The Review of Economics and Statistics

VOL. LVIII MAY 1976 NUMBER 2

THE DYNAMICS OF HOUSEHOLD BUDGET ALLOCATION TO FOOD EXPENDITURES

J. Benus, J. Kmenta, and H. Shapiro*

I

PREVIOUS studies of the determinants of household food consumption have typically been based either on cross-sectional or on aggregate time series observations. The cross-sectional studies have been characterized (often necessarily) by the assumption that all households in the sample are operating in the same market, making it impossible to study the effect of price variation on the family food budget. Further, such data sets could not be used to distinguish short run versus long run behavior patterns. On the other hand, the studies based on aggregate time series could not shed any light on such important "household-specific" questions as the effect of family composition on food expenditures, or on the economies of scale in food consumption. In the present study we are able to relax many of the important limitations of both the cross-sectional and the aggregate time series studies by basing our analysis on an extensive set of sample observations on households over a number of periods of time. The actual data underlying this study are drawn from the Panel Study of Income Dynamics — a data set collected by the Survey Research Center at the University of Michigan. This panel contains budget and other data on five thousand U.S. households over a period of five years, 1968–1972.[1]

Our model of household behavior distinguishes between various sources of income, family composition and various price effects, and allows for the dynamics of adjustment in household behavior. Further, with regard to functional form, we make use of the procedure suggested by Box and Cox (1964) to determine the appropriate degree of curvature. With respect to estimation procedure, the variance-covariance matrix of disturbances is specified in accordance with the assumed nature of these pooled cross-section and time series observations. In addition, the size of our data allows us to separate carefully the hypothesis searching and the hypothesis testing phases of our study. Our procedure is to divide the entire sample into two *independent* subsamples, using one as an input in the process of specifying our model and saving the other for testing. The initial subsample, containing approximately one-quarter of the households, is used for the purpose of weeding out potential explanatory variables, determining the functional form of the relationship, and specifying the form of the variance-covariance matrix of the disturbances. On the assumption that all observations are mutually independent, there is no presumption that the households *not* included in the initial subsample should behave in the same way as the households that are included, *unless* the regularity in behavior which we want to test actually exists.

II

In formulating our model, we start with the linear expenditure system proposed by Stone (1954) which can be expressed as follows:

$$P_iX_i = c_iP_i + b_i\left(Y - \sum\nolimits_{j=1}^{n} c_jP_j\right) \qquad (1)$$

where $P =$ price, $X =$ quantity consumed, and $Y =$ current income. All subscripts refer to

Received for publication February 12, 1975. Revision accepted for publication June 18, 1975.

* We thank Professors Arnold Zellner and Arthur Goldberger and the anonymous referee for their helpful comments.

[1] See Morgan (1974). Ours is the second study to use this set of data for the analysis of household food expenditures. An initial study done by Hymans and Shapiro was confined to the determination of the "equilibrium" or "normal" level of household expenditure patterns from simple five year averages of all the data. In both studies all households with incomplete information for the five-year period have been eliminated from the initial five-thousand-households sample.

individual commodities.[2] We reparameterize the above function, specialize its reference to a particular household with members of different ages, and allow for a stochastic disturbance. In an equilibrium situation we have, for household i in period t,

$$C^*_{it} = \alpha_{i0} + \sum_{k=1}^{K} \alpha_{ik} N_{ikt} + \beta_i Y_{it} + \sum_{q=1}^{Q} \gamma_{iq} P_{iqt} + \epsilon_{it} \qquad (2)$$

where now C^*_{it} = equilibrium expenditure on food of the i^{th} household at time t, N_{ikt} = number of household members in the k^{th} age category, the subscript q refers to the q^{th} commodity, and ϵ_{it} = stochastic disturbance. Since, in general, the collective decision unit is the household and not a single family member, the above model aggregates over all household members and does not allow for each member to have his or her own parameters reflecting his or her own specific utility function. The term Y_{it}, therefore, represents total household income in period t.

The model as specified in (2) has to be modified and extended in a number of ways to approach a reasonable reflection of reality and to facilitate the testing of interesting hypotheses. In the first place, it may be presumed that the consumer's decisions with respect to food consumption may be influenced by the type of income he receives.[3] To the extent that different types of income can be associated with a different degree of "permanency," our disaggregation of income may be viewed as allowing for this effect. Second, the model as specified in (2) above, reflects a simple aggregation over all members of a given household and thus does not allow for any kind of economies (or diseconomies) of scale in expenditure. One way of rectifying this is to allow expenditure to be a quadratic function of the number of household members of each age group rather than a simple linear function as outlined above.[4] Third, there is also the possibility that the consumer may react differently to an income increase than to an income decrease. This hypothesis can be formulated by introducing into equation (2) the following terms:

$$\sum_{m=1}^{M} \delta_{im} Y_{imt} Z_{imt}$$

where

$$Z_{imt} = 1 \text{ if } Y_{imt} \geqslant Y_{im,t-1}$$
$$= 0 \text{ otherwise.}$$

