

Estimation of the Linear Expenditure System

Author(s): Robert A. Pollak and Terence J. Wales

Source: *Econometrica*, Oct., 1969, Vol. 37, No. 4 (Oct., 1969), pp. 611-628

Published by: The Econometric Society

Stable URL: <https://www.jstor.org/stable/1910438>

REFERENCES

Linked references are available on JSTOR for this article:

https://www.jstor.org/stable/1910438?seq=1&cid=pdf-reference#references_tab_contents

You may need to log in to JSTOR to access the linked references.

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <https://about.jstor.org/terms>



The Econometric Society is collaborating with JSTOR to digitize, preserve and extend access to *Econometrica*

JSTOR

ESTIMATION OF THE LINEAR EXPENDITURE SYSTEM¹

BY ROBERT A. POLLAK AND TERENCE J. WALES

In this paper we estimate a complete system of demand equations making full use of the restrictions implied by economic theory. Our theoretical model is based on the Klein-Rubin linear expenditure system which was first estimated by Stone. We place primary emphasis on maximum likelihood estimates obtained using annual time series observations of prices and per capita consumption for the U.S. economy in the period 1948–1965. The plan of the paper is as follows: Section 1 begins with a discussion of the problems involved in making systematic use of economic theory to estimate demand functions; this is followed by a brief description of the linear expenditure system and discussion of the specification of its dynamic and stochastic structure. In Section 2 we describe three methods of estimating the linear expenditure system, including the maximum likelihood procedure which we believe is most appropriate. We report our results in Section 3 and our conclusions in Section 4.

1A. INTRODUCTION

The pure theory of consumer behavior is concerned with individual demand functions. An individual's preferences are assumed to be representable by a well behaved utility function, $U(x_1, \dots, x_n)$, where x_i denotes the rate of consumption of the i th good. He is supposed to maximize U subject to the budget constraint

$$(1) \quad \sum_{k=1}^n p_k x_k = \mu$$

where p_i is the price of the i th good and μ denotes total expenditure. The utility maximizing quantities of the various goods are functions of all prices and total expenditure; we write $x_i = h^i(P, \mu)$ where P denotes the price vector, (p_1, \dots, p_n) and the functions (h^1, \dots, h^n) are the ordinary demand functions. These demand functions satisfy the budget constraint and are homogeneous of degree zero in all prices and total expenditure; in addition, the implied Slutsky substitution matrix is symmetric and negative semidefinite. Furthermore, since any set of demand functions that satisfies these conditions is derivable from a well behaved utility function, we call such a set a "complete system of theoretically plausible demand functions."

The data we use in this study are annual observations on prices and per capita consumption, so we are concerned with "market" or "aggregate" demand functions. Unfortunately, a complete system of market demand functions need not be theoretically plausible even if every individual's demand functions are. Nevertheless, as a matter of research strategy, we shall assume that market demand functions are theoretically plausible, since this assumption reduces substantially the number

¹ This research was supported in part by a grant from the National Science Foundation, GS-1462. We are grateful to J. R. Behrman, D. S. Brady, M. D. McCarthy, E. C. Prescott, R. Summers, and the anonymous referees for helpful comments.

of parameters to be estimated.² Our procedure is to begin with an “aggregate utility function,” derive the corresponding aggregate demand functions, and estimate them.

We have chosen to work with four broad categories of goods: food, clothing, shelter, and miscellaneous. Our reasons are twofold: first, the utility function underlying our demand functions is additive, implying a type of “independence” among “goods.” On *a priori* grounds, additivity is a more reasonable working hypothesis when applied to broad aggregates than to more narrowly defined commodity groups. Second, because we are estimating an interrelated system of equations simultaneously, the computer requirements are substantial and economizing on computer time is a nontrivial consideration.³ In the absence of this constraint, we would have experimented with disaggregating the miscellaneous category somewhat further—perhaps treating medical care as a separate category—but (in any event) the scope for further disaggregation is not great.

In its simplest form, demand theory is concerned with the allocation of total expenditure among goods in a single period. In order to discuss either “saving” or purchases of durable goods in terms of utility maximization, it is necessary to formulate an explicitly intertemporal model, recognizing the role of future consumption, assets, future prices, and future income. Instead of proceeding in this direction, we shall work with a single period model and analyze the allocation of total expenditure on nondurables among broad categories of nondurables. Avoiding the problems of saving and purchases of consumer durables in this way requires strong assumptions about tastes, for the stock of durables may influence the pattern of consumption of nondurables (e.g., the demand for electricity is related to the stock of electrical appliances) and intertemporal complementarity may make future consumption plans relevant to today’s consumption decisions. Technically, our approach is justified if and only if the marginal rates of substitution involving current consumption of nondurables are all independent of current consumption of the services of durables and of all future consumption.⁴

1B. FUNCTIONAL FORM

In this section we discuss the static, nonstochastic form of the demand functions we intend to estimate.

We begin with a utility function of the form

$$(2) \quad U(X) = \sum_{k=1}^n a_k \log(x_k - b_k), \quad a_i > 0, \quad \sum_1^n a_k = 1, \quad x_i - b_i > 0.$$

² It might seem sensible to estimate aggregate demand functions without imposing the requirement of theoretical plausibility and test whether the estimated set of demand functions is theoretically plausible. The difficulty with this approach is that there is no way to distinguish between misspecification of the functional form of the aggregate demand functions and failure of the aggregate demand functions to satisfy the Slutsky symmetry conditions. We remind the reader that the Slutsky symmetry conditions are restrictions on the partial derivatives of the demand functions, and not on finite first differences. The behavior of estimates of these partial derivatives may be very sensitive to the *a priori* specification of the functional form.

³ We discuss this problem briefly in the appendix.

⁴ Since the demand for “transportation services” and “gasoline and oil” are closely related to the stock of automobiles, we have excluded them from our analysis.

Maximizing (2) subject to the budget constraint (1) yields demand functions of the form

$$(3) \quad h^i(P, \mu) = b_i - \frac{a_i}{p_i} \sum_{k=1}^n p_k b_k + \frac{a_i}{p_i} \mu \quad (i = 1, \dots, n).$$

Any system of demand functions implies a corresponding system of expenditure functions, $(e^1(P, \mu), \dots, e^n(P, \mu))$ defined by $e^i(P, \mu) = p_i h^i(P, \mu)$. The expenditure functions corresponding to (3) can be written as

$$(4) \quad e^i(P, \mu) = p_i b_i + a_i \left[\mu - \sum_{k=1}^n p_k b_k \right] \quad (i = 1, \dots, n).^5$$

If the b 's are all positive and income is greater than $\sum p_k b_k$, we may describe the individual as purchasing necessary quantities of the various goods (b_1, \dots, b_n) and then dividing his remaining or "supernumerary" income $(\mu - \sum p_k b_k)$ among the goods in fixed proportions (a_1, \dots, a_n) .⁶ Although it is theoretically possible for some or all of the b 's to be negative, it is unlikely in the present context. If b_i is negative, the demand for the i th good is elastic with respect to its own price; this seems improbable for any of the broad categories of goods with which we are dealing. Positive b 's imply inelastic demand.

