1}, \ldots$. This condition follows directly from equation (202). By (201) we see that if u_t is uncorrelated with $v_t, v_{t-1}, \ldots$, Phillips' fourth condition is satisfied.

Phillips' four conditions must be satisfied if we are to use least squares to estimate the coefficients of (192), but his conditions do not suggest what method or methods should be employed if they are not satisfied.

45/ See Wold and Jureen (36, p. 14).

This difficulty results because Phillips does not specify a model which generates a distributed lag or lags, but only assumes the existence of such lags. When we do specify a model (especially if we specify one of the types discussed above), the problem of what to do if least squares is not appropriate is readily suggested.

An Approach Which Can Be Used If the Method of Least Squares Is Not Directly Applicable

We have shown that when distributed lags are generated by the models discussed in this paper, equations containing them may be reduced to equations containing only discrete lags, albeit at the expense of introducing serial correlation in the residuals of the reduced equations if there is none to begin with in the original equations. An equation or group of equations may be reduced whether the equations deal with demand or supply; consequently we may write any system of equations containing distributed lags as a system of equations containing only discrete lags. Provided we can neglect the problem of serial correlation in the residuals of the reduced equations, standard techniques are available for dealing with the problem of simultaneity.

Before concluding this section, we briefly indicate the operation of this method in a particular case. Suppose we have a demand function

$$x_t = b\, p_t^* + c\, y_t^* \tag{203}$$

a consumption function

$$C_t = 0 + c'\, y_t^* \tag{204}$$

and a supply function

$$x_t = d\, p_t^{**} + e\, z_t \tag{205}$$

where z_t is an exogenous variable determining supply and p_t^{**} is the producers' expected normal price. Note that we assume that (1) income is exogenous, (2) current price and current income do not affect demand, (3) current price does not affect supply, and (4) that producers' and consumers' expectations are different. Let β be the consumers' coefficient of expectations for price, α be the consumers' coefficient of expectations for income, and $\beta*$ be the producers' coefficient of expectations. We neglect the problem of serial correlation. Equation (66) (see page 33) shows that the reduced equations of (203)-(204) are

$$x_t = b\,\beta\, p_t + c\,\alpha\, y_t + (1-\beta)\, x_{t-1} +$$

$$[(1-\alpha) = \frac{b}{c'}(1-\beta)]\, C_{t-1} \tag{206}$$

$$C_t = c'\ \alpha\, y_t + (1-\alpha)\, C_{t-1} \tag{207}$$

Equation (97) (see page 40) shows that the reduced equation of (205) is

$$x_t = d\beta * p_t + e\, z_t + (1-\beta *)\, x_{t-1} - e(1-\beta *)\, z_{t-1} \tag{208}$$

Equations (206) and (208) jointly determine p_t and x_t. y_t, x_{t-1}, C_{t-1}, z_t, and z_{t-1} are predetermined; consequently, equation (207) may be estimated by least squares. Equations (206) and (208) each are overidentified (whether or not we count C_t as a variable in the system) and may be estimated by the method of limited information.

DEMAND FOR DURABLE CONSUMPTION GOODS AND THE DISTINCTION BETWEEN CONSUMPTION AND SAVING 46/

The distinguishing characteristic of a durable commodity is that utility is derived from its services over time rather than from consumption of the commodity itself at a point in time. As a definition of a durable consumption good, however, this distinction is rather loose. We do not observe a consumer's utility derived from the consumption of goods and services but only his expenditures on the number of physical units of each which he purchases. Expenditure on goods and services frequently is called consumption, and is called the "cost of living" or "real income" by Fisher (14, pp. 6-8).

The utility which consumers derive from their expenditures on goods and services may not be accurately measured by those expenditures. Fisher (14, pp. 8-9) put the point well:

"The only ... discrepancy worth carefully noting is that which occurs when ... money spent is not simply for the temporary use of some object but for the whole object, which means merely for all its possible future uses. If a house is not rented but bought, we do not count the purchase price as all spent for this year's shelter. We expect from it many more years of use. Hence out of the entire purchase price, we try to compute a fair portion of the purchase price to be charged up to this year's use. In like manner, the statisticians of cost of living [that is, consumption] should distribute by periods the cost of using a person's house furnishings, clothing, musical instruments, automobiles and other durable goods, and not charge the entire cost against the income of the year of purchase. To any given year should be charged only that year's upkeep and replacement, which measures, at least roughly, the services rendered by the goods in question during that particular year [underlining mine]. The true real annual income from [that is, consumption of] such goods is the equivalent approximately of the cost of the services given off by those goods each year.

"Strictly speaking, then, in making up our income (consumption) statistics, we should always calculate the value of services, and never the value of the objects rendering those services. It is true that, in the case of short-lived objects like food, we do not ordinarily need, in practice, to go to the

---

46/ This topic is partially covered in Nerlove (27), with special reference to the demand for automobiles.

trouble of distinguishing their total cost from the cost of their use. A loaf of bread is worth ten cents [this was written in 1930] because its use is worth ten cents. We cannot rent food; we can only buy it outright. Yet there is some discrepancy ... in the case of foods that keep, such as flour, preserved foods and canned goods. These we may buy in one year but not use until a later year, and in such cases the money given for the food might almost be said to be invested rather than spent, like the money given for a house. A man who buys a basket of fruit and eats it within an hour is certainly spending his money for the enjoyment of eating fruit. But, if he buys a barrel of applies in the fall to be eaten during the winter, is he spending his money or is he investing it for a deferred enjoyment? Theoretically the barrel of apples is an investment comparable to a house or any other durable good. Practically it is classed as an expenditure, although it is a border-line case.

"Spending and investing differ only in degree, depending on the length of time elapsing between the expenditure and the enjoyment. To spend is to pay money for enjoyments which come very soon. To invest is to pay money for enjoyments which are deferred to a later time. We spend money for our daily bread and butter or for a seat at the theater, but we invest money in the purchase of bonds, farms, dwellings, or automobiles, or even suits of clothes."

It is important to note that Fisher uses investment in the sense that we usually use saving, namely as consumers' additions to their assets. Thus the quotation from Fisher not only sheds light on the nature of durable goods purchases but also on the distinction between consumption and saving. Purchases of durable goods are partially in the nature of saving just as are bond purchases, life insurance premiums, retirement fund contributions, social security payments, and additions to one's bank account. Expenditures on durable goods in excess of the use value of services rendered by them should be treated as savings and not consumption. 47/

Purchases of individual durable goods, automobiles, clothing, and the like, are therefore not entirely susceptible to the usual type of demand analysis. By treating the problem of the demand for durable consumption goods in terms of the flow of services from those goods, we can indeed apply the usual type of analysis. This approach, however, leaves a good deal to be desired, for the industry producing a durable good is directly interested not in the demand for the services of those goods but in the number of new units purchased or the total current expenditure for new units. Thus the cotton farmer or the wool grower is interested in the demand for the services of cotton or wool textiles only insofar as this demand affects new purchases of cotton or wool textiles.

47/ It follows that the Commerce Department's definition of aggregate consumption is not entirely appropriate for use in demand analyses of the type suggested in this paper.

At the beginning of this paper we emphasized the role which the durability of certain goods plays in causing changes in the quantities of various goods and services consumed to lag behind changes in prices or incomes; we argued that the durability of certain goods helped to implement the psychological, technological, and institutional factors causing distributed lags. In subsequent paragraphs we develop this idea more fully and show how the methods discussed in preceding sections of this paper may be used to study the new purchases of durable commodities. We do this by making use of an example concerning automobiles. 48/

An Example Relating to Automobiles

Given a homogeneous group of automobiles with given age and quality characteristics, we may reasonably assume that the flow of services is proportional to the number of automobiles in the group. This is just another way of saying that identical automobiles are perfect substitutes, so that the marginal rate of substitution between any two automobiles equals one. It is clear that we cannot consider the total number of automobiles in the hands of consumers as a homogeneous group: a 1937 car in the moderately-priced class can hardly be considered as yielding the same net flow of services as a 1956 car in a high-priced class. Cars of different makes or different ages are not perfect one-for-one substitutes; consequently the flow of services from the cars now in the hands of consumers cannot be considered proportional to their numbers. A possible solution to the problem of investigating the demand for automobiles is to consider the market for cars as a set of inter-related markets for close but not perfect substitutes. 49/ On the other hand, evidence presented by Chow (7), suggests that cars of different ages and/or makes may be perfect substitutes albeit with marginal rates of substitution different from one.

Estimating the stock of automobiles.--Let the stock of automobiles at any given time be a number to which the flow of services is proportional. We can derive the stock of automobiles from the numbers of different makes and ages in the hands of consumers only if such automobiles are perfect substitutes. If the market for new cars were perfect, in an economic sense, we could develop an index of new car purchases based on the value of new cars sold each year, but to do this we must assume that an automobile selling for $5,000, say, yields twice the flow of services as one selling for $2,500. Under this approach, we use the market prices to determine the marginal rates of substitution (which are constant if all new automobiles are perfect substitutes)

48/ The demand for automobiles has been ably treated by Roos and von Szeliski (29), Farrell (12), and Chow (7). The latter contains a useful summary of the studies of Roos and von Szeliski and Farrell. The discussion of long-run automobile demand presented here is not meant to be a study of the demand for automobiles but is designed only for illustrative purposes.

49/ This is the approach taken by Farrell (12).

among automobiles of different makes and, on this basis, form an aggregate in which make or quality is held constant. Although it is by no means clear that the market for new cars is perfect, we neglect the problems which may arise because of imperfections in the new car market.

If cars of different ages are perfect substitutes and if the used-car market were perfect, we could presumably develop an index of the stock of automobiles, adjusted for both age and make, based on an index of their total value. Observation suggests, however, that the market for used automobiles is far from perfect; so that, while we might stretch the facts a bit to assume that the new car market is perfect, we are less justified in assuming the used car market to be perfect. The difference in the degree of perfection of "new" and "used" markets for other durable goods is even more marked, with clothing as perhaps the most extreme example. Nevertheless, the assumption of constant marginal rates of substitution between automobiles enables us to derive, with one additional assumption, a stock of automobiles, adjusted for age, from the purchases of new cars.

The magnitudes of the marginal rates of substitution between cars of different ages are roughly indicated by the way in which automobiles depreciate. It is common knowledge that cars depreciate rapidly the first few years and more slowly thereafter; hence, constant percentage depreciation seems to indicate the pattern of substitution between automobiles of different ages. Let d be the percentage rate of depreciation, then, by our assumption, 1-d is roughly the marginal rate of substitution between two cars of a standard make differing in age by one year. Assume that new car purchases are adjusted for make and model. Let s_t be the stock of automobiles during period t, that is, the number of automobiles to which the flow of services is proportional; let x_t be new car purchases during t; let x_{t-1} be new car purchases during t-1; and so on. Under our assumptions

$$s_{t-1} = x_t + (1-d)\ x_{t-1} + (1-d)^2\ x_{t-2} + \dots \tag{209}$$

Thus, the stock of automobiles at any time is a special kind of function of new car purchases taken with a distributed lag. As discussed previously, this form of distributed lag allows the reduction of (209) to

$$s_t = x_t + (1-d)\ s_{t-1} \tag{210}$$

Under our assumptions, the demand for automobiles may be considered as the demand for a stock of automobiles adjusted for both age and make, that is, s_t.