We reason that, as the consumer's income increases, he will acquire a taste for a higher consumption level of some commodities, and that he may find it hard to return to the previous level of consumption of these commodities when his income goes down. If food is one of the commodities for which the consumer finds it hard to make a downward adjustment in expenditure in response to a decrease in income, then δ_{im} should be positive, otherwise it will be negative (or zero).[5] With these modifications equation (2) becomes

$$C^*_{it} = \alpha_{i0} + \sum_{k=1}^{K} (\alpha_{ik} N_{ikt} + \theta_{ik} N_{ikt}^2) + \sum_{m=1}^{M} \beta_{im} Y_{imt} + \sum_{m=1}^{M} \delta_{im} Y_{imt} Z_{imt} + \sum_{q=1}^{Q} \gamma_{iq} P_{iqt} + \epsilon_{it}. \qquad (3)$$

So far we have been concerned only with the determination of equilibrium expenditure. In reality, however, adjustments to equilibrium are not instantaneous but are delayed by habit, time needed to gather information, etc. One way of taking this delay into account is by adopting a "habit persistence" or "partial adjustment" model of the form

[2] This model, now firmly entrenched in consumer demand theory, for us simply represents a convenient starting point. Any of the well known extensions of this system (e.g., Lluch, 1973) could serve equally well, but offer few advantages in the context of our study.

[3] See Holbrook and Stafford (1971).

[4] The difficulty of appropriately allowing for the effects of household composition has characterized almost all attempts to model household budget behavior since the very earliest studies; see Stigler (1954). An interesting new conceptual approach to this problem, however, has been presented by Muellbauer (1974). Under certain conditions the model presented above is consistent with Muellbauer's scheme. A potential difficulty with our formulation is that economies of scale are attributed to the presence of multiple members in a given age group. The implication of this formulation is that increases in family size resulting from the addition of the first member in a particular age group do not lead to economies of scale. This formulation is appropriate if the types of foods consumed differ by age group (i.e., baby food for infants and hamburgers for teenagers).

[5] In a country such as France where a high value is placed on culinary aspects of life, we would be surprised if δ_{im} were negative. With respect to the United States, we are not so sure.

$$C_{it} - C_{i,t-1} = \Phi_i(C^*_{it} - C_{i,t-1}) + \xi_{it}, (0 < \Phi_i \leqslant 1) \quad (4)$$

where C_{it} = actual food expenditure of household i in period t, and ξ_{it} = stochastic disturbance. Substituting for C^*_{it} from (4) into (3) and rearranging terms, we get the following expression for household food expenditure:

$$C_{it} = \alpha^*_{i0} + \sum_{k=1}^{K} (\alpha^*_{ik}N_{ikt} + \theta^*_{ik}N_{ikt}^2) + \sum_{m=1}^{M} \beta^*_{im}Y_{imt} + \sum_{m=1}^{M} \delta^*_{im}Y_{imt}Z_{imt} + \sum_{q=1}^{Q} \gamma^*_{iq}P_{iqt} + \Phi^*_iC_{i,t-1} + \epsilon^*_{it} \quad (5)$$

where

$$\alpha^*_{i0} = \Phi_i\alpha_{i0}, \alpha^*_{ik} = \Phi_i\alpha_{ik}, \text{etc.},$$

and

$$\Phi^*_i = 1 - \Phi_i, \epsilon^*_{it} = \Phi_i\epsilon_{it} + \xi_{it}.$$

In this formulation each household is characterized by its own set of coefficients. This is probably unnecessarily general and, in any case, this number of parameters is impossible to estimate. It seems reasonable to conjecture that the parametric differences between households are due to the differences in socio-economic status. As an approximation we postulate that the socio-economic status of each household is measured by its average total income (Y'_i) over the five periods for which we have observations. We assume that each of the coefficients of equation (5) except Φ^*_i is a linear function of Y'_i so that we can write

$$\alpha^*_{i0} = \alpha_{10} + \alpha_{20}Y'_i, \alpha^*_{ik} = \alpha_{1k} + \alpha_{2k}Y'_i, \text{etc.}$$