1C. DYNAMIC SPECIFICATION

To use observations from different time periods to estimate demand functions one must either assume that the demand functions are the same in all periods or make fairly specific assumptions about how they change. From a technical standpoint, it is relatively simple to incorporate changing b 's into the linear expenditure system, because the b 's enter the demand functions linearly.⁷ And if one takes the necessary basket interpretation seriously, it seems plausible that the b 's should vary over time.

We write the demand functions (3) with varying b 's as

$$(5) \quad x_{it} = b_{it} - \frac{a_i}{p_{it}} \sum_{k=1}^n p_{kt} b_{kt} + \frac{a_i}{p_{it}} \mu_t,$$

where b_{it} is the necessary quantity of good i in period t . The demand functions of

⁵ The demand functions (3) were introduced by L. R. Klein and H. Rubin, "A Constant-Utility Index of the Cost of Living," *Review of Economic Studies*, XV (2), No. 38 (1947-48), pp. 84-87, and first estimated by Richard Stone, "Linear Expenditure Systems and Demand Analysis; An Application to the Pattern of British Demand," *Economic Journal*, Vol. LXIV, 255 (Sept., 1954), pp. 511-527. Stone has pioneered in empirical applications of the linear expenditure system in a series of papers. See, for example, "Demand Analysis and Projections for Britain, 1900-1970; a Study in Method," with A. Brown and D. A. Rowe, in J. Sandee, ed., *Europe's Future Consumption*, Vol. 2, Amsterdam, North-Holland, 1964, pp. 200-225.

⁶ For stylistic reasons, we shall hereafter refer to μ as "income."

⁷ Stone has estimated the linear expenditure system under assumptions which permit the a 's and b 's to vary; he has made use of the linear time trend and has suggested using the lagged consumption dynamic specifications. See Stone, "Linear Expenditure Systems . . .," op. cit., p. 522, and "Demand Analysis . . .," op. cit., p. 205.

period t are generated by a utility function of the form

$$(6) \quad U^t(X_t) = \sum_{k=1}^n a_k \log(x_{kt} - b_{kt}), \quad a_i > 0, \quad x_{it} - b_{it} > 0, \quad \sum a_k = 1.$$

The regularity conditions are identical to those of (2) and they must be satisfied in every period. It is easy to verify that, *ceteris paribus*, a higher level of b_{it} implies a higher level of x_{it} (and a lower level of x_{jt} , $j \neq i$).

The simplest way to permit the b 's to vary is to assume that b_{it} is a linear function of time:

$$(7) \quad b_{it} = b_i^* + \beta_i t.$$

Although we do estimate the linear expenditure system with this dynamic specification, the use of a time trend is not very satisfactory because it gives so little insight into the structure of the economic system. Furthermore, it implies that taste change would continue unabated (i.e., the necessary quantities would continue to increase) even if prices and income remained constant over a long period of time.

A more satisfactory dynamic specification—one which attempts to deal directly with the mechanism underlying changes in tastes—is based on the concept of “habit formation.” Habit formation can be incorporated into the model by allowing the b 's to depend on past consumption. The simplest habit model is based on the assumption that b_{it} is a linear function of consumption of the i th good in period $t - 1$,

$$(8a) \quad b_{it} = b_i^* + \beta_i x_{it-1}.$$

The constant b model is a special case of (8a) with the β 's all equal to zero. Another special case of (8a) which we consider is the proportional habit model

$$(8b) \quad b_{it} = \beta_i x_{it-1}.^8$$

In general, the linear habit formation model can be written as

$$(9) \quad b_{it} = b_i^* + \beta_i z_{it-1}$$

where z_{it-1} is a variable representing consumption of the i th good prior to period t . In addition to (8), we consider two specifications in which z_{it-1} depends on the level of past consumption and two in which it depends on the rate of growth of consumption. The two based on the level of consumption take z_{it-1} to be: (a) the highest level of consumption of the i th good during the three years prior to period t , and, (b) the average level of consumption of the i th good during the three years prior to period t .

The two models based on the rate of growth of consumption, g_{it} ,

$$(10) \quad g_{it} = \frac{x_{it}}{x_{it-1}} - 1,$$

⁸ Existence and stability of the long run equilibrium in these models have been analyzed by Robert A. Pollak, “Habit Formation and Dynamic Demand Functions,” Discussion Paper Number 79, Department of Economics, University of Pennsylvania, March, 1968, mimeographed.

take z_{it-1} to be (c) the sum of the g_{it} 's up to period $t - 1$ (so $\Delta b_{it} = \beta_i g_{it-1}$), and, (d) the sum of $(g_{it} + 1)$ up to period $t - 1$ (so $\Delta b_{it} = \beta_i + \beta_i g_{it-1}$).⁹

1D. STOCHASTIC SPECIFICATION

Although the linear expenditure system has been estimated by a number of investigators, little systematic attention has been given to the problems of specifying an appropriate error structure.¹⁰ We assume throughout that the disturbance terms enter the demand functions additively. That is,

$$(11) \quad x_{it} = b_{it} - \frac{a_i}{p_{it}} \sum_{k=1}^n p_{kt} b_{kt} + \frac{a_i}{p_{it}} \mu_t + v_{it}$$

where v_{it} is a random variable. It is sometimes more convenient to work with the expenditure functions

$$(12) \quad p_{it} x_{it} = p_{it} b_{it} - a_i \sum_{k=1}^n p_{kt} b_{kt} + a_i \mu_t + w_{it},$$

where $w_{it} = p_{it} v_{it}$ denotes the disturbance term of the i th expenditure equation.

It would be convenient to assume that the v 's (or the w 's) are mutually independent; but this assumption is inconsistent with the budget constraint which requires that

$$(13) \quad \sum_{k=1}^n p_{kt} v_{kt} = \sum_{k=1}^n w_{kt} = 0 (t = 1, \dots, T).$$

To satisfy (13), the covariance matrix of disturbance terms for each period must be singular. We shall discuss three general ways of specifying the distributions of the v 's (w 's) which satisfy these requirements.