Factors that affect the demand for the services of automobiles.--This demand for the services of automobiles depends on a variety of factors. These include: (1) the price of automobiles relative to other commodities, (2) real disposable income, (3) population, (4) the extent and quality of the highway network, and (5) the degree of urbanization and/or suburbanization. The long-run elasticities of demand with respect to these variables may differ from the corresponding short-run elasticities, and the quantity demanded may depend on expected normal price and income as well as on the current values of these variables. We may therefore formulate a variety of models.

Let p_t be the current price of automobiles relative to other commodities, y_t be current disposable income, and z_t be any other current factors which influence the demand for automobiles. 50/ The simplest model is to assume no difference between long- and short-run elasticities of demand due either to technological or institutional rigidities or to uncertainty. In this case the demand for the services of automobiles may be written in linear form as

$$s_t = a_1 p_t + a_2 y_t + a_3 z_t \tag{211}$$

For convenience we neglect the constant and residual terms.

Substituting (211) and (211) lagged one period in (210) we have

$$\begin{aligned} s_t &= a_1 p_t + a_2 y_t + a_3 z_t \\ &= x_t + (1-d) [a_1 p_{t-1} + a_2 y_{t-1} + a_3 z_{t-1}] \end{aligned} \tag{212}$$

Hence, x_t, new car purchases in period t, may be related to p_t, y_t, z_t and p_{t-1}, y_{t-1}, and z_{t-1} as follows

$$\begin{aligned} x_t = {} & a_1 p_t + a_2 y_t + a_3 z_t - \\ & a_1 (1-d) p_{t-1} - a_2 (1-d) y_{t-1} - a_3 (1-d) z_{t-1} \end{aligned} \tag{213}$$

under suitable conditions, the coefficients in (213) may be estimated by means of an ordinary least squares regression perhaps with the restraint that

$$\frac{a_1}{a_1 (1-d)} = \frac{a_2}{a_2 (1-d)} = \frac{a_3}{a_3 (1-d)} \tag{214}$$

To develop a model which includes rigidities of a technological or institutional nature, we introduce the notion of long-run equilibrium stock of automobiles, s_t^*, and suppose that consumers try to change their current stock in proportion to the difference between their long-run equilibrium stock and the stock they currently hold, that is, we suppose

$$s_t - s_{t-1} = \delta [s_t^* - s_{t-1}], \; 0 < \delta \leq 1 \tag{215}$$

Replacing s_t by s_t^* in (211) we have

$$s_t^* = a_1 p_t + a_2 y_t + a_3 z_t \tag{216}$$

50/ If both stock and income are taken on a per capita basis, z_t need not include population. Otherwise z_t includes population as well as other factors.

If we do not know the appropriate rate of depreciation, d, we cannot reduce equations (215) and (216) in such a way that we obtain equations involving only observable variables; instead we proceed as follows: Lag (215) one period, multiply through by (1-d) and subtract the result from (215); thus,

$$s_t = (1-d)\ s_{t-1} = \delta\ [s_t^* - d\ s_{t-1}^*] + (1-\delta)\ [s_{t-1} - d\ s_{t-2}] \tag{217}$$

From (217) and (210) we have,

$$x_t = \delta\ [s_t^* - d\ s_{t-1}^*] + (1-\delta)\ x_{t-1} \tag{218}$$

Substituting from (216) for s_t^* and s_{t-1}^* in (218) we have

$$x_t = a_1\ \delta\ p_t - a_1\ \delta\ (1-d)\ p_{t-1} + a_2\ \delta\ y_t - a_2\ \delta\ (1-d)\ y_{t-1} + a_3\ \delta\ z_t = a_3\ \delta\ (1-d)\ z_{t-1} + (1-\delta)\ x_{t-1} \tag{219}$$

which contains only observable variables. Under appropriate conditions, the coefficients in (219) may be estimated by ordinary least squares, perhaps with the restriction that

$$\frac{a_1\delta}{a_1\ \delta\ (1-d)} = \frac{a_2\delta}{a_2\ \delta\ (1-d)} = \frac{a_3\delta}{a_3\ \delta\ (1-d)} \tag{220}$$

We now drop the assumption that rigidities are due to technological or institutional causes, and assume, instead, that they are present due to uncertainty. Let p_t^* and y_t^* be expected normal price and expected normal income, respectively. For simplicity, let us assume that current price and current income have no effect on the demand for automobiles. Then we may write the demand equation as

$$s_t = a_1\ p_t^* + a_2\ y_t^* + a_3\ z_t \tag{221}$$

Substituting (221) in (210) we have

$$x_t = a_1\ p_t^* - a_1\ (1-d)\ p_{t-1}^* + a_2\ y_t^* - a_2\ (1-d)\ y_{t-1}^* + a_3\ z_t - a_3\ (1-d)\ z_{t-1} \tag{222}$$

or, if $H_t = x_t - a_3\ z_t + a_3\ (1-d)\ z_{t-1}$ and $g_1 = a_1$, $g_2 = -a_1\ (1-d)$, $g_3 = a_2$, $g_3 = -a_2\ (1-d)$, then

$$H_t = g_1\ p_t^* + g_2\ p_{t-1}^* + g_3\ y_t^* + g_4\ y_{t-1}^* \tag{223}$$

Our expectational equations are

$$p_t^* = \beta\, p_t + (1-\beta)\, p_{t-1}^* \tag{224}$$

and

$$y_t^* = \alpha\, y_t + (1-\alpha)\, y_{t-1}^* \tag{225}$$

By treating p_{t-1}^* and y_{t-1}^* as separate expectational variables and writing (224) and (225), lagged one period, as their expectational equations, we could, in principle, reduce (223) by the single equation method. As indicated in that discussion, the resulting reduced equation, which contains only discrete lags, includes 52 variables in addition to x_t, since each of the $4^2 = 16$ added non-expectational variables includes lagged values of z_t and z_{t-1} as well as x_t. The single equation method, directly applied, is clearly unmanageable.

Since we have specified no additional demand equations in which p_t^*, p_{t-1}^*, y_t^*, and y_{t-1}^* enter, we cannot directly apply the multiple equation approach. The situation, however, may be somewhat alleviated by specifying a consumption function, say

$$C_t = g_5\, y_t^* \tag{226}$$

Equations (223), (226), and (226) lagged one period form a separable system of demand equations; these may be separated prior to reduction. Although, as we have seen, such a procedure is not to be recommended generally, we have no choice in this case. Separation of (223), (226), and (226) lagged one period leads to the following equation:

$$H_t = g_1\, p_t^* + g_2\, p_{t-1}^* + \frac{g_3}{g_5}\, C_t + \frac{g_4}{g_5}\, C_{t-1} \tag{227}$$

The equation

$$K_t = H_t - \frac{g_3}{g_5}\, C_t - \frac{g_4}{g_5}\, C_{t-1} = g_1\, p_t^* + g_2\, p_{t-1}^* \tag{228}$$

may be reduced by the single equation method of reduction, taking (224) and (224) lagged one period as expectational equations. By (56) (see page 29), the result is

$$\begin{aligned} K_t = {} & g_1\, \beta\, p_t + g_2\, \beta\, p_{t-1} - g_1\, (1-\beta)\ \beta\, p_{t-1} - \\ & g_2\, (1-\beta)\ \beta\, p_{t-2} + 2\,(1-\beta)\, K_{t-1} - (1-\beta)^2\, K_{t-2} \end{aligned} \tag{229}$$

where K_t is defined in (228) and g_1 and g_2 are defined in (223). Substitution for K_t (and for H_t in K_t) shows that (229) involves 17 variables in addition to x_t. While this is a considerable improvement over the 52 obtained by applying the single equation method directly, it is still a large number of variables.

Conclusions

If we think that rigidities of an expectational nature are important to the understanding of the demand for consumer durables, we would do well to specify a system of demand equations for durable commodities before proceeding with any estimation. When distributed lags are introduced because of technological or institutional rigidities, it is not necessary to specify demand equations for additional durable commodities.

FRIEDMAN'S PERMANENT INCOME HYPOTHESIS AND ITS IMPLICATIONS FOR DEMAND ANALYSIS 51/

In our discussion of Friedman's hypothesis, we take up first the theoretical formulation which led Friedman to believe that, if consumption and income are defined correctly, consumption should tend to be a constant proportion of income independent of the absolute level of income. We next turn to Friedman's statistical specification, which is actually the key to the whole permanent income hypothesis. Following this discussion we examine some of the implications of Friedman's hypothesis, especially in regard to estimating the income elasticities of demand for individual categories of consumption and the appropriate procedure to be followed when combining cross-section and time-series data. In conclusion, we present a simple test of the adequacy of the permanent income hypothesis as applied to individual categories of consumption. Although the test is a crude one, it suggests that Friedman's hypothesis is inadequate in certain respects.

Theoretical Formulation

Friedman (16, Chap. II) first considers a simple 2-period model: Under conditions of perfect certainty, a consumers' tastes at a point in time, say period 1, may be summarized by a 2-dimensional system of indifference curves, as in figure 6. C_1, measured along the vertical axis, is the money value, measured in period 1 prices, of services consumed in period 1 (n.b., not expenditures on goods in period 1); C_2, measured along the horizontal axis, is the money value at period 2 prices of the services consumed in period 2. A point on an indifference curve thus represents a particular combination of consumption in the two periods. We must suppose that underlying each point there is a prior maximization process, that is, that the expenditures represented by C_1 and C_2 are optimally distributed at the prices expected in periods 1 and 2.

51/ Friedman (16) presents a detailed exposition of the permanent income hypothesis and many of its implications, complete with empirical applications and a discussion of several alternative theories. The present discussion does not aim at completeness; for further details the reader is advised to consult Friedman's own discussion.