The coefficient of $C_{i,t-1}$, Φ^*_i, was thought to depend not only on Y'_i but also on the direction of the adjustment, i.e.,

$$\Phi^*_i = \Phi_1 + \Phi_2Y'_i + \Phi_3Z_{imt}.$$

Equation (5) was derived entirely on the basis of a priori considerations. To obtain further information about the model, we utilized one-quarter of the sample observations for experimentation, the main purpose of which was to weed out some apparently irrelevant variables. This experimentation was limited to the sets of variables specified in the model derived above. As a result of this initial data analysis several variables were deleted from the model. For example, the terms

$$(\delta_{1m} + \delta_{2m}Y'_i)Y_{imt}Z_{imt} \quad (m = 1,2,\ldots,M)$$

were found to make individually insignificant contributions to the explanation of C_{it}. These terms together had expressed the idea that the response of households' food expenditure budgets to changes in income depends on: (a) the source of the income change, (b) the direction of the change, (c) socio-economic status (measured by Y'_i) and (d) overall level of different income flows. They were replaced by

$$\delta_1Y_{it}Z_{it} \text{ and } \delta_2Y_{it}Y'_{it},$$

where

Y_{it} = total income of the i^{th} household in the t-period

and

$$Z_{it} = 1 \text{ if } Y_{it} \geqslant Y_{i,t-1}$$
$$= 0 \text{ otherwise.}$$

Thus we proceed with a model where the effects of socio-economic status and the direction of income change are considerably simplified. Further, it was also found that our measure of the socio-economic status of the household had no role in determining the price coefficients of the model and thus could be dropped. With these and other minor modifications the resulting equation becomes

$$C_{it} = \alpha_0 + \sum_{k=1}^{K}(\alpha_{1k}N_{ikt} + \theta_{1k}N_{ikt}^2) + (\Phi_1 + \Phi_2Y'_i + \Phi_3Z_{it})C_{i,t-1} + \sum_{m=1}^{M}\beta_{1m}Y_{imt} + \delta_1Y_{it}Z_{it} + \delta_2Y_{it}Y'_i + \sum_{q=1}^{Q}\gamma_{1q}P_{iqt} + \epsilon_{it}. \quad (6)$$

At this stage we have as yet not directly introduced any considerations regarding the appropriate functional form of the relationship. The linear-in-parameters form of equation (6) is certainly a candidate, but not one without serious competitors. In their pioneering study, Prais and Houthakker (1955) considered the following functional forms as eligible on a priori grounds:

$$v_i = \alpha + \beta/v_0 \quad (7a)$$

$$v_i = \alpha + \beta v_0 \quad (7b)$$

$$\log v_i = \alpha + \beta \log v_0 \quad (7c)$$

$$\log v_i = \alpha + \beta/v_0 \quad (7d)$$

$$v_i = \alpha + \beta \log v_0 \quad (7e)$$

where v_i is the expenditure on the i^{th} commodity and v_0 is total expenditure (taken to be a surrogate for income). Prais and Houthakker made their final choice of functional form on the basis of more or less ad hoc tests and other considerations, but some of the recent developments in the area of transformations make it possible to approach the problem in a more rigorous and formal way.[6] In fact, it is possible to represent all of the five different functional forms in the one equation

$$\frac{v_i^\lambda - 1}{\lambda} = \alpha^* + \beta^* \left(\frac{v_0^\mu - 1}{\mu}\right) \qquad (8)$$

where the new parameters λ and μ determine the degree and type of nonlinearity, and α^* and β^* are expressions involving α and β. The elasticity of v_i with respect to v_0 is given as

$\eta_{v_i|v_0} = \beta^*(v_0^\mu / v_i^\lambda)$.

Each of the functional forms (7a) through (7e) can be represented as a special case of (8). Specifically,

if $\lambda = 1$ and $\mu = -1$, then (8) = (7a);
if $\lambda = 1$ and $\mu = 1$, then (8) = (7b);
if $\lambda \to 0$ and $\mu \to 0$, then (8) $\to$ (7c);
if $\lambda \to 0$ and $\mu = -1$, then (8) $\to$ (7d);
if $\lambda = 1$ and $\mu \to 0$, then (8) $\to$ (7e).

This transformation was applied to the food expenditure, income, and price variables of equation (6) to yield[7]

$$\begin{aligned}\frac{C_{it}^\lambda - 1}{\lambda} &= \alpha_0 + \sum_{k=1}^{K} (\alpha_{1k} N_{ikt} + \theta_{1k} N_{ikt}^2) \\ &+ (\Phi_1 + \Phi_2 Y'_i + \Phi_3 Z_{it}) \left(\frac{C_{i,t-1}^\lambda - 1}{\lambda}\right) \\ &+ \sum_{m=1}^{M} \beta_{1m} \left(\frac{Y_{imt}^\mu - 1}{\mu}\right) \\ &+ \delta_1 Z_{it} \left(\frac{Y_{it}^\mu - 1}{\mu}\right) \\ &+ \delta_2 Y'_i \left(\frac{Y_{it}^\mu - 1}{\mu}\right) \\ &+ \sum_{q=1}^{Q} \gamma_{1q} \left(\frac{P_{iqt}^\mu - 1}{\mu}\right) + \epsilon_{it}. \end{aligned} \qquad (9)$$