If we assume that the w 's in the first $n - 1$ expenditure equations are mutually independent, then the disturbance term in the last equation is the negative of the sum of the first $n - 1$ w 's. Expenditure on the n th good is a residual in the sense that it adjusts passively to the independent disturbances in the other equations so as to satisfy the budget constraint. We have two objections to this stochastic specification. First, we are not happy with the assumption that the disturbances in $n - 1$ of the expenditure equations are mutually independent. We believe that the disturbance terms in the various expenditure equations are interrelated, and that a random shock which causes an increase in the consumption of one good is likely to affect consumption of many other goods. Second, this specification of the error

⁹ Specification (c) implies that if past consumption has been constant, then b_{it} will remain constant; (d) implies that if past consumption has been constant, then b_{it} will grow in the manner described by the linear time trend (7). We were led to experiment with (d) because of plausible results obtained using (7).

¹⁰ With the exception of the work of Theil and Barten, very little attention has been given to the problem of specifying error structures for empirical demand analysis. See H. Theil, *Economics and Information Theory*, North-Holland Publishing Company, Amsterdam, 1967, especially pp. 227-233, and A. P. Barten, "Estimating Demand Equations," *Econometrica*, Vol. 36, No. 2, April, 1968, pp. 213-251.

structure forces us to decide which good is to play the role of residual, but it offers no criterion for making the decision.

If we assume that the covariance matrix of the w 's is the same in each period, then its form need not be specified.¹¹ The elements of the covariance matrix can be estimated together with the other parameters.¹² This procedure has two obvious drawbacks. It is expensive in terms of degrees of freedom, and the assumption that the covariance matrix is constant is not an appealing one. More specifically, it seems to us that as per capita consumption and expenditure increase over time, the variance of the expenditure equations should also increase. Furthermore, the specification of a constant variance-covariance matrix implies that if all prices and income were to increase proportionally, then the variance of each expenditure equation would remain unchanged, while the variances of the demand equation disturbances would decrease. On theoretical grounds, we prefer a specification of the error structure in which the covariance matrix of the demand equation disturbances (the v 's) is unaffected by proportional changes in all prices and income. One way of specifying such an error structure is to assume that the covariance matrix of the w 's in each period is equal to a constant matrix (the same in each period) multiplied by the square of income.

We now present a third method of specifying the structure of the disturbances which we feel is superior to the two methods described above. Formally, our method is based on replacing b_{it} in each demand equation by $b_{it} + u_{it}$, where u_{it} is a random variable. The implied stochastic demand functions are of the form (11) where v_{it} is given by

$$(14) \quad v_{it} = u_{it} - \frac{a_i}{p_{it}} \sum_{k=1}^n p_{kt} u_{kt}.$$

We can rationalize these stochastic demand functions by postulating a stochastic utility function of the form

$$(15) \quad U^i(X_t) = \sum_{k=1}^n a_k \log(x_{kt} - b_{kt} - u_{kt}), \quad a_i > 0, \quad (x_{it} - b_{it} - u_{it}) > 0, \\ \sum a_k = 1.$$

The u 's can be interpreted as random variations in the necessary basket, but it is neither necessary nor especially useful to interpret this error structure in terms of a

¹¹ The covariance matrix of the v 's must depend on prices for (13) to hold; hence, the covariance matrix of the v 's cannot be the same in each period. Formally, the covariance matrix Ω_t is given by $\Omega_t = E(v_t v_t')$. Since $p_t' v_t = 0$, we have $p_t' \Omega_t = 0$. That is, p_t is an eigenvector of Ω_t corresponding to eigenvalue 0. If Ω is constant and if there are n linearly independent price vectors in the set $\{p_1, \dots, p_T\}$, then the covariance matrix is of rank zero. This strongly suggests that the covariance matrix is not constant.

¹² This is the procedure followed by R. W. Parks for the constant b model; Richard W. Parks, "Systems of Demand Equations: An Empirical Comparison of Alternative Functional Forms," Center for Mathematical Studies in Business and Economics, University of Chicago, Report 6808, Feb., 1968. Also see: A. P. Barten "Maximum Likelihood Estimation of a Complete System of Demand Equations," Center for Operations Research and Econometrics, Université Catholique de Louvain, Discussion Paper No. 6709, August, 1967.

stochastic utility function. A specification of the error structure should be judged on its implications for the stochastic demand functions.¹³

We first observe that u_{it} is directly related to v_{it} ; a higher value of u_{it} implies, *ceteris paribus*, a higher value of v_{it} (and hence x_{it}) and a lower value of v_{jt} (and hence x_{jt}) for all $j \neq i$. Second, the adding-up condition (13) is automatically satisfied by the v 's implied by this error structure, regardless of the distribution of the u 's. Third, proportional changes in all prices and income (providing these changes do not affect the distribution of the u 's) will not affect the distribution of the v 's. Fourth, the Slutsky substitution matrix implied by these stochastic demand functions is symmetric and negative semidefinite regardless of the values assumed by the u 's. Fifth, this method of specifying the error structure treats all goods in a symmetric manner.

The simplest assumptions about the distribution of the u 's are:

$$(16) \quad E(u_{it}) = 0,$$

$$(17) \quad E(u_{it}^2) = \sigma_i^2,$$

$$(18) \quad E(u_{it}u_{jt}) = 0, \quad i \neq j,$$

$$(19) \quad E(u_{it}u_{jt+\tau}) = 0, \quad \tau \neq 0.$$

That is, the expected value of each u_{it} is zero, its variance is constant over time, and the u 's are independent across goods and time periods. Finally, we assume that the u 's have a multivariate normal distribution.

The implications for the distribution of the v 's of our assumptions about the distributions of the u 's are easily derived. Since the v 's are linear combinations of the u 's, they also have a multivariate normal distribution.¹⁴ The covariance matrix of the v 's depends on prices, so it is not constant over time. The variance of the disturbance term in the i th demand equation is independent of income and it is not directly related to consumption of the i th good, although it is inversely related to the price of the i th good. Finally, v 's from different periods are mutually independent. The *a priori* plausibility of these last two properties deserves critical scrutiny.

It seems unlikely that the variance of the disturbance for the i th good would be independent of income. If prices remain constant and income increases, causing an increase in consumption of each good, then the variance of the disturbance terms in each demand equation will probably also increase. One way to incorporate this *a priori* belief into our specification of the error structure is to replace assumption (17)

¹³ In its most general form, our approach to the specification of the error structure is the following. Consider a nonstochastic utility function, $U(X, a, b)$, where a and b are vectors of unknown parameters to be estimated; let $H(P, \mu, a, b)$ denote the demand functions corresponding to U (in vector form). We postulate that a subset of the parameters are random variables (i.e., $b = b^* + u$, where u is a vector of random variables with zero mean) and estimate the stochastic demand functions $H(P, \mu, a, b + u)$.