INDIFFERENCE CURVES AND A BUDGET LINE

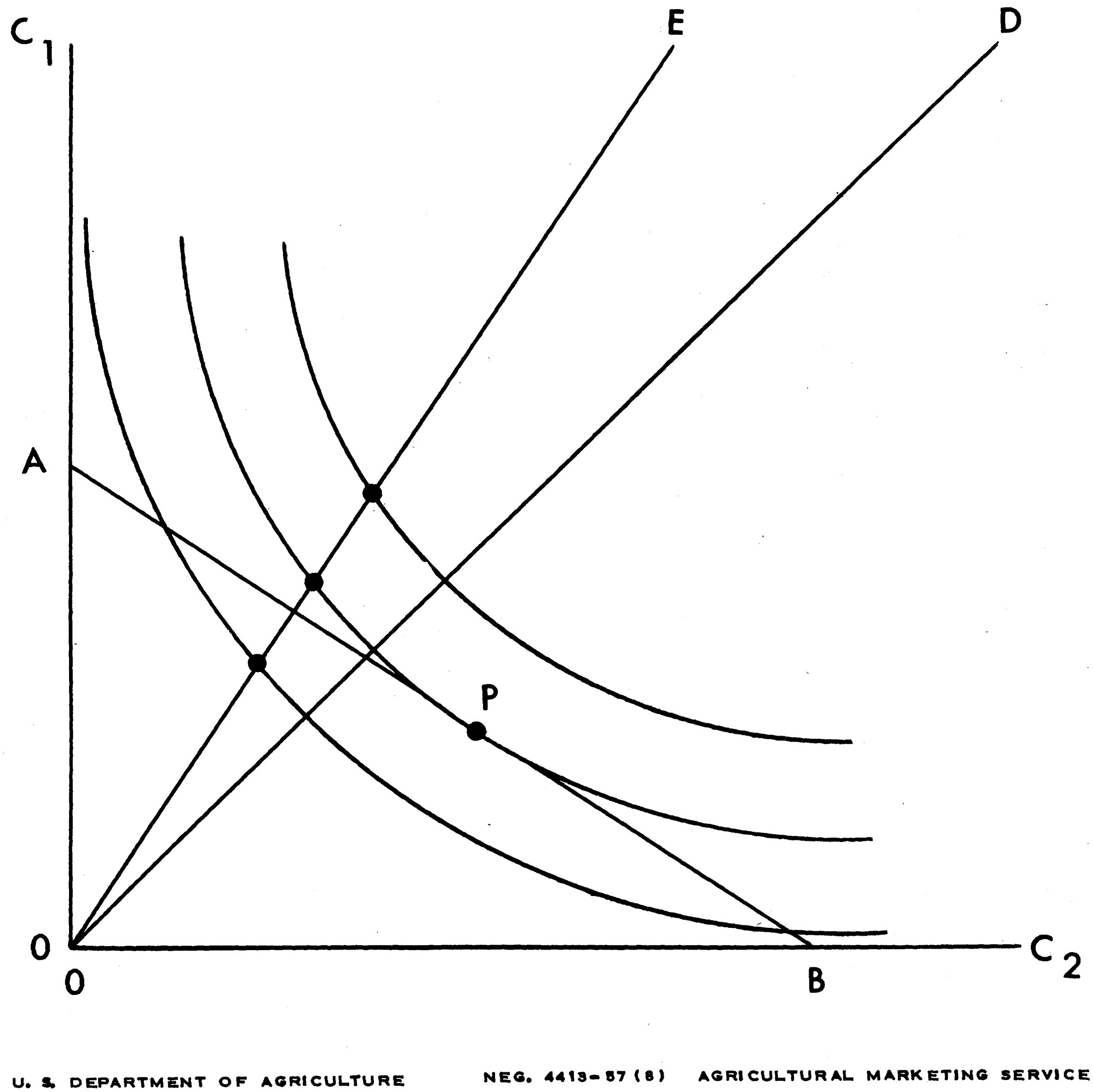


Figure 6.--Hypothetical indifference curves and a budget line of a consumer unit for consumption in two time periods. (From figure 1 in Friedman (16, p. 8).)

Let R_1 and R_2 be the consumer's expected receipts in periods 1 and 2 and let i be the rate of interest at which the consumer may borrow or lend freely. Then the maximum amount the consumer may consume in period 1, represented by the distance OA in figure 6, is

$$W_1 = R_1 + \frac{R_2}{1+i} \tag{230}$$

and the maximum he may consume in period 2, represented by OB, is

$$W_2 = R_1\ (1+i) + R_2 \tag{231}$$

The line AB thus represents the budget line of the consumer. It is clear from the diagram that the opportunities of the consumer, which we have represented by three variables R_1, R_2, and i, actually depend on only two, W_1 and i, or W_2 and i, that is, the slope of the budget line and its height.

As Friedman points out, this elementary formulation sheds considerable light on the usual view about the consumption function: "What we have been calling receipts in ... [period] 1 (R_1) or some slight modification thereof, is usually, and particularly in statistical budget studies, called 'income' and taken as the variable on which consumption depends. Now in our simple case it is clear that consumption in ... [period] 1 does not depend directly on R_1 at all; a change in R_1 affects consumption only through its effect on W_1 and, if accompanied by an appropriate opposite change in R_2, may not affect consumption at all. This is clearly eminently sensible: if a consumer unit knows that its receipts in any one year are unusually high and expects lower receipts subsequently, it will surely tend to adjust its consumption to its 'normal' receipts rather than to its current receipts." (16, pp. 9-10.) *

Friedman also points out the inadequacy of the usual measure of consumption expenditures: "We have been using the term consumption to designate the value of the *services* that it is *planned* to consume during the period in question, which, *under conditions of certainty*, would also equal the value of the services actually consumed. The term is generally used in statistical studies to designate actual expenditures on goods and services. It therefore differs from the value of services it is planned to consume on two counts; first, because of additions to or subtractions from the stock of consumer goods, second, because of divergencies between plans and their realization." (16, p. 11.)

It can be seen from the simple 2-period model that we already have the germ of the type of theory enunciated in this paper concerning distributed lags due to uncertainty. As Friedman has stated in his theory, however, it does not bear a *direct* relation to the theory discussed at the beginning of this paper. The latter actually provides additional reasons for supposing that current consumption and current income, as usually measured, do not relate directly to the pure theory of consumer behavior.

In addition to supposing that the indifference curves of figure 6 are negatively sloped and convex to the origin (for the usual reasons), Friedman supposes that there is an absence of "time preference proper," that is, that the indifference curves are symmetrical around OD (the 45° line) so that C_1 and C_2 could be interchanged without altering the curves. It is not unreasonable to suppose, further, that the indifference curves have common slopes at the point where they intersect the line OD; thus, for a given level of W_1 or W_2 and a given interest rate, a consumer tends to consume the same proportion of receipts in period 1, whether those receipts are large or small. That is, in this simple case, the rate at which the individual is willing to substitute consumption in period 2 for consumption in period 1 depends only on the ratio of consumption in the two periods, not on the absolute level of consumption. In other words, Friedman supposes the consumption function to be homogeneous of degree one in W_1 or W_2. Friedman argues that this is sensible since C_1 and C_2 are of the "same stuff, differing only in dating." Since, "... it is hard to see any reason why this difference should have an asymmetrical effect ...," we may reasonably question the "initially plausible conjecture that the ratio of consumption to income decreases with income, if income is appropriately defined as a flow that can be permanently maintained [that is, income defined in terms of W_1 or W_2]." (16, p. 13.) *

For indifference curves that satisfy Friedman's assumptions, the consumption function assumes a particularly simple form

$$C_t^* = k(i) \cdot y_t^* = k(i) \cdot i \cdot W_t \tag{232}$$

where the asterisks on the variables C_t^* and y_t^* indicate consumption and income variables, respectively, which may differ from consumption and income as usually measured. $k(i)$ is a constant, supposed to depend only on the rate of interest, 52/ and the last equation indicates that y_t^* is related to the amount of wealth (both human and non-human) at the disposition of the consumer. In writing (232), Friedman has generalized from the 2-period case, while still abstracting from uncertainty. In subsequent discussion, Friedman goes considerably beyond the simple formulation in (232), but for our purposes it should suffice.

Statistical Specifications

The real interest in the permanent income hypothesis lies, not in Friedman's formulation of the pure theory of consumption behavior as expounded above, but in his treatment of the relation between the unobservable variables, C_t^* and y_t^* (called by Friedman "permanent" consumption and "permanent" income, respectively) and the variables C_t and y_t, which are the consumption and income we actually measure. Friedman treats this relation in a chapter entitled "The Permanent Income Hypothesis" (Chap. III), and this chapter is the key to the permanent income hypothesis. Friedman states the basic problem well: "The magnitudes termed 'permanent income' and 'permanent consumption' that play such a critical role in the theoretical analysis cannot be observed directly for any individual consumer unit. The most that can be observed are actual receipts and expenditures during some finite period, together, perhaps, with some verbal statements about expectations for the future. The theoretical constructs are *ex ante* magnitudes; the empirical data are *ex post*. Yet in

52/ In actual empirical analyses of time series data Friedman neglects the effects of changes in the rate of interest.

order to use the theoretical analysis to interpret empirical data, a correspondence must be established between the theoretical constructs and the observed magnitudes." The usual way of establishing such a correspondence has been simply to treat current consumption expenditures and current income as if they were the theoretical constructs. Friedman's approach is different.

Let y_t represent the measured income for a given period of time, and let C_t represent his measured consumption expenditures for the same period. Friedman proposes the division of these measured magnitudes into two parts: 53/ The first may be called the permanent component, y_t^* or C_t^*; the second may be called the transitory component, y_t^T or C_t^T.

Thus the permanent income hypothesis now consists of three equations:

$$C_t^* = k(i)\ y_t^* \tag{232}$$

$$y_t = y_t^* + y_t^T \tag{233}$$

$$C_t = C_t^* + C_t^T \tag{234}$$

However, the system of equations (232)-(234) is not, perhaps, the most reasonable form of the permanent income hypothesis. On several grounds the logarithmic form seems preferable; consequently we may write

$$\log C_t^* = \log k(i) + \log y_t^* \tag{232'}$$

$$\log y_t = \log y_t^* + \log y_t^T \tag{233'}$$

$$\log C_t = \log C_t^* + \log C_t^T \tag{234'}$$

In his theoretical development Friedman frequently switches from the linear to the logarithmic form and back, but in his empirical applications he usually uses the logarithmic form.

The interpretation of the permanent and transitory components of either income or consumption is slightly different depending on whether we think of the hypothesis primarily in the context of budget studies (cross-section data) or primarily in the context of time-series analyses of the aggregate consumption function. In his chapter III, "The Permanent Income Hypothesis," Friedman gives the following interpretation:

53/ As Friedman points out, this division into two components is arbitrary. At one point in his monograph, Friedman generalizes his assumption so that the observed magnitudes are divided into an arbitrary number of components. The choice of two components was made primarily on the grounds of simplicity. (See footnote 2 (16, p. 22) and (16, p. 186).)