It should be noted that the partial adjustment process implied by equation (9) is of the form

$$\left(\frac{C_{it}^\lambda - 1}{\lambda}\right) - \left(\frac{C_{i,t-1}^\lambda - 1}{\lambda}\right) = \Phi^*_i \left[\left(\frac{C^{*\lambda}_{it} - 1}{\lambda}\right) - \left(\frac{C_{i,t-1}^\lambda - 1}{\lambda}\right)\right] + \xi_{it} \qquad (10)$$

where Φ^*_i is as defined above. If $\lambda = 1$ the above adjustment process reduces to that given in (4), i.e., C_{it} is a weighted arithmetic mean of C^*_{it} and $C_{i,t-1}$. If $\lambda \to 0$, the adjustment is related to *proportions* rather than to absolute differences, and C_{it} becomes a weighted *geometric* mean of C^*_{it} and $C_{i,t-1}$. For values of λ between 0 and 1 the process reflects yet other kinds of averaging. Since there are no a priori reasons for imposing one type of partial adjustment process in preference to another, we let the data decide.

The above model would be incomplete without the specification of the characteristics of the disturbance term ϵ_{it}. One specification applicable to pooled cross-section and time series observations is that of the "error components model" in which the disturbance can be decomposed into an effect specific to a time period but not to a cross-sectional unit and an effect specific to a cross-sectional unit but not to a time period. Then we can write

$\epsilon_{it} = \omega_i + \nu_t + u_{it}$.

Under standard assumptions each of the components is specified to have zero mean, to be homoskedastic, and to be uncorrelated with any other component. Further, there is no correlation between ω's of different cross-sectional units, between ν's of different time periods, and between u's of different cross-sectional units *or* different time periods. The implication of this specification is that the regression disturbances ϵ_{it} are homoskedastic but correlated across cross-sectional units and over time. The correlation over time remains unchanged no matter how far apart in time the disturbances are.[8] We modify these assumptions in a number of ways. First, the component specific to each household is omitted on the grounds that the household-specific effects have been captured by allowing for variation in the regression coefficients in

[6] See Box and Cox (1964), and Zarembka (1974).

[7] In previous empirical research the same power transformation has been applied to both the dependent and independent variables. In the present model the search for an appropriate functional form is somewhat more flexible inasmuch as the dependent and independent variables may enter the function with different power transformations. An alternative but related functional form has been proposed by Ramsey (1972).

[8] See Kmenta (1971), pp. 514–516.

accordance with socio-economic status. Second, we view the time-specific effect shared by all households as representing periodic shifts in the regression intercept which can also be characterized by a dummy variable for each time period involved. If, apart from the influence of the explanatory variables, there is a systematic "trend" in household food expenditure over time (e.g., change in tastes), this will be reflected in the regression coefficients of the time-period dummy variables. Finally, we assume that the disturbance follows a first-order autoregressive scheme over time and drop the assumption of homoskedasticity. With these modifications, the error term becomes

$$\epsilon_{it} = \sum_{t=1}^{T} \pi_t V_t + u_{it}$$

where

$V_t = 1$ for the t^{th} period
0 otherwise

and

$$u_{it} = \rho u_{i,t-1} + v_{it}$$

with v_{it} being a normally and independently distributed random variable which has zero mean and is uncorrelated with $u_{i,t-1}$. Further, it is assumed that

$$E(v_{it}^2) = \sigma^2 d_{it}$$

for all i, t, where d_{it} is a constant which may take on different values for different groups of observations, with the grouping being done in accordance with the predicted value of the dependent variable.

In order to test the hypothesis that $\rho = 0$, and $d_{it} = 1$, we proceeded as follows. First, the equation (9) was estimated from the data in the initial one-quarter sample by maximizing a likelihood function for normally and independently distributed random variables.[9] This maximization was accomplished by a two-dimensional search over various values of λ and μ. An examination of the likelihood surface for the different values of λ, μ revealed a single peaked function with the peak at $\lambda = 0.4$ and $\mu = 0.8$. However, since the likelihood function was relatively flat in the neighborhood of the maximizing values of λ and μ, we chose $\lambda = 0.5$ and $\mu = 1$ as the preferred functional form. Our choice was guided largely by the relative simplicity of the functional form implied. Second, using the the estimated residuals from the resulting equation, we carried out a "test" for autoregression by using the method proposed recently by Durbin (1970). This method involved regressing the residual $\hat{u}_{it}$ on all of the explanatory variables in the equation and $\hat{u}_{i,t-1}$. Lack of significance of the estimated coefficient attached to $\hat{u}_{i,t-1}$ indicates an absence of autocorrelation. This, indeed, turned out to be the case with our initial quarter-sample observations, so that we could then directly proceed to the "test" for homoskedasticity. For this purpose we used the same residuals as those used in the preceding test, but divided them into several groups according to the magnitude of food expenditure. For each group we calculated the corresponding variance and used a chi-square test to see whether their differences were significant.[10] This test failed to reject the hypothesis of homoskedasticity.[11]