The Theil-Barten approach, which they describe as the "marginal utility shock model" appears to us to be equivalent to the following procedure. Again starting with a nonstochastic utility function, $U(X, a, b)$, define a new stochastic utility function, $V(X, a, b, u)$, by $V(X, a, b, u) = U(X, a, b) + \sum_{k=1}^n u_k x_k$; then derive the stochastic functions corresponding to V and estimate them.

¹⁴ For a discussion of singular linear transformations of a multivariate normal vector see T. W. Anderson, *Introduction to Multivariate Statistical Analysis*, New York, John Wiley and Sons, 1958.

by

$$(20) \quad E(u_{it}^2) = \sigma_i^2 \hat{x}_{it}^2$$

where \hat{x}_{it} is the expected value of x_{it} (i.e., the nonstochastic portion of (11)). We use \hat{x}_{it}^2 rather than x_{it}^2 in order to preserve the additivity of the error structure. A higher level of income will cause a higher level of \hat{x}_{it}^2 and hence increase the variance of u_{it} which, in turn, increases the variance of v_{it} .¹⁵

The assumption that the u 's are uncorrelated over time implies that the v 's are uncorrelated over time. We believe that, in the context of the habit models which depend on lagged consumption, (8a) and (8b), and also in the habit models (b), (c) and (d), this is a plausible assumption. One would expect autocorrelation of the v 's if a higher level of consumption of the i th good yesterday is associated with a higher level of consumption of the i th good today. But in the habit models which depend on lagged consumption, this relationship has already been taken into account. In all of these models, a higher level of v_{it-1} implies a higher level of x_{it-1} , which in turn implies a higher level of b_{it} and x_{it} . Thus, there is no reason to assume that the v 's are autocorrelated in the habit models which depend on consumption in the previous period. In the constant b model, the linear time trend, and the habit model which depends on previous peak consumption, we would expect the v 's to exhibit autocorrelation. A higher level of v_{it-1} does not imply a higher value of b_{it} in the first two cases, and need not in the third. Unfortunately it would be extremely complicated to estimate the linear expenditure system with an error structure incorporating autocorrelation, and we have not attempted to move in this direction. Consequently, parameter standard errors in these models should be viewed with extreme caution.

2. ESTIMATION PROCEDURES

The most straightforward procedure for obtaining estimates of the linear expenditure system and the one used by most investigators is to minimize the sum of squared residuals over all expenditure equations and time periods.¹⁶ It is appealing because of its simplicity, because it is a straightforward generalization of single equation ordinary least squares, and because it requires no *a priori* specification of the error structure. Its major drawback is that, because it does not rest on a specification of the error structure, the properties of the estimator are not known. It should be noted, however, that it is not a maximum likelihood procedure since a maximum likelihood interpretation requires a disturbance covariance matrix proportional to the identity, whereas in fact the covariance matrix of the system (although unknown) is singular.

An alternative least squares estimation procedure is to minimize the sum of squared expenditure residuals after omitting one equation. The major drawback of

¹⁵ This argument may be misleading. If the variance of the v 's is only slightly sensitive to changes in income or consumption, then the constant variance specification (17) may be more appropriate than (20).

¹⁶ For example, this is the technique used by Stone.

this method is that the estimates depend on which equation is omitted. That is, n sets of parameter estimates are obtained by estimating the system with a different equation omitted each time. This procedure yields maximum likelihood estimates if (i) the disturbances associated with all but the omitted expenditure equation are mutually independent and (ii) the variances associated with the disturbances in the retained equations are equal and constant over time. If these two conditions are satisfied, the covariance matrix of the retained equations is proportional to the identity matrix. As noted in Section 1 the difficulty with this specification of the error structure is that it is asymmetric, and there is no basis for deciding which good should play the role of residual.¹⁷

The stochastic specification which appears to us most appropriate and the one for which we obtain maximum likelihood estimates of the parameters assumes that the disturbances are associated with the b 's. The demand equations for any period (11) may be written in matrix form as

$$\begin{aligned} x &= (I - \gamma p')(b + u_t) + \gamma \mu \\ &= (I - \gamma p')b + \gamma \mu + (I - \gamma p')u_t, \end{aligned}$$

where x is an $n \times 1$ vector of quantities, u_t is an $n \times 1$ vector of disturbances, p is an $n \times 1$ vector of prices, μ is a scalar, equal to total expenditure, b is an $n \times 1$ vector to be estimated, and γ is an $n \times 1$ vector with elements a_i/p_i , where the a_i are parameters to be estimated.¹⁸ As discussed above, the vector u_t is assumed to be multivariate normal with mean 0 and covariance matrix D_t given by

$$D_t = E(u_t u_t') = \text{diag}(\sigma_1^2 \hat{x}_{1t}, \dots, \sigma_n^2 \hat{x}_{nt}^2)$$

where \hat{x}_{it} is the nonstochastic component of x_{it} .

The disturbance vector of the demand equations, v_t , is given by

$$v_t = (I - \gamma p')u_t = M_t u_t.$$

Clearly v is a linear transformation of u , with the covariance matrix of the v 's, Ω_t , given by $\Omega_t = M_t D_t M_t'$. Since the transformation M_t is singular, the covariance matrix Ω_t is also singular.¹⁹

To obtain maximum likelihood estimates of this system, we drop one equation and maximize the likelihood function of the reduced system. We show in the appendix that this procedure yields maximum likelihood estimates for the full system, and that the estimates obtained are independent of which equation is omitted. If we redefine, u , v , and M , after omitting the n th good, as

$$\begin{aligned} \tilde{u}_t &= (u_{1t}, \dots, u_{n-1,t}), \\ \tilde{v}_t &= (v_{1t}, \dots, v_{n-1,t}), \\ \tilde{M}_t &= M_t \text{ with the } n\text{th row deleted,} \end{aligned}$$

¹⁷ As discussed in Section 1 this problem does not arise if the elements of the covariance matrix are estimated together with the other parameters using generalized least squares.

¹⁸ The assumption that the b 's are constant over time is easily modified. We have suppressed the time subscripts on the p 's, x 's, γ 's, and μ .