"The permanent component [of income] is to be interpreted as reflecting the effect of those factors that determine the consumer unit's capital value or wealth: the non-human wealth it owns; the personal attributes of the earners in the [consumer unit under consideration] ... such as their training, ability, personality; the attributes of the economic activity of the earners, such as occupation followed, the location of the economic activity, and so on. It is analogous to the 'expected' value of a probability distribution. The transitory component is to be interpreted as reflecting all 'other' factors, factors that are likely to be treated by the unit affected as 'accidental' or 'chance' occurences, though they may, from another point of view, be the predictable effect of specifiable forces, for example, cyclical fluctuations in economic activity." (16, pp. 21-22.)

Two types of forces produce the transitory component: The first is that specific to an individual consumer unit. The second is not specific to an individual unit but affects all or part of the group of consumer units under consideration. For the group as a whole, the transitory factors affecting specific consumer units tend to cancel out by the law of large numbers, so that the mean transitory component, if caused by factors of the first type, generally tends towards zero. On the other hand, transitory factors which affect all or a large number of the members of the group under consideration do not tend to cancel one another. In our earlier discussion of distributed lags due to uncertainty, our primary emphasis was on transitory factors which do not tend to cancel out. These are likely to be important in time series analyses and their importance depends on the nature of the period covered. Factors specific to individual consumer units are of importance primarily in the analysis of budget data and they depend on the nature of the group being studied.

The transitory forces affecting consumption are likely to be specific to particular consumer units. Consequently, Friedman argues that the transitory component of consumption may be assumed to have zero mean when we deal with time series data. However, on the basis of our discussion in an earlier section of this paper, it is clear that rigidities in consumer behavior may arise from causes of an institutional or technological nature. In this case, we may well regard current consumption expenditures for an entire group of consumer units as having transitory components whose mean is not zero. It therefore seems unreasonable to assume before hand that the transitory component of consumption averages zero when we deal with time series data.

In the general form stated in equations (232)-(234) or (232')-(234'), the permanent income hypothesis is empty: two additional equations have been specified but so have two additional variables. No empirical data could contradict the hypothesis as it stands. Additional assumptions are therefore necessary. The particular additional assumptions that Friedman makes are as follows:

Assumption I: The transitory components of income and consumption are uncorrelated with one another and with the corresponding permanent components.

Assumption II: The mean transitory components of consumption and income are both zero. 54/

Assumption II is clearly unjustified for time-series data. Its chief justification for cross-section data is that it enables Friedman to explain, interpret, and predict a wide variety of empirical phenomena. The justification for Assumption I is given by Friedman as follows:

"[The assumption that the correlation between the transitory components and the corresponding permanent components] ... are zero seem[s] very mild and highly plausible. Indeed, by themselves, they have little substantive content and can almost be regarded as simply completing or translating the definitions of transitory and permanent components: the qualitative notion that the transitory component is intended to embody is of an accidental and transient addition to or subtraction from income, which is almost equivalent to saying an addition or subtraction that is not correlated with the rest of income. ... The assumption that [the correlation between the two transitory components of income and consumption] ... is zero is a much stronger assumption. It is primarily this assumption that introduces important substantive content into the hypothesis and makes it susceptible of contradiction by a wide range of phenomena capable of being observed. ... The common notion that savings ... are a 'residual' speaks strongly for the plausibility of the assumption. For this notion implies that consumption is determined by rather long term considerations, so that any transitory changes in income lead primarily to addition to assets or to the use of previously accumulated [cash] balances rather than to corresponding changes in consumption. ... Yet from another point of view, the assumption seems highly implausible. Will not a man who receives an unexpected windfall use at least some part of it in 'riotous living,' in consumption expenditures? Is it not unreasonable to suppose that he will add the whole of it to his wealth? In answering these questions much depends on how 'consumption' is defined. The offhand affirmative answer reflects in large measure, I believe, an implicit definition of consumption in terms of purchases, including durable goods, rather than in terms of the value of services. If the latter definition is adopted, as seems highly desirable in applying our hypothesis to empirical data, much that one offhand classifies as consumption is reclassified as savings. Is not the windfall likely to be used for the purchase of durable goods? Or, to put it differently, are not the timing of the replacement of durable goods and additions to the stock of such goods likely to some extent to be adjusted so as to coincide with windfalls?" (16, pp. 26-28.)

54/ Assumption II actually is not necessary for the development of the permanent income hypothesis. Friedman indicates this but makes the assumption primarily to facilitate exposition. We make it here for the same reason: It renders the distinction between the permanent income hypothesis as applied to budget data and the permanent income hypothesis as applied to time series data more clear-cut. Assumption I applies to both cross-section and time-series data; Assumption II is relevant only in the case of cross-section data. (See (16, p. 30).)

Friedman warns us not to interpret the permanent components as corresponding to average life-time values: "It is tempting to interpret the permanent components as corresponding to average lifetime values, and transitory components as the difference between such life-time averages and the measured values in a specific time period. Such an interpretation is not, however, appropriate, and this for two reasons. In the first place, the experience of one unit is itself but a small sample from a more extensive hypothetical universe, so there is no reason to suppose that transitory components average out to zero over the unit's lifetime. In the second place, and more important, it seems neither necessary nor desirable to decide in advance the precise meaning to be attached to 'permanent.' The distinction between permanent and transitory is intended to interpret actual behavior: we ... treat consumer units <u>as if</u> they regarded their income and their consumption as the sum of two such components, and <u>as if</u> the relation between the permanent components is the one suggested by our theoretical analysis. This general approach is suggested by theoretical considerations, but the precise line to be drawn between permanent and transitory components is best left to be determined by the data themselves, to be whatever seems to correspond to consumer behavior." (<u>16</u>, p. 23.)

Thus the notion of permanent income corresponds to what we would call "expected normal income;" as such it should be interpreted, not as average lifetime earnings, but as the income a consumer unit expects as his normal income, where his expectations hold only for a finite period into the future. The length of this period may be called the particular person's economic horizon. The concept of "permanent" is related to the length of this horizon: the shorter the horizon, the more of any given income change will be considered permanent.

When the hypothesis is interpreted in logarithmic form, Assumptions I and II should, of course, be construed in logarithmic terms as well.

Specification of the permanent income hypothesis in the form (232)-(234) or (232')-(234') under Assumptions I-II enables Friedman to interpret the nature of the usual statistical relationship between measured consumption and measured income. Suppose that we compute the least squares regression of C_t on y_t, then we have, say

$$C_t = a + b\ y_t \tag{235}$$

where a and b are the estimated coefficients. The regression may be taken over individuals or over time periods, however, maintenance of Assumption II suggests that we should consider (235) in the context of a budget study.

Friedman shows that the least squares regression coefficient b may be interpreted as

$$b = k\ (i) \cdot P_y \tag{236}$$

where P_y is the fraction of the total variance of income in the group [or over the period, if we are dealing with time series] contributed by the permanent component of income. As Friedman states: "The regression coefficient b measures the difference in consumption associated, on the average, with a one dollar difference between consumer units in measured income. On our permanent income hypothesis, the size of this difference in consumption depends on two things: first, how much of the difference in measured income is also a difference in permanent income, since only differences in permanent income are regarded as affecting consumption systematically; second, how much of permanent income is devoted to consumption. P_y measures the first; k, the second; so their product equals b. If P_y is unity, transient factors are either entirely absent or affect the incomes of all members of the group by the same amount; a one dollar difference in measured income means a one dollar difference in permanent income and so produces a difference of k in consumption; b is therefore equal to k. If P_y is zero, there are no differences in permanent income; a one dollar difference in measured income means a one dollar difference in the transitory component of income, which is taken to be uncorrelated with consumption [by Assumption I]; in consequence, this difference in measured income is associated with no systematic difference in consumption; b is therefore zero." (16, p. 32.)

If Assumption II holds, that is, if the mean transitory components of both income and consumption are zero, and if permanent consumption is proportional to income, as is continually assumed in the permanent income hypothesis, (236) yields an extremely simple interpretation of the regression of consumption on income. Let η_{cy} be the elasticity of measured consumption on measured income as computed at the mean values of measured consumption and measured income, for the linear case, or as the coefficient of log y_t, in the logarithmic case. Friedman shows that

$$\eta_{cy} = P_y \tag{237}$$

at mean measured consumption and income in the linear case and throughout in the logarithmic case. That is, the income elasticity of aggregate consumption, as measured from budget data, measures not the elasticity of permanent consumption to permanent income (the theoretically relevant variables) but the proportion of the variance of measured income in the sample contributed by variation in the permanent component! On Friedman's interpretation, therefore, the regression of measured consumption on measured income tells us nothing about the relation of consumption to income but rather something about the relation between the distributions of wealth and of measured income in the sample under consideration.

It should now be clear why, in dealing with cross-section data, Assumption II is an integral part of the theory, although, as we have remarked, it is not a necessary part. If the mean transitory components of consumption and income equal zero for the sample under consideration, then the ratio of the average consumption for the group to the average income for the group measures k(i). The elasticity of measured consumption on measured income measures P_y. Both parameters can be identified. If Assumption II does not hold, neither P_y nor k can be measured separately; only their product can be measured.

A similar situation arises in the case of application of the permanent income hypothesis to time-series data. As we have indicated, Assumption II is unreasonable when applied to aggregate consumption and income over time; consequently, if only Assumption I could be made, the regression of measured aggregate consumption on measured aggregate income over a period of time would tell us something about the product of $k(i)$ with P_y, where P_y measures the contribution of the permanent component to the total variance of income over the period in question; it would not tell us anything about $k(i)$ and P_y separately: $k(i)$ could not be measured as the ratio of average measured aggregate consumption to average measured aggregate income and P_y could not be measured by the elasticity of measured consumption with respect to measured income. An additional assumption or assumptions must therefore be introduced.

In Chapter V, "Consistency of the Hypothesis with Existing Evidence on the Relation between Consumption and Income: Time Series Data" (16, pp. 115-156), Friedman specifies additional assumptions appropriate in the context of time-series data. On the basis of an examination of existing time-series studies of the relation between income and consumption, Friedman concludes that permanent income should be susceptible of representation as a weighted average of incomes for current and past years. The transitory component of consumption is taken to be a purely random error term; that is, no systematic relationship between the transitory components of different time periods exists and the transitory components have mean zero. The possible existence of rigidities of consumer behavior of a technological or institutional nature makes this assumption seem somewhat doubtful.

The particular model which Friedman constructs to represent the relation between what people consider the permanent component of income to be and past measured income is closely related to the model derived from Hicks' definition of the elasticity of expectations. Friedman assumes that

$$y_t^* = \alpha \int_{-\infty}^{t} e^{-\alpha [t-\lambda]} y_\lambda \, d\lambda \quad \underline{55}/ \tag{238}$$

Equation (238) may be derived from the following equation:

$$\frac{dy_t^*}{dt} = \alpha [y_t - y_t^*] \tag{239}$$

55/ Friedman actually allows for the expectation of a time-trend in permanent income. But this refinement need not concern us here.