With respect to the selection of particular variables to represent household composition, income, and prices, the data analysis performed on the initial sample of households led to the choice of the following variables (in addition to the interaction terms and dummy variables defined above).[12]

Household Composition
- Number of Adults (*NAD*)
- Number of Children, Age 0–4 (*N*04)
- Number of Children, Age 5–9 (*N*59)
- Number of Children, Age 10–12 (*N*1012)
- Number of Children, Age 13–17 (*N*1317)

[9] In the case where income was to enter in a reciprocal way we used the reciprocal of a linear combination of various types of income rather than treating the reciprocal of each type of income separately.

[10] See, e.g., Hoel (1954).

[11] A questionable aspect of this procedure is whether the transformation takes care of the nonlinearities in the regression equation or whether it actually serves to remove heteroskedasticity. Fortunately, Zarembka (1974) has shown that if the transformation leads to homoskedasticity of the error term, then the parameters of the relationship have been consistently estimated. This appears to be the case with our data.

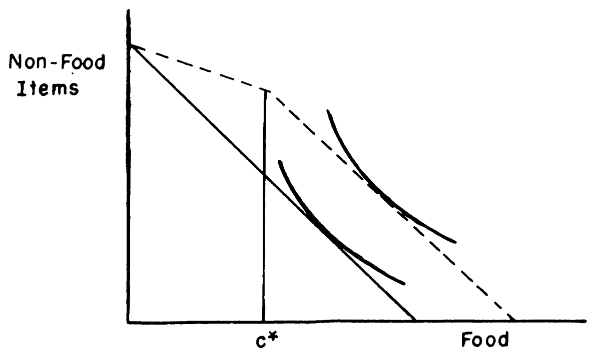
[12] In the presence of various food subsidy programs, expenditures on food understate the value of food consumption. In order to have a more accurate measure of consumption, we have added the reported value of such items as "savings due to meals purchased at school," etc., to the basic food expenditure variable.

Income[13]
- Total Income (Y)
- Transfer Income (Y_{TR})
- Imputed Income from Net Benefits of Food Subsidy Programs (Y_{FS})
- Basic Income (Y_O) $= (Y - Y_{TR} - Y_{FS})$

Prices[14]
- Food (P_F)
- Housing (P_H)
- Clothing (P_C)
- Transportation (P_T)
- Health and Recreation (P_{HR})

Before proceeding further, we should add a note of explanation concerning the income variable Y_{FS}. This variable represents an aggregate of benefits derived by a household unit from a wide variety of food subsidy programs (e.g., food stamps, food provided at work and/or school, etc.). Unfortunately, from the data available to us it was impossible to specify a variable, or a set of variables, that would adequately represent the price and income effects of these programs in all cases. We did, however, select a procedure that does deal correctly with these effects in what is probably the great majority of the cases. This procedure involves specifying the imputed dollar value of the benefits received as a component of income and as a component of consumption, and making no change in the price variable itself. As long as the household consumption of food exceeds that provided under the various subsidy programs, the marginal cost of food is appropriately measured by the existing market price and the household's new budget line is parallel to the old (without the food subsidy program) in the relevant range. Thus, simply adding the net benefit as a component of income is appropriate. In general, however, the appropriate price variable is not even uniquely defined as the budget line contains a "kink." This is illustrated by the following diagram.

FIGURE 1. — BUDGET ALLOCATION WITH FOOD SUBSIDY



where c^* = food consumed under various subsidy programs

The dotted line is the budget constraint after the implementation of a food subsidy program. As long as household food consumption is to the right of c^*, the slope of the budget constraint is unchanged.

Finally, as the last point, we have decided that the model should also take into account the opportunity of the household manager to provide food for the household as cheaply as possible. It can be surmised that this opportunity will be reduced if both the wife and the husband (or, in the case of single-adult households, the household head) are gainfully employed. Such households can be differentiated from others by the introduction of a dummy variable:

$W_{it} = 1$ if both the wife and the husband (or, in the case of single-adult households, the household head) hold a job,

$= 0$ otherwise.

The final form of the model to be tested then becomes:

[13] Total income is measured in current prices and is net of federal income taxes. In addition to labor, capital, and transfer income, this figure also includes the following important imputations: (1) savings due to performing own car and house repairs, (2) imputed rent for owner occupied housing, (3) savings generated by participation in food subsidy programs, (4) savings from growing own food, (5) savings from rent subsidy programs, and (6) other sources of income in kind.