¹⁹ Proof: $p'v = p'(I - \gamma p')u = p'u - (p'\gamma)p'u = 0$ so $\Omega p = E(vv'p) = 0$; hence 0 is an eigenvalue of Ω and $|\Omega| = 0$.

then \tilde{v}_t has a multivariate normal distribution with mean 0 and covariance $\tilde{\Omega}_t = \tilde{M}_t D_t \tilde{M}_t'$. Further, since the vectors $\tilde{v}_1, \dots, \tilde{v}_T$ are assumed to be independently distributed, the likelihood of $\tilde{v}_1, \dots, \tilde{v}_T$ is given by

$$L(\tilde{v}_1, \dots, \tilde{v}_T) = \prod_1^T L(\tilde{v}_t) = \prod_1^T (2\pi)^{-(n-1)/2} |\tilde{\Omega}_t|^{-\frac{1}{2}} e^{-\frac{1}{2} \tilde{v}_t' \tilde{\Omega}_t^{-1} \tilde{v}_t}$$

and the logarithm of the likelihood by

$$\mathcal{L}(\tilde{v}_1, \dots, \tilde{v}_T) = -\frac{(n-1)T}{2} \ln(2\pi) - \frac{1}{2} \sum_1^T (\ln |\tilde{\Omega}_t| + \tilde{v}_t' \tilde{\Omega}_t^{-1} \tilde{v}_t).$$

Let $\tilde{G} = \sum_1^T (\ln |\tilde{\Omega}_t| + \tilde{v}_t' \tilde{\Omega}_t^{-1} \tilde{v}_t)$; then maximizing L is equivalent to minimizing \tilde{G} . Consequently we obtain maximum likelihood estimates of b_i and σ_i for $i = 1, \dots, n$ and of a_i for $i = 1, \dots, n-1$ by minimizing \tilde{G} . The estimate of a_n is obtained by subtracting the estimates of a_1, \dots, a_{n-1} from 1.

Asymptotic standard errors of the parameters are obtained in the usual manner from the matrix of second partial derivatives of $\mathcal{L}(\tilde{v}_1, \dots, \tilde{v}_T)$ evaluated at the critical point.

For models in which the b 's are not constant, but are given by $b_{it} = b_i^* + \beta_i z_{it-1}$ where z_{it-1} is any predetermined variable, \tilde{G} is minimized with respect to the b_i^* 's and β_i 's in addition to the σ_i 's and a_i 's.

We assume that prices and total expenditure are nonstochastic, or if stochastic then independent of the disturbance terms. This avoids simultaneous equation problems.

3. RESULTS

The results of estimating the linear expenditure system assuming four different dynamic specifications for the b 's, and assuming that $Eu_{it}^2 = \sigma_i^2 \hat{x}_{it}^2$ are reported in Table I. Values in parentheses are asymptotic standard errors. As a rough rule of thumb, parameter estimates are considered "significant" (that is, significantly different from zero) if they are more than twice their standard errors. The reported R^2 statistic for each equation was computed as one minus the ratio of the variance of the disturbances to that of the dependent variable. The R^2 values are generally high, which is not surprising since the data are time series, and a large number of parameters are used in the equations. In any event this statistic is not of particular interest since the estimation procedure involves a system of equations, and a "least squares" criterion is not employed. The R^2 estimates are presented primarily for comparison with other studies of the linear expenditure system.

For the model with constant b 's the estimated marginal budget shares (a 's) are all positive, less than one, and significantly different from zero. The estimates of the b 's, however, although significant, exceed consumption values in every time period. This result is not consistent with the underlying utility function, suggesting that the constant b model is inappropriate.

The values of σ_i^2 range from zero for shelter to 9.8×10^{-4} for clothing. The estimate of zero suggests that in our stochastic formulation the u disturbance for

TABLE I^a
LINEAR EXPENDITURE SYSTEM, 1948–1965

Stochastic assumption: $Eu_{it}^2 = \sigma_i^2 \hat{x}_{it}^2$.						
	Food	Clothing	Shelter	Misc.		
I. Constant b : $b_{it} = b_i$						
a	.190 (.026)	.079 (.018)	.566 (.022)	.165 (.017)		
b	526 (64)	202 (28)	641 (153)	250 (41)		
σ^2	1.51 (.82)	9.86 (4.57)	0 (6.22)	3.80 (1.87)		
R^2	.90	.66	.99	.98		
II. Linear Time Trend: $b_{it} = b_i^* + \beta_{it}$						
a	.205 (.053)	.446 (.187)	.151 (.065)	.199 (.139)		
b^*	328 (99)	–16 (27)	208 (56)	47 (34)		
β	.068 (.178)	.089 (.412)	.142 (.121)	.082 (.136)		
σ^2	.84 (.64)	0 (.12)	1.36 (.48)	.71 (3.35)		
R^2	.96	.89	.99	.99		
III. Proportional Habit Formation: $b_{it} = \beta_i x_{it-1}$						
a	.343 (.055)	.239 (.036)	.286 (.047)	.131 (.031)		
β	.969 (.016)	.943 (.026)	.996 (.019)	.987 (.017)		
σ^2	1.19 (.49)	2.00 (1.00)	.56 (.35)	.95 (.44)		
R^2	.96	.94	.99	.99		
IV. Linear Lagged Consumption Habit Formation: $b_{it} = b_i^* + \beta_i x_{it-1}$						
a	.348 (.057)	.219 (.044)	.302 (.042)	.130 (.024)		
b^*	136 (66)	49 (49)	59 (43)	23 (26)		
β	.749 (.135)	.797 (.188)	.928 (.037)	.939 (.074)		
σ^2	1.21 (.52)	2.34 (.95)	.35 (.23)	.49 (.25)		
R^2	.96	.93	.99	.99		

^a In Tables I–IV a_i , β_i , and σ_i are pure numbers, while b_i and b_i^* are expressed in 1958 dollars. The initial value and increment for the time variable (t) is 100, while x is per capita expenditure in 1958 dollars. All estimates of σ^2 have been multiplied by 10^4 before being entered in the tables. In view of the fact that the second derivative matrix of the log likelihood function is based on numeric derivatives (see Appendix), the standard errors should be viewed with caution.

shelter is essentially zero. This does not mean that the demand for “shelter” is non-stochastic, since the disturbance in each demand equation involves all the u ’s.

One of the simplest dynamic specifications involving the b ’s is the assumption that they change by a constant amount each year—this implies a linear time trend of the form (7). None of the estimated β values and only two of the b^* values differ significantly from zero. The annual increase in b is the largest for shelter, followed by clothing, miscellaneous, and food. The estimated marginal budget shares differ considerably from those of the constant b model, which is not surprising in view of the difference in specification. Finally, calculated b_{it} values are less than consumption in all time periods, for all goods.