In the actual fitting procedure, (238) is approximated by a discrete summation. The fitting procedure that Friedman uses is essentially the same as the iterative procedure described above.

which is the differential equation analogue of (33) (see page 24). In the case of (238), however, static expectations, that is, the belief that the permanent component of income is equal to measured income, is given by $\alpha = \infty$ rather than $\alpha = 1$, as when the discrete representation of (33) is assumed.

Implications of Friedman's Hypothesis

Of the many implications of Friedman's permanent income hypothesis for the analysis of the demand for individual commodities, we can discuss only a few of the most general: (1) The effect of the type of group covered by the sample on the income elasticities for individual items of consumption derived from cross-section data; (2) the effect of the length and type of period covered on the income elasticities derived from time-series data; and (3) the valid way to combine income elasticities derived from cross-section data with other time-series data in a demand analysis.

Measurement of income elasticities from cross-section data.--Consider the planned expenditures on, or the planned quantity to be consumed of, an individual item of consumption like food. We may expect this quantity to be related, via consumer tastes and preferences, to the prices (current and/or expected) of food and other items, and to the income the family expects to receive, or the permanent component of income. If current and expected prices may be taken as the same, we may suppose, without loss of generality, that each consumer unit in a cross-section faces approximately the same prices [current and/or expected]. 56/

In Chapter VIII, Friedman describes the situation as follows: "[A consumer unit's] ... measured expenditures on food differs from its planned expenditures because of a transitory component of food expenditures, and its measured income differs from its permanent income because of a transitory component of income. When the regression of measured expenditures on measured income is computed from budget data for a group of families--the regression that has come to be called an 'Engel curve'--the transitory component of food expenditures tends to average out, 57/ but the transitory component of income

---

56/ This is, of course, the usual assumption with regard to prices in any cross-section analysis. Dropping the assumption of static price expectations leads to difficulties similar to those which are encountered when we drop the assumption that measured income is the relevant income variable. In the latter case, we assume that diversity of past experience leads to expected incomes systematically different from measured incomes; we neglect the distribution of coefficients of expectations among the members of the sample. In the case of prices this distribution is of principal importance. A discussion of the implications of non-static price expectations in cross-section analysis would take us too far afield here, however.

57/ As indicated on page 102, this may not be the case if rigidities exist in consumer behavior of a technological or institutional nature.

does not. ... In consequence, the elasticity of measured expenditures with respect to measured income reflects not only ... [consumers'] tastes and preferences but also the transitory components of income.

"Let c_f stand for the mean observed consumption on food of families with a given measured income, and assume that the transitory component of food expenditures is uncorrelated with the permanent or transitory component of income and averages zero for the group as a whole, so that c_f be regarded as the mean permanent component of food expenditures. The elasticity of c_f with respect to measured income then is

$$\eta_{c_f y} = \frac{dc_f}{dy} \cdot \frac{y}{c_f} = \frac{dc_f}{dy_f^*} \cdot \frac{dy_f^*}{dy} \cdot \frac{y}{y_f^*} \cdot \frac{y_f^*}{c_f}$$

$$= \frac{dc_f}{dy_f^*} \cdot \frac{y_f^*}{c_f} \cdot \frac{dy_f^*}{dy} \cdot \frac{y}{y_f^*}$$

$$= \eta_{c_f y_f^*} \cdot \eta_{y_f^* y} \quad \text{58/} \tag{240}$$

[where y_f^* represents the permanent component of measured income based on an economic horizon for food expenditures]. 59/

"But, on our hypothesis, $y^* = \frac{1}{k(i)} c^*$, 60/ which means that [if $y_f^* = y^*$]

$$\eta_{y_f^* y} = \frac{dy_f^*}{dy} \cdot \frac{y}{y_f^*} = \frac{1}{k(i)} \frac{dc^*}{dy} \cdot \frac{k(i)y}{c^*}$$

$$= \frac{dc^*}{dy} \cdot \frac{y}{c^*} = \eta_{c^*y} \tag{241}$$

58/ The notation in (240) is altered slightly so as to correspond to the notation used elsewhere in this paper.

59/ The rationale for distinguishing between permanent components appropriate to various categories of consumption and that appropriate to total consumption is given below. Although Friedman explicitly assumes that the same concept of the permanent component applies equally to total consumption and its individual categories, he is in some doubt as to its validity. As we shall see, differences in the concepts of permanent income appropriate to various categories of consumption is one of the key reasons for doubting the adequacy of Friedman's permanent income hypothesis.

60/ See equation (232) on page 96.

so that

$$\eta_{c_f\,y} = \eta_{c_f\,y_f^*} \cdot \eta_{c^*y} \quad \underline{61}/ \tag{242}$$

The first elasticity on the right hand side, between permanent food expenditures and permanent income, reflects the influence of tastes and preferences proper; the second, the influence of transitory factors affecting income.

"It follows that the differences among groups of families in the observed income elasticity of particular categories of consumption cannot be interpreted as reflecting solely the influence of differences in tastes or of differences in prices or similar factors affecting opportunities [such as income, sic!]; they may [also] reflect a third set of forces, namely, differences in a particular characteristic of the income distribution, the importance of transitory components of income. 62/" (16, pp. 206-207.)

If the elasticity of measured food expenditures, or, for that matter, expenditures on any particular category of consumption, with respect to measured income depends in part on the contribution of the permanent component to the variance of measured income, it follows that the income elasticity obtained from a budget study depends crucially on the group of households covered by the sample.

For example, consider two groups one of which typically has highly variable incomes and the other of which typically has stable incomes (for example, farmers versus civil servants). On the basis of the permanent income hypothesis we would expect the income elasticity for a particular consumption good to be lower for the first group than for the second, <u>even if the distribution of tastes, income, and the like were the same for the two groups</u>. The reason for this is simply that, for the group whose incomes are typically highly variable, the permanent component varies much less than measured incomes; whereas, for the group whose incomes are quite stable, the variation of the measured income within the group is accounted for almost entirely by the variation of the permanent component. Thus we would expect both η_{y^*y} and P_y to be smaller for the first group than for the second.

61/ The notation in (241) and (242) is altered slightly from Friedman's original notation.

62/ This statement in qualitative form follows directly from (240). If we wish to express it in usable form

$$\eta_{c_f\,y} = \eta_{c_f\,y^*} \cdot P_y \tag{i}$$

we must assume, first, that $y_f^* = y^*$ and, second, that the mean transitory components of food consumption and total consumption equal zero. Both assumptions are questionable.

Similarly, if we compare the elasticities of expenditure on particular categories of consumption with respect to measured income for two representative samples, one of the urban population of the United States and the other of the total population of the United States, we might expect to find the former higher than the latter, since a sample of the urban population excludes the farm population and this population might be expected to have more highly variable incomes than those of the urban population. In general, we may suppose that the more typically stable a group's incomes the more nearly will an elasticity of measured consumption expenditures (for total consumption or for an individual category of consumption) on measured income tend to approximate the elasticity with respect to the permanent component of income <u>appropriate to that group</u>.

<u>Measurement of income elasticities from time-series data</u>.--In the case of cross-section data, the characteristics of the group sampled are of crucial importance in the interpretation of income elasticities; in the case of income elasticities computed from time-series data, the characteristics and, particularly, the length of the period considered are of crucial importance. In Chapter V, Friedman makes this point in reference to total consumption as follows:

"The length of the period is important because, other things the same ... [the contribution of the permanent component to the variance of measured income], and so the observed income elasticity, can be expected to be larger, the longer the period covered, provided that the society in question is undergoing a systematic secular change in income. The total variance of [measured] income equals the variance contributed by the transitory component plus the variance contributed by the permanent component, given our assumption that the two components are uncorrelated. The variance contributed by the transitory component is not systematically affected by lengthening the period: by definition, the transitory components are largely random and short-lived. True, the variance may be larger at one time than another--this is why the historical characteristics of the period are important--but there is no reason why it should be systematically larger or smaller for a long than for a short period. 63/ The variance contributed by the permanent component, on the other hand, tends to be systematically larger, the longer the period covered, for the more widely separated two dates are, the larger will tend to be the secular difference in income between them. ... the ratio of the variance contributed by the permanent component to the total variance [of measured income], will therefore tend to be higher, the longer the period, and to approach unity as the period is indefinitely lengthened. If secular change were the only

---

63/ Friedman points out that this statement should be taken as referring to the variance of logarithmic components rather than the actual values. With a secular increase in income, we expect the variance of the transitory component, expressed in actual values, to increase over time; hence, unless we express our variables in logarithms, we expect the variance of the transitory component to be systematically larger for longer than for shorter periods.

source of variation in the permanent component the lower limit of ... [the ratio of the variance contributed by the permanent component to total variance] would be zero and this limit would tend to be approached as the length of the period covered approached zero. Since there are other sources of variation in the permanent component [over time], all one can say is that ... a lower limit greater than zero [will tend to be approached] as the length of the period approaches zero." (16, pp. 125-127.)

Friedman finds that this expectation is well fulfilled by the elasticities computed from regressions of measured total consumption on measured income for periods of different lengths: the elasticities are systematically lower for shorter than for longer periods. With certain qualifications, this implication of the permanent income hypothesis may be extended to individual categories of consumption: If we are willing to accept the assumptions underlying equation (242) (see page 105), then the elasticity of the demand for food with respect to measured income, obtained from time series data, tends to be 64/ the product of the elasticity with respect to permanent income and the proportion of the variance of measured income contributed by the permanent component. Thus, for example, the income elasticity computed for the period between the World Wars I and II should be lower than the elasticity computed for a period including both the interwar and postwar periods. 65/

The only reservation we may have to this as a general statement is that the assumptions on which (242) is based may not be fulfilled when we deal with elasticities estimated from time-series: First, the permanent component appropriate to the particular category of consumption under consideration may not be the same as that appropriate to total consumption. Thus, if the permanent component appropriate to, say, food is all of measured income, the income elasticity computed for the interwar period might be greater than, less than, or equal to the elasticity computed for both interwar and postwar periods. Second, systematic transitory components in the expenditure devoted to the particular category may occur, for the reasons mentioned above; in general, however, this might be expected to strengthen the qualitative conclusions based on the permanent income hypothesis but the quantitative relationship, (242), would no longer hold.

Combining income elasticities from cross-section data with other time-series data.--Several recent studies of the demand for individual commodities, for example, Stone (30), Wold and Jureen (36), and Tobin (33), attempt to combine cross-section and time-series data. The procedure is generally to obtain an income elasticity from a cross-section sample and to assume that this elasticity applies, with or without certain minor adjustments, to the aggregates

---

64/ "Tends to be" rather than "is", since the correlations between prices and measured income are not taken into account.