[14] The price variables are all components of the CPI and vary not only over time, but across households at a point in time due to geographical variation in prices. Specifically, each year individuals who lived in the 23 largest SMSA's were assigned their SMSA's consumer price index. This applied to approximately half of our sample. The remaining cases were assigned their regional CPI for each year of the analysis. These price variables, while capturing different rates of inflation for different areas, do not capture the differences in the cost of living. That is, two places with identical CPI's could differ in their cost of living.

$$\frac{\sqrt{C_{it}}-1}{1/2} = \alpha_0 + \alpha_{11}NAD_{it} + \alpha_{12}N04_{it} + \alpha_{13}N59_{it} + \alpha_{14}N1012_{it} + \alpha_{15}N1317_{it} + \Theta_{11}(NAD)_{it}^2$$
$$+ \Theta_{12}(N04)_{it}^2 + \Theta_{13}(N59)_{it}^2 + \Theta_{14}(N1012)_{it}^2 + \Theta_{15}(N1317)_{it}^2 + \Phi_1\left(\frac{\sqrt{C_{i,t-1}}-1}{1/2}\right)$$
$$+ \Phi_2 Y'_i\left(\frac{\sqrt{C_{i,t-1}}-1}{1/2}\right) + \Phi_3 Z_{it}\left(\frac{\sqrt{C_{i,t-1}}-1}{1/2}\right) + \beta_{11}(Y_{0,it}-1)$$
$$+ \beta_{12}(Y_{FS,it}-1) + \beta_{13}(Y_{TR,it}-1) + \delta_1 Z_{it}(Y_{it}-1) + \delta_2 Y'_i(Y_{it}-1)$$
$$+ \gamma_{11}(P_{F,it}-1) + \gamma_{12}(P_{H,it}-1) + \gamma_{13}(P_{C,it}-1) + \gamma_{14}(P_{T,it}-1)$$
$$+ \gamma_{15}(P_{HR,it}-1) + \sum_{t=2}^{5}\pi_t V_t + \psi W_{it} + u_{it}. \quad (11)$$

III

For purposes of testing the set of hypotheses embodied in our final model (equation (11)) we make use of a set of independent data, a sample consisting of approximately three-quarters of the households in our basic data set. Table 1 contains estimates of the parameters of this model along with associated test and summary statistics generated by the data in this second subsample. It can be seen that the data in this sample fail to support any significant role for two of the price terms incorporated in our initial model — the price of clothing and the price of health and recreation — and also for one of the economies-of-scale and one of the period dummy variables. In all other cases the variables incorporated in the model appear to play a significant role and to operate in the expected fashion, at least on the basis of the standard statistical tests. These estimated parameters provide a set of inferences regarding certain aspects of the household budget allocation process and we now turn to a discussion of some of the more important and interesting of these.

TABLE 1. — PARAMETER ESTIMATES OF HOUSEHOLD CONSUMPTION EXPENDITURE (EQUATION (11)) ON SECOND SUBSAMPLE OF HOUSEHOLDS

Independent Variables	Coefficient	t-Ratio
CONSTANT	3.38	
NAD	7.80	17.24
$(NAD)^2$	—.87	11.30
$N04$	2.42	6.40
$(N04)^2$	—.22	1.48
$N59$	4.55	12.49
$(N59)^2$	—.73	5.94
$N1012$	3.18	6.78
$(N1012)^2$	—.42	2.10
$N1317$	4.90	13.70
$(N1317)^2$	—.87	7.41
P_F	—.44	5.57
P_H	.25	4.71
P_C	.06	.88
P_T	.24	5.70
P_{HR}	.03	.74
Y_{TR}	14.3*10^{-4}	13.69
Y_{FS}	121.7*10^{-4}	27.47
Y_0	9.5*10^{-4}	12.95
Y*Z	—1.13*10^{-5}	2.11
Y*Y'	—.03*10^{-6}	10.16
V_1	2.19	1.66
V_2	2.66	3.32
V_3	1.13	2.15
W	2.93	8.45
C_{t-1}	.33	29.49
C_{t-1}*Y'	9.2*10^{-6}	12.34
C_{t-1}*Z	3.5*10^{-2}	5.84

R^2 = .705. S.E.E. = 12.48. * Represents a multiplication sign.

Economies of Scale in Food Expenditure Budgets

The results presented in table 1 clearly suggest the prevalence of economies of scale in family food budgets. All the quadratic terms relating to the age distribution of the household have the negative sign that reflects the realization of economies of scale in food expenditure. These effects, however, are not uniform over all age groups with such economies being least important for the youngest age group (0 to 4 years). The estimated parameters also indicate, as expected, that older children (13 to 17 years) and adults made the largest per capita claim on the family food budget. Infants and pre-adolescents (10 to 12 years) made the smallest such claims.