The habit formation model which assumes that b is proportional to lagged consumption (8b) yields significant estimated proportionality coefficients which are very close to one for all goods. The implied estimates of b_{it} are less than consumption in all time periods for shelter, clothing, and miscellaneous, and in all

but the first time period for food. The marginal budget shares differ from those in both the preceding models. The fact that the β 's are approximately unity means that supernumerary income is very small in every time period. But this does not mean that the marginal budget shares are of no interest. First, in the dynamic process, although supernumerary income is small in every period, its allocation among the different goods helps to determine future b values (through the effect of lagged consumption) and hence consumption in succeeding periods. Second, the short term effect on the consumption of various goods of a policy change which increases income (for example a tax cut) will depend solely on the marginal budget shares.

The habit formation model in which b is a linear function of lagged consumption (8a) yields marginal budget share estimates which are all significant and less than one. The β 's are also significant, but the model is not acceptable because calculated b_{it} values exceed the corresponding consumption values in every time period for all goods.

Several other versions of the linear habit formation model were estimated using different variables to represent past consumption. First, each b was linearly related to a three year moving average of past consumption. Second, each b was linearly related to the largest annual consumption value in the preceding three years. Finally, two models were estimated in which b was related to the growth rate of consumption. Although all models yielded significant parameters, the estimated b_{it} 's exceeded consumption in many time periods. Parameter estimates are not presented for any of these models because of this inconsistency with the underlying utility maximization framework.

Finally, several of the models were estimated under the assumption that the variance of the u 's is constant over time, although different for each good. The results obtained by estimating the models under this specification do not differ significantly from those reported in Table I.

It is interesting to compare the results given above with those obtained by assuming a common σ^2 value for all goods: $E(u_{it}^2) = \sigma^2 \hat{x}_{it}^2$. The advantage of this procedure is that it economizes on degrees of freedom. The results of estimating the four basic models under this assumption are presented in Table II. The proportional habit formation model again yields acceptable results, and indeed the parameter estimates differ only slightly. Similarly, the linear time trend estimates are approximately as given above, although the β values are all slightly lower and the b^* values slightly higher than those reported in Table I. Nevertheless the calculated b_{it} values are less than consumption in all periods. For the linear lagged consumption habit model the marginal budget share estimates are very close to those of Table I while the β and b^* estimates differ considerably. However, calculated b values still exceed consumption for sixty-four of seventy-two observations. Finally for the constant b model, although the marginal budget share estimates do not differ appreciably from those given above, the b values do. The latter are at least twenty per cent larger for each good than previously, and consequently exceed consumption values by more than this amount.

As mentioned above an alternative estimation procedure is to minimize the sum of squared residuals over all expenditure equations and time periods. The

TABLE II
LINEAR EXPENDITURE SYSTEM, 1948-65

Stochastic assumption: $E u_{it}^2 = \sigma^2 \hat{x}_{it}^2$.				
	Food	Clothing	Shelter	Misc.
I. Constant $b: b_{it} = b_i$				
a	.199 (.026)	.102 (.028)	.554 (.028)	.144 (.019)
b	623 (87)	258 (61)	896 (187)	308 (38)
R^2	.83	.76	.99	.97
II. Linear Time Trend: $b_{it} = b_i^* + \beta_i t$				
a	.202 (.052)	.423 (.053)	.161 (.049)	.214 (.030)
b^*	345 (34)	25 (66)	271 (20)	59 (31)
β	.058 (.023)	.065 (.043)	.136 (.016)	.074 (.018)
R^2	.96	.91	.99	.99
III. Proportional Habit Formation: $b_{it} = \beta_i x_{it-1}$				
a	.342 (.053)	.247 (.033)	.284 (.057)	.127 (.031)
β	.975 (.014)	.952 (.023)	1.002 (.017)	.993 (.014)
R^2	.96	.94	.99	.99
IV. Linear Lagged Consumption Habit Formation: $b_{it} = b_i^* + \beta_i x_{it-1}$				
a	.354 (.051)	.235 (.032)	.297 (.051)	.115 (.029)
b^*	91 (58)	25 (39)	37 (39)	8 (20)
β	.816 (.102)	.888 (.149)	.952 (.037)	.994 (.062)
R^2	.96	.94	.99	.99

results of estimating the basic models using this technique for the postwar period appear in Table III. The R^2 values are tabulated for each equation; standard errors, however, cannot be calculated. It is interesting to study the acceptability of these estimates in terms of the underlying utility function, and to compare them with the maximum likelihood results of Table I.

Marginal budget shares are positive and less than unity for all goods and models. For the constant b formulation, b_{it} values exceed consumption in all time periods, thus rendering this formulation unacceptable. The three other models, however, yield b_{it} values which in almost all instances are acceptable.

For the constant b model, marginal budget shares reported in Table III are approximately the same as those in Table I, while the least squares b values are uniformly lower than the maximum likelihood estimates, although they still exceed consumption. For the linear time trend model, the marginal budget shares of Tables I and III differ considerably for clothing and shelter; the maximum likelihood β estimates exceed, and b^* estimates fall below, those obtained using least squares. The two sets of estimates obtained for the proportional habit model are very similar. In the lagged consumption habit model the least squares β estimates exceed, and b^* estimates fall below, the corresponding maximum likelihood estimates. The b_{it} values implied by the least squares estimates are slightly less than consumption, while those obtained from the maximum likelihood estimates are slightly greater than consumption. Nevertheless, the least squares

TABLE III
LINEAR EXPENDITURE SYSTEM, 1948-65

Least Squares Estimates.				
	Food	Clothing	Shelter	Misc.
I. Constant b : $b_{it} = b_i$				
a	.182	.090	.556	.172
b	526	208	648	257
R^2	.90	.66	.99	.98
II. Linear Time Trend: $b_{it} = b_i^* + \beta_i t$				
a	.185	.367	.249	.199
b^*	383	104	233	101
β	.028	.063	.111	.041
R^2	.96	.94	.99	.98
III. Proportional Habit Formation: $b_{it} = \beta_i x_{it-1}$				
a	.358	.275	.242	.125
β	.974	.946	1.006	.997
R^2	.96	.94	.99	.99
IV. Linear Lagged Consumption Habit Formation: $b_{it} = b_i^* + \beta_i x_{it-1}$				
a	.380	.232	.273	.115
b^*	77	-5	1	-10
β	.769	.941	.976	1.036
R^2	.97	.93	.99	.99

estimates imply that the linear lagged consumption habit model is acceptable, while this conclusion is not warranted by the maximum likelihood procedure.