65/ This is in fact true for meat; see Nerlove, Marc, The Predictive Test as a Tool for Research: The Demand for Meat in the United States, M.A. thesis, the Johns Hopkins University, 1955. Numerous other examples could probably be cited.

over time. The income elasticity so obtained is inserted into the demand function and the remaining parameters are estimated from time-series data. It is clear that this procedure is inconsistent with the permanent income hypothesis; as Friedman states in Chapter V: "On our hypothesis, income elasticities of [total] expenditures computed from time series and from budget data are estimates of different things. Neither tells us anything directly about consumption behavior. ... Both measure instead a feature of the income structure, and they measure different features. The budget elasticity measures the fraction of the variance among incomes of a group of consumer units <u>at a point in time</u> contributed by differences in permanent components. The time series elasticity measures the fraction of the variance among aggregate or per capita incomes of a <u>series of time units</u> contributed by differences in permanent components. ... our conclusion that the elasticities computed from budget data and from time series are estimates of different magnitudes applies also to the elasticity for a particular category. ... The income elasticity computed from budget data can be expected to be the same (on the average) as that computed from time series data for a particular span of years only if transitory components of income have the same importance for the two bodies of data. There is no reason to expect the transitory components to have the same importance, and ..., if they do for one span of years, they will not for a longer or shorter span." (<u>16</u>, pp. 136-137.)

On the basis of equation (242) Friedman suggests a way to combine cross-section and time series data which is consistent with the permanent income hypothesis: Simply divide the income elasticity for a particular commodity by the income elasticity of total consumption expenditures, both estimated from the same budget study. In this way we obtain an estimate of the elasticity of expenditures on, or demand for, the particular commodity under consideration with respect to permanent income. This estimate is a valid one only on two assumptions: (1) the same concept of permanent income appropriate to the particular commodity is appropriate to total consumption; and (2) the mean transitory components of expenditure on, or consumption of, the commodity and of total consumption are zero. An estimate of aggregate permanent income over time may be constructed from the procedure which Friedman uses to estimate the consumption function from time series data, or by the non-iterative procedures suggested in this paper. The resulting series and the estimated elasticity with respect to permanent income may be combined with other time-series data to obtain estimates of the other parameters which appear in the demand function for the commodity under consideration.

We have already indicated the reason why the second of the two assumptions underlying the procedure discussed above may be suspect; Friedman himself suggests that the relation between the permanent and transitory components of income may be interpreted in terms of the length of the economic horizon of the consumer, hence: "One possible source of difficulty with this approach [that is, the procedure discussed above] is the necessity of taking permanent income to mean the same thing for the different categories of consumption. We have interpreted the exact meaning of permanent income in terms of the horizon of the consumer unit. Now there seems no reason why the horizon should be the same for all individual categories of consumption and some why it should differ systematically. For example, it seems highly plausible that housing expenditures are planned in terms of a longer horizon, and so a different concept of permanent income, than expenditures on, say, food. If this turns out to be a meaningful way of looking at the problem, the concept of permanent income applicable to total consumption will have to be regarded as an average of the concepts applicable to each category.

If both consumption and income are properly defined, it does not seem reasonable that the horizons for different categories of consumption should differ greatly from one another. Only if indivisabilities, difficulties of short-run substitution, or various institutional factors are introduced would the concept of differing horizons appear to be useful, but, as we have indicated elsewhere in this paper, it may be useful to treat this type of rigidity differently from expectational rigidity in consumer behavior. We thus think of expenditures for housing in terms of rental (or rental value) rather than in terms of purchase price; if it were not for imperfections in the capital and housing markets, an individual consumer experiencing a rise in his income which he considered to be permanent with respect to total consumption, or any of its categories, would immediately adjust his housing expenditure. Even if we accept Friedman's view that the concept of permanent income might be different for different categories of consumption, it is plausible that it would not differ greatly for highly similar commodities. In any case the attractive simplicity of Friedman's permanent income hypothesis is greatly reduced if one must assume different concepts of permanent income for different categories of consumption.

Tests of the Permanent Income Hypothesis

First test.--The notion that the concept of permanent income appropriate to individual categories of consumption ought really to be the same from category to category at once suggests a simple test of the adequacy of the permanent income: When dealing with time-series, Friedman's model suggests that both total consumption and each individual category of consumption depends on income taken with a distributed lag. No mention is made of whether the consumption of particular commodities depends on prices taken with a distributed lag, but it is explicitly assumed that the transitory components of total consumption and the consumption of individual items have zero mean and are *independently* distributed. Thus, Friedman's permanent income hypothesis appears to assume only a distributed lag in income.

If Friedman's hypothesis is to provide a useful tool for the analysis of the demand for individual commodities, the distribution of lag should be the same for each individual commodity and for total consumption, or, at least, for similar commodities or groups of commodities. 66/ If we assume the

---

66/ It is interesting to note that this test of Friedman's hypothesis was suggested by Friend and Kravis (17, p. 547). Their work came to the attention of the author only after his results had been obtained.

distribution of lag to be generated by a model of expectation formation, such as (239) or (33), the distribution of lag can be summarized by the value of a single parameter, namely the coefficient or elasticity of expectations, . Thus should be the same for every individual commodity and for total consumption. In order to simplify the computations, the model represented by equation (33), rather than by (239), was used.

In addition to total consumption, the demand for all food and the demand for meat were investigated. Meat and all food are similar commodities as compared, say, to housing expenditures or clothing expenditures, and we would not, therefore, expect the "horizons" appropriate to these two commodities to differ greatly, although we might allow some difference between food and meat, on the one hand, and total consumption on the other.

Let $y(t)$ = log of observed income during period t

$y^*(t)$ = log of permanent or "expected normal" income

$C(t)$ = log of observed aggregate consumption

$q_f(t)$ = log of the consumption of all food

$q_m(t)$ = log of the consumption of meat

$p_f(t)$ = log of the price of food

$p_m(t)$ = log of the price of meat

For statistical purposes these variables, except for $y^*(t)$, were defined as follows:

$y(t)$: Per capita disposable personal income (Commerce definition) deflated by the BLS consumer price index (1947-49=100)

$C(t)$: Per capita personal consumption expenditures (Commerce definition) deflated by the CPI

$q_f(t)$: The Agricultural Marketing Service index of per capita civilian food consumption at retail (not expenditure)

$q_m(t)$: Total civilian meat consumption per capita, in pounds, excluding lard

$p_f(t)$: The Bureau of Labor Statistics index of food prices at retail deflated by the CPI

$p_m(t)$: An Agricultural Marketing Service index of the retail prices of all meat excluding lard, deflated by the CPI

The data on observed total consumption and income are not entirely appropriate for this test: the Commerce definition of consumption includes many items which are properly savings, and the Commerce definition of disposable income excludes many items which might properly be considered as income (for example, social security taxes). Friedman has constructed series on consumption and income more appropriate for work of this kind, but these series were, at the time of writing, unavailable to the author of this paper. The computational difficulty of constructing such series precluded their use in the simple test presented here.

The basic equations to be estimated are a consumption function, a demand function for all food, and a demand function for meat:

$$C(t) = a_{00} + a_{01}\, y^*(t) + u_0(t) \tag{243}$$

$$q_f(t) = a_{10} + a_{11}\, p_f(t) + a_{12}\, y^*(t) + u_1(t) \tag{244}$$

$$q_m(t) = a_{20} + a_{21}\, p_m(t) + a_{22}\, y^*(t) + u_2(t) \tag{245}$$

where $u_0(t)$, $u_1(t)$, and $u_2(t)$ are residual terms. As our expectational equation we have

$$y^*(t) - y^*(t-1) = \alpha\,[y(t) - y^*(t-1)] \tag{246}$$

where α is an elasticity of expectations (since $y^*(t)$ and $y(t)$ are expressed in logarithmic form).

The system (243)-(245), with expectational equation (246), may be reduced by the single or multiple equation methods. 67/ Performing the reduction, we have

$$\begin{aligned} C(t) = {} & a_{00}\,\alpha + a_{01}\,\alpha\, y(t) + (1-\alpha)\, C(t-1) + \\ & u_0(t) - (1-\alpha)\, u_0(t-1) \end{aligned} \tag{247}$$

$$\begin{aligned} q_f(t) = {} & a_{10}\,\alpha + a_{12}\,\alpha\, y(t) + (1-\alpha)\, q_f(t-1) + \\ & a_{11}\, p_f(t) - a_{11}\,(1-\alpha)\, p_f(t-1) + \\ & u_1(t) - (1-\alpha)\, u_1(t-1) \end{aligned} \tag{248}$$

$$\begin{aligned} q_m(t) = {} & a_{20}\,\alpha + a_{22}\,\alpha\, y(t) + (1-\alpha)\, q_m(t-1) + \\ & a_{21}\, p_m(t) - a_{21}\,(1-\alpha)\, p_m(t-1) + \\ & u_2(t) - (1-\alpha)\, u_2(t-1) \end{aligned} \tag{249}$$

67/ Since only one expectational variable enters each equation, the two methods amount to the same thing in this case.

The system (243)-(245) is a separable system and may be separated prior to reduction, that is, equation (243) may be used to eliminate $y^*(t)$ from (244) and (245). In this case reduction after separation is not necessary since the separated equations contain no further expectational variables. Separation leads to the following two equations, alternative to (242) and (249), respectively:

$$q_f(t) = (a_{10} - \frac{a_{12}a_{00}}{a_{01}}) + \frac{a_{12}}{a_{01}} C(t) + a_{11} p_f + u_1(t) - \frac{u_0(t)}{a_{01}} \qquad (250)$$

$$q_m(t) = (a_{20} - \frac{a_{22}a_{00}}{a_{01}}) + \frac{a_{22}}{a_{01}} C(t) + a_{21} p_m(t) + u_2(t) - \frac{u_0(t)}{a_{01}} \qquad (251)$$

On Friedman's permanent income hypothesis, the coefficient a_{01} should be one, since all the variables are expressed in logarithms.

In an earlier section we suggested that, as a general rule, a system of equations should not be separated prior to estimation because an element would then be introduced which would tend to bias the estimates of certain coefficients in the estimated equation. This principal is illustrated by (250) and (251): Since $C(t)$ is positively correlated with $u_0(t)$ by equation (243), it tends to be negatively correlated with $u_1(t) - u_0(t)$ or $u_2(t) - u_0(t)$, the residuals in (250) and (251), respectively. There are two reasons, however, why separation should not introduce serious correlation between $C(t)$ and the residuals of (250) and (251): (1) the correlation between total consumption expenditures and expected normal income is likely to be high, and (2) the transitory components of food and meat consumption are likely to be highly correlated with the transitory component of total consumption expenditures. Since

$$r^2_{C(t)[u_i(t)-u_0(t)]} = \frac{\text{var } u_0(t) - \text{cov } [u_0(t), u_i(t)]}{\text{var } [u_0(t) - u_i(t)] \text{ var } C_t} \qquad (252)$$

for i=1, 2, it follows that this correlation is low if var $u_0(t)$ is small and if cov $[u_0(t), u_i(t)]$ does not exceed var $u_0(t)$. Wold (35) shows that under these circumstances the bias in the least squares estimates is not great. Consequently, it is possible to justify estimating equations (250) and (251).