Speed of Adjustment

The estimated parameters presented above clearly reveal a mechanism of partial adjustment in the response of household food budgets to shifts in the determining variables incorporated in our model. The estimated speed-of-adjustment coefficient ($\hat{\phi}_i$) can be expressed as follows from the figures in table 1:

$$\hat{\phi}_{it} = 1.0 - 0.3273 - (9.2) \times 10^{-6}(Y'_i) - .0352(Z_{it}).$$

Evaluating $\hat{\phi}_{it}$ at the sample means yields a value of 0.5716. Thus, somewhat under one-half of any indicated adjustment in the household's food budget takes place in the initial time period. As is clear from the above expression, however, the speed of adjustment is an inverse function of the family's average income and will, therefore, fall as income rises. This is due to the effect of rising income on Y'_i and on Z_{it}. Thus, households with higher and rising incomes adjust their food expenditures relatively more slowly than others. For households in the lowest quartile (by income), $\hat{\phi}_{it} = 0.64$, while for those in the highest quartile, $\hat{\phi}_{it} = 0.46$. The reader should also note that since the speed of adjustment varies with Y'_i and Z_{it} so will, for example, the relationship between the impact and steady state price elasticities. We will discuss this more fully below.

Price Effects

All the price variables enter with the expected sign indicating a negative elasticity with respect to food prices and positive cross-elasticities with respect to other prices. As noted above, however, in the case of clothing (P_C) and health and recreation (P_{HR}) we cannot reject the hypothesis that these prices play no important role. The parameter estimates in table 1 yield the following measures of the impact and long run price elasticities (evaluated at appropriate mean values) of household food expenditure.

Price Elasticity	Impact	Long Run
Food (P_F)	−2.22	−3.14
Housing (P_H)	.71	1.39
Transportation (P_T)	.68	1.18

Since our dependent variable represents expenditure on food and not quantity, the price elasticity of demand for food was obtained by deducting unity from the price elasticity of expenditure on food. No adjustment was necessary for other elasticities. The absolute values of the price elasticities — and in particular the short and the long run price elasticities for food — revealed by our sample appear to be larger than could be expected on a priori grounds or on the basis of past empirical results. The explanation may be sought in one or more of the following:

(a) As noted earlier, our measures of the price variables are given by the values of the Consumer Price Index for various regions. These measures do not allow for price differences of regional prices as they existed in the base period. That is, our food price variable P_F is, in fact, measured by the Consumer Price Index as

$$\frac{P_{F,it}}{P_{F,i0}} \times 100$$

where $t = 0$ refers to the year 1967 (i.e., the year before our first sample period). The same applies to the measurement of prices of other commodities. This would introduce no bias into our elasticity estimates if either there were no regional price differences in 1967 or if the rates of price change were the same in every region. Since neither of these conditions did apply, our elasticity estimates are likely to be biased. It can be shown that, under certain fairly plausible assumptions, the bias in the estimated coefficient of the price of food is negative (and that the coefficients of the household composition and income are not affected).

(b) Our "food expenditure" variable includes expenditure on food eaten in restaurants, etc., which contains a substantial service component. The demand for this part of "food expenditure" is likely to be considerably more price elastic than that for food prepared at home.

(c) Some distortion may be due to the fact that we include the value of all food subsidy programs in income and none in price. While, as mentioned above, this is likely to be quite appropriate for the great majority of households, there probably exist some cases for which this introduces an error of measurement into the price variable.

Income Elasticities and the Marginal Propensity to Consume

From the parameter estimates presented in table 1, it is also possible to derive point estimates of various income elasticities as well as to make a number of inferences regarding certain key relationships between the household food budget and the level, direction of change, and source of family income. The following points emerge.

a) A steady state elasticity of food expenditure with respect to basic income (Y_0) is in the neighborhood of 0.2. This is substantially lower than in many previous studies (e.g., Houthakker, 1957), but is more consistent with the results reported by Hathaway (1974), Hymans and Shapiro (1974), Brandow (1961), Girshick and Haavelmo (1947), and Tobin (1950).[15] Our estimate, however, is clearly at the lower end of the spectrum. Although our sample clearly reveals that households with higher incomes spend a lower proportion of their budget on food, there seems to be little evidence that the income elasticity falls with rising incomes. The situation, however, is somewhat different in the case of transfer income (Y_T) and income derived from food subsidy programs. First, the elasticity of food expenditures with respect to these income sources is substantially lower. Second, these elasticities do tend to fall with rising household income levels.

b) It is perhaps interesting to look at these income responses in another dimension — by considering the marginal propensity to consume (MPC) from these different income sources. Our estimates of the steady state or equilibrium MPC range from 0.05 with respect to the household's basic income (Y_0) to 0.86 with respect to income generated from food subsidy programs. Thus, household units use a significant share (14% in the steady state) of such subsidy income to increase their general purchasing power through substitutions in other portions of their budget. This result is qualitatively similar to that obtained by Hymans and Shapiro (1974), but reflects rather less substitution than this latter study. Our estimate for the MPC out of transfer income (Y_{TR}) is 0.082. The reader should recall that approximately one-half of the steady state response is felt in the initial period.