The final estimation procedure considered involves minimizing the sum of squared residuals with one equation omitted. The constant b model for the postwar period under the stochastic assumption of least squares was estimated four times with a different good omitted each time. The marginal budget share estimates do not differ greatly for any good, with the largest discrepancy in absolute terms being .05 for the miscellaneous category. The b estimates on the other hand depend crucially on which good is omitted from the estimation procedure. In the shelter category, for example, b values range from 418 to 880 while per capita consumption in the period ranges from 286 to 486 (in 1958 dollars). Also, when shelter is omitted from the estimation procedure, the b estimates fall below consumption in the final four or five time periods for all goods, whereas b estimates exceed consumption in all other cases.

It has often been argued that consumer tastes changed during World War II. It is interesting, therefore, to compare the basic models using data from the prewar and postwar periods. Table IV contains estimates based on data for the time period 1930-41, under the assumption that $E(u_{it}^2) = \sigma_i^2 \hat{x}_{it}^2$. A casual comparison with Table I reveals a striking difference in parameter values. Without resorting to rigorous tests it is clear that the prewar and postwar observations are not drawn from the same population.

TABLE IV
LINEAR EXPENDITURE SYSTEM, 1930-41

Stochastic assumption : $E u_{it}^2 = \sigma_i^2 \hat{x}_{it}^2$.								
	Food		Clothing		Shelter		Misc.	
I. Constant b : $b_{it} = b_i$								
a	.480	(.026)	.175	(.016)	.213	(.015)	.132	(.007)
b	.279	(32)	.122	(11)	.195	(10)	.87	(7)
σ^2	13.35	(3.72)	9.90	(3.47)	.83	(1.22)	0	(3.00)
R^2	.93		.79		.96		.98	
II. Linear Time Trend : $b_{it} = b_i^* + \beta_i t$								
a	.348	(.022)	.258	(.013)	.248	(.015)	.146	(.008)
b^*	.223	(36)	.105	(26)	.179	(18)	.75	(13)
β	.087	(.072)	.004	(.052)	.013	(.038)	.011	(.026)
σ^2	2.97	(1.49)	1.84	(1.87)	.74	(.45)	.22	(.68)
R^2	.99		.97		.97		.99	
III. Proportional Habit Formation : $b_{it} = \beta_i x_{it-1}$								
a	.399	(.040)	.233	(.021)	.214	(.022)	.154	(.010)
β	.887	(.060)	.818	(.060)	.921	(.026)	.857	(.049)
σ^2	3.56	(1.81)	4.30	(1.86)	.90	(.34)	0	(.96)
R^2	.99		.96		.97		.99	
IV. Linear Lagged Consumption Habit Formation : $b_{it} = b_i^* + \beta_i x_{it-1}$								
a	.417	(.041)	.217	(.022)	.216	(.024)	.150	(.010)
b^*	.44	(33)	.19	(17)	.43	(24)	.19	(11)
β	.757	(.108)	.709	(.104)	.722	(.114)	.679	(.110)
σ^2	3.55	(1.99)	3.29	(1.49)	.85	(.36)	0	(.91)
R^2	.99		.97		.97		.99	

Both the proportional and linear lagged consumption habit specifications are acceptable in the prewar period, in that calculated b values are less than consumption in all cases. The time trend model is again acceptable, while for the constant b model consumption falls below the corresponding b value in only eight of thirty-two cases. Also in this model the postwar estimates of b are from two to three times as large as the prewar estimates. For all models the marginal budget shares are less than one and significantly different from zero, while for the other parameters the only insignificant estimates are certain b^* values in the linear habit formation model and the β estimates in the time trend model.

4. CONCLUSIONS

Several important conclusions emerge from our study. First, for our preferred stochastic formulation only the linear time trend and proportional habit formation models are consistent with the underlying utility functions for the postwar period. On the other hand all the dynamic specifications are appropriate in the prewar period.

Second, the dynamic specification of the model is of crucial importance. Different dynamic specifications result in widely differing estimates, not only of the

parameters which characterize the dynamic specification (the b_t^* 's and the β 's), but also of the marginal budget shares. Therefore, in the absence of a criterion for choosing among the dynamic specifications, we can have little confidence in any of the estimated parameters. Further study of the merits of various dynamic specifications is clearly warranted.

Third, the estimation technique is important. The estimates obtained by minimizing the sum of squared expenditure residuals over all equations, over three equations, and by the maximum likelihood technique based on our preferred error structure differ substantially in some cases. The problem with relying on the simpler techniques is that one cannot know beforehand whether they will yield parameter estimates similar to those obtained by the more sophisticated procedure.

Fourth, the different assumptions made about the variance of the u 's affect our parameter estimates only slightly. They might make a difference, however, in estimates based on a time period in which consumption grew substantially.

Finally, it seems to us that future work in empirical demand analysis should experiment with more general specifications of the functional form of the demand equations. The fact that many of our habit models did not yield theoretically plausible parameter estimates may reflect misspecification of the functional form of the demand equations, rather than misspecification of the dynamic or stochastic structure.

University of Pennsylvania

APPENDIX

A. ESTIMATION THEOREM

In this section we show that minimizing \tilde{G} (text, page 620) is equivalent to maximizing the appropriate likelihood function.

Since u_t is assumed to be multivariate normal with covariance matrix D_t , then $v_t = M_t u_t$ is multivariate normal with covariance matrix $\Omega_t = M_t D_t M_t'$, and $w_t = \hat{P}_t v_t$ is multivariate normal with covariance matrix $S_t = \hat{P}_t \Omega_t \hat{P}_t'$, where $\hat{P}_t = \text{diag}(p_{1t}, \dots, p_{mt})$. Both Ω_t and S_t are singular, so the densities of v_t and w_t cannot be expressed directly in terms of Ω_t and S_t . A. P. Barten has shown, however, that in this case the density of w_t (ignoring a factor of proportionality) is given by

$$f(w_t) = |S_t + \ell \ell' |^{-\frac{1}{2}} e^{-\frac{1}{2} w_t' (S_t + \ell \ell')^{-1} w_t}$$

where ℓ is the $n \times 1$ column vector $(1, 1, \dots, 1)$.²⁰ Hence, the density of v_t (again ignoring a factor of proportionality) is given by

$$g(v_t) = |\Omega_t + \hat{P}_t^{-1} \ell \ell' \hat{P}_t^{-1}|^{-\frac{1}{2}} e^{-\frac{1}{2} v_t' (\Omega_t + \hat{P}_t^{-1} \ell \ell' \hat{P}_t^{-1})^{-1} v_t}$$