Regressions based on equations (247)-(251) are presented in table 1. In addition to the estimates of the coefficients in the regression, table 1 gives the square of the multiple correlation coefficient, the number of observations, the Durbin-Watson statistic, the estimated or assumed elasticity of expectations, and the estimated elasticity of total consumption or demand with respect to permanent (or expected normal) income. Each regression was run for two periods: the interwar period, 1920-41, and the interwar and postwar periods, 1920-41 and 1948-55.

Table 1.--Demand for aggregate consumption, food, and meat based on the permanent income hypothesis: Least squares regressions and related statistical data 1/

Item	Aggregate consumption, equation (247)		Food, based on equation--				Meat, based on equation--			
			(248)		(250)		(249)		(251)	
	1920-41	1920-41 and 1948-55	1920-41	1920-41 and 1948-55	1920-41	1920-41 and 1948-55	1920-41	1920-41 and 1948-55	1920-41	1920-41 and 1948-55
Regression coefficient for specified independent variable:										
y(t)	0.729	0.726	0.199	0.210	---	---	0.351	0.387	---	---
	(.056)	(.057)	(.030)	(.029)			(.074)	(.060)		
C(t)	---	---	---	---	0.291	0.296	---	---	0.422	0.554
					(.028)	(.016)			(.101)	(.063)
C(t-1)	.099	.209	---	---	---	---	---	---	---	---
	(.077)	(.064)								
$q_f(t-1)$	---	---	.182	.266	---	---	---	---	---	---
			(.128)	(.108)						
$q_m(t-1)$	---	---	---	---	---	---	.231	.432	---	---
							(.161)	(.117)		
$p_f(t)$	---	---	-.065	-.088	-.133	-.137	---	---	---	---
			(.058)	(.058)	(.045)	(.038)				
$p_m(t)$	---	---	---	---	---	---	-.470	-.485	-.475	-.458
							(.088)	(.076)	(.086)	(.072)
$p_f(t-1)$	---	---	-.094	-.052	---	---	---	---	---	---
			(.045)	(.045)						
$p_m(t-1)$	---	---	---	---	---	---	-.015	.093	---	---
							(.085)	(.067)		
Constant term	.315	.122	1.545	1.323	1.681	1.682	1.774	1.134	2.120	1.855
R^2	.97	.99	.92	.97	.85	.96	.74	.83	.62	.76
Number of observations	22	30	22	30	22	30	22	30	22	30
Durbin-Watson statistic, d	1.56	1.48	1.86	1.89	1.52	2/1.42	2.13	2.21	2/1.08	3/.92
Estimate of--										
α	.90	.79	.82	.73	4/.90	4/.79	.77	.57	4/.90	4/.79
	(.08)	(.06)	(.13)	(.11)			(.16)	(.12)		
The elasticity of the dependent variable with respect to permanent income	.81	.92	.24	.29	.29	.30	.46	.68	.42	.55

1/ Numbers in parentheses beneath the coefficients are their respective standard errors. 2/ Durbin-Watson test inconclusive at the 5 percent probability level. 3/ Significant positive serial correlation. 4/ Assumed.

The regressions based on equation (247) appear to indicate that the elasticity of total consumption expenditures with respect to income is not one as suggested by the permanent income hypothesis. If $a_{01} = 1$, as required by the permanent income hypothesis, the sum of the coefficients of $y(t)$ and $C(t-1)$, in the regressions based on (247), should equal one. Thus we can test the significance of the difference between the two relevant elasticities from one by testing the significance of the difference of the sum of the coefficients of $y(t)$ and $C(t-1)$ from one. A likelihood ratio test may easily be derived to test the null hypothesis that the sum of the coefficients of $y(t)$ and $C(t-1)$ is one. The likelihood ratio for the period 1920-41 is 43, and for the period 1920-41 and 1948-55 is 12. Since the value of Chi-square for one degree of freedom is 11 at the 0.001 probability level, we reject the null hypothesis and conclude that the estimated elasticities of total consumption with respect to permanent income do differ significantly from one. This result, however, should be interpreted with care. First, the use of more-or-less inappropriate data on consumption and income may have led to this result inconsistent with the permanent income hypothesis. Second, the fact that the estimated elasticities rise when a longer period is used is consistent with the permanent income hypothesis, and suggests that some transitory component of income may be affecting the regression. Third, although the Durbin-Watson statistic does not indicate the presence of positive serial correlation, it is sufficiently low to warrant caution. 68/

The most interesting row of table 1 is the second to last. There the estimates of the elasticities of expectations based on the various regressions are presented. Although the elasticities derived from the different equations differ depending on the length of the period, the differences are more marked as between commodities, especially between total consumption and food, on the one hand, and meat, on the other. In addition, when consumption, rather than income, is used in the regressions for food and meat, the multiple correlations are markedly lower for the interwar period, and the same is true for the interwar plus postwar period in the case of meat. The only significant or inconclusive Durbin-Watson tests are also found for these regressions.

Are these differences in the elasticities of expectations among commodities and total consumption significant? If they are, we have reason to doubt the adequacy and/or utility of Friedman's permanent income hypothesis as applied to individual categories of consumption. An F-test, based on a test described in Meinken (25, pp. 100-102), was used to test the significance of the differences between the coefficients of $C(t-1)$, $q_f(t-1)$, and $q_m(t-1)$ in the regressions based on equations (247) to (249). The procedure followed is described in full in the Appendix. A significant F-ratio indicates a significant difference among the coefficients, and, hence, among the elasticities of expectations for total consumption, all food, and meat. The F-ratio for the regressions using data for the period 1920-41 is 20, with 2 and 53 degrees of

---

68/ If positive serial correlation in the residuals of (247) is present, the estimates of the coefficients are biased, since $C(t-1)$ enters as an independent variable.

freedom; the F-ratio for the combined periods 1920-41 and 1948-55 is 261, with 2 and 77 degrees of freedom. Each ratio is highly significant.

We cannot, of course, definitely conclude on the basis of the simple and, perhaps, crude test that the permanent income hypothesis is false; all we can say is that it does not appear useful when applied to individual categories of consumption.

Second test.--A simpler, alternative hypothesis is suggested by considerations discussed in this paper with reference to rigidities in consumer behavior due to technological or institutional factors. If these factors are the sole causes of rigidities, we have distributed lags in both prices and income, but within any equation the distributions of lag is the same for both price and income. In place of equation (246), we have three equations of the same form as equation (23) (see page 18):

$$C(t) - C(t-1) = \delta_c [C^*(t) - C(t-1)] \tag{253}$$

$$q_f(t) - q_f(t-1) = \delta_f [q_f^*(t) - q_f(t-1)] \tag{254}$$

$$q_m(t) - q_m(t-1) = \delta_m [q_m^*(t) - q_m(t-1)] \tag{255}$$

where $C^*(t)$, $q_f^*(t)$, and $q_m^*(t)$ are the long-run equilibrium values of total consumption and the quantities demanded, and δ_c, δ_f, and δ_m are the parameters determining the distributions of lag. Our basic demand equations are:

$$C^*(t) = a_{00} + a_{01}\, y(t) + u_0(t) \tag{256}$$

$$q_f^*(t) = a_{10} + a_{11}\, p_f(t) + a_{12}\, y(t) + u_1(t) \tag{257}$$

$$q_m^*(t) = a_{20} + a_{21}\, p_m(t) + a_{22}\, y(t) + u_2(t) \tag{258}$$

When equations (256)-(258) are reduced we have:

$$C(t) = a_{00}\, \delta_c + a_{01}\, \delta_c\, y(t) + (1-\delta_c)\, C(t-1) + \delta_c\, u_0(t) \tag{259}$$

$$q_f(t) = a_{10}\, \delta_f + a_{11}\, \delta_f\, p_f(t) + a_{12}\, \delta_f\, y(t) + (1-\delta_f)\, q_f(t-1) + \delta_f\, u_1(t) \tag{260}$$

$$q_m(t) = a_{20}\ \delta_m + a_{21}\ \delta_m\ p_m(t) + a_{22}\ \delta_m\ y(t) + (1-\delta_m)\ q_m(t-1) + \delta_m\ u_2(t) \tag{261}$$

Equation (253) suggests a regression of exactly the same form as is suggested by (247), but equations (260) and (261) suggest somewhat different ones than those under the permanent income hypothesis. Comparing (260) with (248) and (261) with (249) we see that $p_f(t-1)$ does not enter a regression based on (260) and $p_m(t-1)$ does not enter a regression based on (261). The fact that the coefficients of these variables in the regressions presented in table 1 are not significantly different from zero or are of the wrong sign suggests that the alternative hypothesis may have some merit. The coefficients of $C(t-1)$, $q_f(t-1)$, and $q_m(t-1)$, of course, need not be the same. Regressions based on equations (259)-(261) are presented in table 2 for the period 1920-41 and the combined periods 1920-41 and 1948-55. As can be seen, the results compare favorably with those presented in table 1.

It is perhaps too early to conclude that technological and institutional rigidities play a greater role in demand analysis than rigidities of an expectational nature. Nonetheless the results presented in this section suggest that this may be the case. If so, it would simplify matters greatly. In order, however, to reach a firm conclusion, further research is necessary. It is hoped that this paper makes a modest contribution to that research.

LITERATURE CITED

(1) Aitken, A. C.
1935. On Least Squares and Linear Combinations of Observations. Proc. Royal Soc. Edinburgh. 55:42-48.

(2) Alt, F. F.
1942. Distributed Lags. Econometrica. 10:113-128.

(3) Arrow, Kenneth J., and Nerlove, Marc
1957. A Note on Expectations and Stability. Tech. Rpt. 41, Dept. Econ., Stanford Univ. 14 pp. [Processed.]

(4) Berger, J.
1953. On Koyck's and Fisher's Methods for Calculating Distributed Lags. Metroeconomica. 5:89-90.

(5) Cagan, Phillip
1956. The Monetary Dynamics of Hyper-Inflations. In Friedman, Milton, ed., Studies in the Quantity Theory of Money. 260 pp., illus. Chicago.

(6) Chernoff, Herman, and Divinsky, Nathan
1953. The Computation of Maximum-Likelihood Estimates of Linear Structural Equations. In Hood, Wm. C., and Koopmans, Tjalling C., ed., Studies in Econometric Method, Cowles Commission for Research in Economics Monogr. 14, 324 pp., illus. New York.