[15] The results reported by Houthakker, however, are total expenditure rather than income elasticities, and there is substantial reason to expect the income elasticity of household food consumption to be lower than the expenditure elasticity.

Labor Force Participation and Food Budgets

Finally, we have investigated the effect on household food expenditures when all the available household managers are participating in the labor force (i.e., both working). Our hypothesis was that if all managers were working, the food budget would be managed less efficiently (i.e., there would be a tendency to eat in restaurants more often, use more expensive convenience foods, etc.) because of the high opportunity cost of the time available for such efforts. The results presented in table 1 clearly support this idea (c.f., estimated coefficient for W_{it}). The estimated magnitude of this effect is interesting. Our coefficient estimate indicates that, *ceteris paribus*, households where all managers are working have food budgets on the average somewhat over 7¼% higher than similar households with at least one household manager out of the labor force.

Concluding Remarks

In the present study we have attempted to identify, from an unusually rich combined time-series and cross-section data set, certain key aspects of the process of household budget allocation to food expenditures. We have investigated the effects of family composition, of incomes differentiated by type, of the price of food and of other commodities, of the level of overall income, and of delayed budget responses. The rich data source noted above has offered substantial opportunities for the application of modern econometric techniques and enabled us to separate the hypothesis searching stage of our study and the model validation process. Although some of our results are in broad agreement with previous studies, a number of interesting new results have been obtained, particularly with respect to the effects of family composition and the dynamic response of households.

REFERENCES

Box, G. E. P., and D. R. Cox, "An Analysis of Transformations," *Journal of the Royal Statistical Society*, series B, vol. 26 (1964), 211–243.

Brandow, G. E., *Interrelations Among Demands for Farm Products and Implications for Control of Market Supply*, Pennsylvania State University College of Agriculture, Agriculture Extension Station, University Park, Pennsylvania, 1961.

Durbin, J., "Testing for Serial Correlation in Least-Squares Regression When Some of the Regressors are Lagged Dependent Variables," *Econometrica* 38 (May 1970), 410–421.

Girshick, M. A., and T. Haavelmo, "Statistical Analysis of the Demand for Food: Examples of Simultaneous Estimation of Structural Equations," *Econometrica* 15 (Apr. 1947), 79–110.

Hathaway, D. E., "Food Prices and Inflation," *Brookings Papers on Economic Activity*, The Brookings Institution, Washington, D.C., 1:1974, 63–109.

Hoel, P. G., *Introduction to Mathematical Statistics*, 2nd edition (New York: John Wiley & Sons, 1954).

Holbrook, R. S., and F. Stafford, "The Propensity to Consume Separate Types of Income: A Generalized Permanent Income Hypothesis," *Econometrica* 39 (Jan. 1971), 1–22.

Houthakker, H. S., "An International Comparison of Household Expenditure Patterns, Commemorating the Centenary of Engel's Law," *Econometrica* 25 (Oct. 1957), 532–552.

Hymans, S., and H. Shapiro, "The Allocation of Household Income to Food Consumption," in *Five Thousand American Families — Patterns of Economic Progress*, vol. II, Institute of Social Research, University of Michigan, Ann Arbor, 1974.

Kmenta, J., *Elements of Econometrics*, (New York: Macmillan Co., 1971).

Lluch, C., "The Extended Linear Expenditure System," *European Economic Review* 4 (1973), 21–32.

Morgan, J. N., ed., *Five Thousand American Families — Patterns of Economic Progress*, vol. I, II, Institute for Social Research, University of Michigan, Ann Arbor, 1974.

Muellbauer, J., "Household Composition, Engel Curves and Welfare Comparisons Between Households," *European Economic Review* 4 (1974), 103–122.

Prais, S. J., and H. S. Houthakker, *The Analysis of Family Budgets* (Cambridge: Cambridge University Press, 1955).

Ramsey, J. B., "Limiting Functional Forms for Market Demand Curves," *Econometrica* 40 (Mar. 1972), 327–341.

Stigler, G., "The Early History of Empirical Studies of Consumer Behavior," *Journal of Political Economy* 62 (Apr. 1954), 95–113.

Stone, R. A., "Linear Expenditure Systems and Demand Analysis: An Application to the Pattern of British Demand," *Economic Journal* 64 (Sept. 1954), 511–527.

Tobin, J., "A Statistical Demand Function for Food in the United States," *Journal of the Royal Statistical Society*, series A, vol. 113 (1950), 113–141.

Zarembka, P., "Transformation of Variables in Econometrics," in P. Zarembka (ed.), *Frontiers in Econometrics* (New York: Academic Press, 1974).