The likelihood of the sample (v_1, \dots, v_T) is therefore

$$L(v_1, \dots, v_T) = \prod_{t=1}^T g(v_t)$$

and the logarithm of the likelihood is

$$\mathcal{L}(v_1, \dots, v_T) = \sum_{t=1}^T \log g(v_t).$$

²⁰ A. P. Barten, *op. cit.*

We define $G(v_1, \dots, v_T)$ by

$$G(v_1, \dots, v_T) = -2\mathcal{L}(v_1, \dots, v_T),$$

so minimizing G is equivalent to maximizing L . Clearly

$$G = \sum_{t=1}^T \log |\Omega_t + \hat{P}_t^{-1} \ell \ell' \hat{P}_t^{-1}| + v'(\Omega_t + \hat{P}_t^{-1} \ell \ell' \hat{P}_t^{-1})^{-1} v_t.$$

We shall show that G and \tilde{G} differ only by a constant so that minimizing \tilde{G} is equivalent to minimizing G . We focus on a typical term of G and drop all time subscripts.²¹

(i) The typical term of G can be written as

$$(A1) \quad \log |\Omega + \hat{P}^{-1} \ell \ell' \hat{P}^{-1}| + v'[\Omega + \hat{P}^{-1} \ell \ell' \hat{P}^{-1}]^{-1} v.$$

(ii) Define a matrix E by

$$(A2) \quad E = I - e_n e_n' - (1/n) \hat{P}^{-1} \ell \ell' \hat{P}$$

where e_n is a vector consisting of all zero elements except the last one which is unity. Then

$$(A3) \quad E(\Omega + \hat{P}^{-1} \ell \ell' \hat{P}^{-1})E' = (I - e_n e_n')\Omega(I - e_n e_n') + (1/p_n^2)e_n e_n' \\ = \begin{bmatrix} \tilde{\Omega} & 0 \\ 0 & 1/p_n^2 \end{bmatrix}$$

where $\tilde{\Omega}$ is the $(n-1) \times (n-1)$ matrix obtained from Ω by deleting the last row and column.

(iii) Taking the determinant of this expression, it is easy to show that

$$|\Omega + \hat{P}^{-1} \ell \ell' \hat{P}^{-1}| = (1/p_n^2)|E|^{-2}|\tilde{\Omega}|$$

where the determinant of E is independent of the parameters to be estimated as follows from the definition of E in (A2). The typical term of G is then given by

$$(A4) \quad -2 \log |E| - 2 \log p_n + \log |\tilde{\Omega}| + v'[\Omega + \hat{P}^{-1} \ell \ell' \hat{P}^{-1}]^{-1} v.$$

(iv) Since $p'v = 0$,

$$(A5) \quad Ev = (I - e_n e_n')v = \begin{bmatrix} \tilde{v} \\ 0 \end{bmatrix}$$

where \tilde{v} is the $(n-1) \times 1$ vector obtained from v by deleting the last element. Now, using (A3) and (A5),

$$(A6) \quad v'[\Omega + \hat{P}^{-1} \ell \ell' \hat{P}^{-1}]^{-1} v = v'E'[E(\Omega + \hat{P}^{-1} \ell \ell' \hat{P}^{-1})E']^{-1}Ev \\ = (\tilde{v}', 0) \begin{bmatrix} \tilde{\Omega}^{-1} & 0 \\ 0 & p_n^2 \end{bmatrix} \begin{pmatrix} \tilde{v} \\ 0 \end{pmatrix} = \tilde{v}'\tilde{\Omega}^{-1}\tilde{v}.$$

Consequently the typical element of G can be written as

$$(A7) \quad -2 \log |E| - 2 \log p_n + \log |\tilde{\Omega}| + \tilde{v}'\tilde{\Omega}^{-1}\tilde{v}.$$

B. DATA

Constant dollar expenditures on the various categories of goods were obtained from Table 2.6 of *The National Income and Product Accounts of the United States, 1929-1965, Statistical Tables*, pp. 48-49. The price variables (p_i) are implicit deflators (1958 = 100) and were obtained from Table 8.6, pp. 162-163, of the same source. In terms of the categories reported in those tables, we defined our four broad categories of goods as follows (numbers in parentheses refer to the two tables cited above):

I. Food:

1. Food and Beverages (15).

II. Clothing:

1. Clothing and shoes (21),
2. Shoe cleaning and repair (54),
3. Cleaning, dyeing, pressing, etc. (55).

²¹ We are indebted to M. D. McCarthy and an anonymous referee for this proof.

III. *Shelter*:

1. Housing (35),
2. Household operation services (39),
3. Semidurable house furnishings (29),
4. Cleaning and polishing preparations, etc. (30),
5. Other fuel and ice (31).

IV. *Miscellaneous*:

1. Tobacco products (27),
2. Toilet articles and preparation (28),
3. Nondurable toys and sport supplies (33),
4. Barbershops, beauty parlors, and baths (56),
5. Medical care services (57),
6. Admission to specified spectator amusements (61),
7. Drug preparations and sundries (32).

Per capita consumption of a good (x_i) was calculated by dividing annual expenditure in 1958 dollars by population. The population figures are "total population residing in the United States" and are taken from Table 2 of the *Statistical Abstract of the U.S., 1967*. Per capita expenditure (μ) was calculated as $\sum_{i=1}^4 p_i x_i$.

C. NUMERICAL PROCEDURE

The algorithm used to maximize the likelihood function was "Grad x," a nonlinear routine developed by Goldfeld, Quandt, and Trotter. This procedure requires the evaluation of first and second derivatives at each iteration. Since the covariance matrix of our system is a product of three matrices, each dependent on the unknown parameters, it is difficult to express the second derivatives of the likelihood function analytically. Consequently, we calculated the second derivatives numerically, that is, by differencing the first derivatives. We also experimented with calculating both the first and second derivatives numerically, and found that numeric first derivatives were close to the analytic ones, and that second partial derivatives based on the numeric first derivatives were close to those based on the analytic first derivatives. All of our reported results are based on numeric first and second derivatives. Numeric first derivatives are convenient because they permit estimation of variants of the basic model without substantial reprogramming, and because they reduce computation time. But even using numeric first derivatives, estimation time was substantial; for example, models involving fifteen parameters and eighteen observations (e.g., the time trend and the linear lagged habit models for the postwar period) required approximately twenty-five seconds per iteration and in many cases ten to twenty for convergence. All computations were carried out on the University of Pennsylvania IBM 360/65 computer.

Because of the nature of the likelihood function and the large number of function evaluations needed to calculate derivatives in the maximization procedure, the computation time required is approximately proportional to the number of observations. Investigators considering longer time periods or a finer classification of goods might find it advisable to use a maximization algorithm which does not require second derivatives.