Table 2.--Demand for aggregate consumption, food, and meat based on an alternative to the permanent income hypothesis: Least squares regressions and related statistical data 1/

Item	Aggregate consumption, equation (259)		Food, equation (260)		Meat, equation (261)	
	1920-41	1920-41 and 1948-55	1920-41	1920-41 and 1948-55	1920-41	1920-41 and 1948-55
Regression coefficient for specified independent variable:						
$y(t)$	0.729 (.056)	0.726 (.057)	0.217 (.031)	0.219 (.028)	0.350 (.071)	0.421 (.056)
$C(t-1)$	.099 (.077)	.209 (.064)	---	---	---	---
$q_f(t-1)$	---	---	.151 (.139)	.237 (.106)	---	---
$q_m(t-1)$	---	---	---	---	.248 (.127)	.371 (.111)
$p_f(t)$	---	---	-.155 (.042)	-.142 (.036)	---	---
$p_m(t)$	---	---	---	---	-.478 (.076)	-.429 (.065)
Constant term	.315	.122	1.563	1.366	1.730	1.263
R^2	.97	.99	.89	.97	.73	.82
Number of observations	22	30	22	30	22	30
Durbin-Watson statistic, d	1.56	1.48	1.91	1.89	2.16	2.04
Estimate of--						
δ	.901	.791	.849	.763	.752	.629
Long-run elasticity with respect to--						
Income	.81	.92	.26	.29	.47	.67
Price	---	---	-.183	-.186	-.636	-.682

1/ Numbers in parentheses beneath the coefficients are their respective standard errors.

(7) Chow, G. C.
1957. Demand for Automobiles in the United States: A Case Study in Consumer Durables. Amsterdam. [In press.]

(8) Cochrane, D., and Orcutt, G. H.
1949. Application of Least Squares Regression to Relationships Containing Auto-correlated Error Terms. Jour. Amer. Statis. Asso. 44:32-61.

(9) Cramér, Harald
1946. Mathematical Methods of Statistics. 575 pp., illus. Princeton, N. J.

(10) Durbin, J., and Watson, G. S.
1950-51. Testing for Serial Correlation in Least Squares Regressions, I and II. Biometrika. 37:409-428 and 38:159-178.

(11) Enthoven, Alain C., and Arrow, K. J.
1956. A Theorem on Expectations and the Stability of Equilibrium. Econometrica. 24:288-293.

(12) Farrell, M. J.
1954. The Demand for Motor-Cars in the United States. Jour. Royal Statis. Soc., Ser. A. 117:171-201, illus.

(13) Fisher, Irving
1925. Our Unstable Dollar and the So-called Business Cycle. Jour. Amer. Statis. Asso. 20:179-202, illus.

(14) ______
1930. The Theory of Interest as Determined by Impatience to Spend Income and Opportunity to Invest It. 566 pp., illus. New York.

(15) ______
1937. Note on a Short-cut Method for Calculating Distributed Lags. Internatl. Statis. Inst. Bull. 29:323-327.

(16) Friedman, Milton
1957. A Theory of the Consumption Function. National Bureau of Economic Research. 243 pp., illus. Princeton, N. J.

(17) Friend, Irwin, and Kravis, Irving B.
1957. Consumption Patterns and Permanent Income. Amer. Econ. Rev. May issue (Proc.), 47:536-555.

(18) Hicks, J. R.
1946. Value and Capital. 2d ed. 340 pp., illus. Oxford, England.

(19) Hurwicz, Leonid
1950. Least-squares Bias in Time-series. In Koopmans, Tjalling C., ed., Statistical Inference in Dynamic Economic Models, Cowles Commission for Research in Economics Monogr. 10, 438 pp. New York.

(20) Klein, Lawrence R.
1953. A Textbook of Econometrics. 355 pp., illus. Evanston, Ill.

(21) Koopmans, Tjalling C., and Hood, William C.
1953. The Estimation of Simultaneous Linear Economic Relationships. In Hood, William C., and Koopmans, Tjalling C., ed., Studies in Econometric Method, Cowles Commission for Research in Economics Monogr. 14, 324 pp., illus. New York.

(22) Koyck, L. M.
1954. Distributed Lags and Investment Analysis. 111 pp., illus. Amsterdam.

(23) Mann, H. B., and Wald, A.
1943. On the Statistical Treatment of Linear Stochastic Difference Equations. Econometrica. 11:173-220.

(24) Marshall, Alfred
1936. Principles of Economics, 8th ed. 871 pp., illus. London.

(25) Meinken, Kenneth W.
1953. The Demand and Price Structure for Oats, Barley, and Sorghum Grains. U. S. Dept. Agr. Tech. Bull. 1080. 102 pp., illus.

(26) Nerlove, Marc
1956. Estimates of the Elasticities of Supply of Selected Agricultural Commodities. Jour. Farm Econ. 38:496-509.

(27) ______
1957. A Note on Long-run Automobile Demand. Jour. Mkt. 22:57-64, illus.

(28) Phillips, A. W.
1956. Some Notes on the Estimation of Time-forms of Reactions in Interdependent Dynamic Systems. Economica. 23:99-113.

(29) Roos, C. F., and von Szeliski, Victor
1939. Factors Governing Changes in Automobile Demand. In General Motors Corporation, Dynamics of Automobile Demand, 139 pp., illus. New York.

(30) Stone, Richard
1954. The Measurement of Consumers' Expenditures and Behavior in the United Kingdom, 1920-1938. 448 pp., illus. Cambridge, England.

(31) Tinbergen, J.
1949. Long-term Foreign Trade Elasticities. Metroeconomica. 1:174-185.

(32) ______
1951. Econometrics. Transl. from the Dutch by H. Ryken Van Olst. 258 pp., illus. Philadelphia, Pa.

(33) Tobin, James A.
1950. A Statistical Demand Function for Food in the U.S.A. Jour. Royal Statis. Soc., Ser. A. 113:113-140, illus.

(34) Wilks, S. S.
1950. Mathematical Statistics. 284 pp., illus. Princeton, N. J. [Processed.]

(35) Wold, Herman, and Faxer, P.
1957. On the Specification of Error in Regression Analysis. Annals Math. Statis. 28:265-267.

(36) ______ and Jureen, Lars
1953. Demand Analysis: A Study in Econometrics. 358 pp., illus. New York.

APPENDIX

A Test of Significance for the Equality of the Elasticities of Expectations Obtained for Several Commodities

Equations (245)-(249) may be written

$$C(t) = \pi_{00} + \pi_{01}\, y(t) + \pi_{02}\, C(t-1) + v_0(t) \tag{262}$$

$$q_f(t) = \pi_{10} + \pi_{11}\, y(t) + \pi_{12}\, p_f(t) + \pi_{13}\, p_f(t-1) + \pi_{14}\, q_f(t-1) + v_1(t) \tag{263}$$

$$q_m(t) = \pi_{20} + \pi_{21}\, y(t) + \pi_{22}\, p_m(t) + \pi_{23}\, p_m(t-1) + \pi_{24}\, q_m(t-1) + v_2(t) \tag{264}$$

We wish to test whether

$$\pi_{02} = \pi_{14} = \pi_{24} = \pi_{..} \tag{265}$$

where $\pi_{..}$ is the common value of the coefficients of $C(t-1)$, $q_f(t-1)$, and $q_m(t-1)$.

The test may be carried out in the following steps:

<u>Step 1</u>: Compute the adjusted sums of squares of the dependent variables

$$M_{C(t)C(t)} - \hat{\pi}_{01}\, M_{y(t)C(t)} \tag{266}$$

$$M_{q_f(t)q_f(t)} - \hat{\pi}_{11}\, M_{y(t)q_f(t)} - \hat{\pi}_{12}\, M_{p_f(t)q_f(t)} - \hat{\pi}_{13}\, M_{p_f(t-1)q_f(t)} \tag{267}$$

$$M_{q_m(t)q_m(t)} - \hat{\pi}_{21}\, M_{y(t)q_m(t)} - \hat{\pi}_{22}\, M_{p_m(t)q_m(t)} - \hat{\pi}_{23}\, M_{p_m(t-1)q_m(t)} \tag{268}$$

Add these together and call the sum M^*_{11}.

<u>Step 2</u>: Compute the adjusted moments of $C(t)$ or $C(t-1)$; $q_f(t)$ on $q_f(t-1)$; $q_m(t)$ on $q_m(t-1)$ <u>after the effects of the other variables have been removed</u>. That is, compute

$$M_{C(t)C(t-1)} - \hat{\pi}_{01}\, M_{y(t)C(t-1)} \tag{269}$$

$$M_{q_f(t)q_f(t-1)} - \hat{\pi}_{11}\, M_{y(t)q_f(t-1)} - \hat{\pi}_{12}\, M_{p_f(t)q_f(t-1)} - \hat{\pi}_{13}\, M_{p_f(t-1)q_f(t-1)} \tag{270}$$

$$M_{q_m(t)q_m(t-1)} - \hat{\pi}_{21} M_{y(t)q_m(t-1)} - \hat{\pi}_{22} M_{p_m(t)q_m(t-1)}$$
$$- \hat{\pi}_{23} M_{p_m(t-1)q_m(t-1)} \tag{271}$$

Add these together and call the sum M^*_{12}.

Step 3: Add the moments $M_{C(t-1)C(t-1)}$, $M_{q_f(t-1)q_f(t-1)}$, and $M_{q_m(t-1)q_m(t-1)}$ together and call the sum M^*_{22}.

Step 4: Compute the residual sum of squares for the combined simple regression:

$$S_1 = M^*_{11} - \frac{(M^*_{12})^2}{M^*_{22}} \tag{272}$$

This residual sum of squares has N-10-1 degrees of freedom, where N is the total number of observations (that is, the sum of the numbers for each regression). The number of degrees of freedom lost in computing the adjusted moments is 10, and in computing the combined regression is 1. The common value we suppose π_{02}, π_{14}, and π_{24} to have is

$$\pi_{..} = \frac{M^*_{12}}{M^*_{22}} \tag{273}$$

Step 5: Compute the residual sums of squares for each of the three regressions

$$S_2 = M_{C(t)C(t)} (1-R_1^2) +$$
$$M_{q_f(t)q_f(t)} (1-R_2^2) +$$
$$M_{q_m(t)q_m(t)} (1-R_3^2) \tag{274}$$

where R_1^2, R_2^2, and R_3^2 are the squares of the respective coefficients of multiple correlation. S_2 has N-13 degrees of freedom, where the 13 is obtained by counting one for each coefficient and constant term in the three regressions.

Step 6: Compute $S_3 = S_1 - S_2$. S_3 has N-11-(N-13) = 2 degrees of freedom.

Step 7: Compute

$$F' = \frac{S_3/2}{S_2/N-13}$$

F' is distributed as F with 2 and N-13 degrees of freedom. If $F' \leq F_{.01}(2, N-13)$ we accept the hypothesis that

$$\pi_{02} = \pi_{14} = \pi_{24} = \pi_{..}$$

56

☆ U. S. GOVERNMENT PRINTING OFFICE : 1958 O