

THE ALLOCATION OF HOUSEHOLD INCOME TO FOOD CONSUMPTION*

Saul H. HYMANS and Harold T. SHAPIRO

University of Michigan, Ann Arbor, Mich. 48104, U.S.A.

Received September 1974, revised version received April 1975

Low-income households – those toward which various income supplement programs are aimed – not only spend a large share of their incomes on food, but exhibit a higher income elasticity of demand for food than does the rest of the population. Further, a greater proportion of the marginal income generated via welfare payments and food subsidy programs (e.g., food stamps) is devoted to food consumption than is true of wage income. These are among the major conclusions emerging from an extensive Engel Curve analysis applied to the data generated by a five-year (1968–1972) panel study of 5000 U.S. households.

1. Introduction

Differences in the level and pattern of consumption in poor and rich families have drawn the attention of social reformers for centuries, and the empirical analysis of family budgets ranks as one of the oldest branches of quantitative economic and social research.¹ The strong continuing interest in budget studies and the associated controversies over their design and interpretation have led to important advances in our understanding of certain economic phenomena, and to the further improvement of statistical techniques. Our general objective in the present study is to clarify and extend our knowledge regarding the determinants of household expenditures on food, on the basis of the analysis of data on 5,000 U.S. household budgets generated by a five-year (1968–1972) panel study conducted by the Survey Research Center at the University of Michigan (1974). The availability of so rich a data set provides many potentially fruitful paths of analysis, not all of which can be exploited in this study. The detailed purposes of our study can be enumerated as follows:

*The first version of this paper was published as a Special Study in James N. Morgan, ed., *Five thousand American families – patterns of economic progress*, Vol. II (Institute for Social Research, University of Michigan, 1974). The research was supported by a contract with the Office of Economic Opportunity and by National Science Foundation Grant GS-36932X. We are particularly grateful to our research assistant, Mr. Jay Cherlow, for his contributions to many aspects of this study; and we wish to acknowledge the valuable advice of our colleague, Professor James N. Morgan.

¹For an excellent historical survey of empirical studies of consumer behavior, see Stigler (1954).

- (i) to derive estimates of the income elasticity of food consumption for households with various sources of income,
- (ii) to investigate specifically the effect of transfer payments (welfare, social security and retirement, unemployment insurance compensation, etc.) on the level of food consumption,
- (iii) to investigate the effect of various food subsidy programs (e.g., food stamps) on the allocation of household resources,
- (iv) to investigate whether various demographic characteristics of households (e.g., age, sex, race, education, etc.) represent useful groupings for purposes of isolating a stable relationship between income and food consumption.
- (v) to compare and contrast the household food consumption/income relation for the 'neediest' quintile of our sample, with that for the remainder of the sample.

The basic approach used in this study is multivariate regression analysis aimed at the isolation of the appropriate Engel Curve. Within this framework the five-year panel of household budget data allows two alternate but not necessarily competing procedures. First, one could pool the Panel data in a manner designed to reveal the short-run dynamic response of households to various stimuli of interest. Second, one could employ the Panel data in a way designed to reveal the 'equilibrium' or 'normal' level of household expenditure patterns as a function of alternative possible configurations of the determining variables. We have chosen in this initial study to concentrate on the second approach, leaving the matter of dynamics to a future study.²

Since we are attempting to isolate the normal relationship of food consumption to the level and composition of household income, it is necessary to abstract from (or control against) behavior dominated by adjustment to a changing or transient socio-economic environment within the household unit. In order to focus the data on the kind of consumption behavior we wish to analyze, we have restructured the sample in the following fashion. First, we limited the data to be used to those households which had the same 'head and spouse' throughout the five-year period. Second, we felt that there was a need to smooth the data in order to avoid the effects of substantial year-to-year swings in the level and component sources of household income.³ We did this by averaging income, consumption, and all other such continuous variables over the five-year span,

²It should be noted that the sample is a probability sample especially constructed to over-represent low-income households. It is possible to 'weight' the data so that the resulting weighted sample is representative of the population as a whole. It would have been important to do so if we had been attempting to draw macro-economic conclusions from the data. We prefer in this study, however, not to permit the data on poorer families to be swamped by those of the far more numerous non-poor.

³Smoothing is especially important in the context of our data set since the years 1968-1972 include years of rapid growth in income as well as a period of recession (1969-1970). In addition, food subsidy programs increased in importance during the period and substantial changes occurred in the level of social security retirement benefits.

for each household with the same head and spouse.⁴ Finally, in the presence of various food subsidy programs and the option to grow one's food, *expenditures* on food understate the value of food *consumption*. In order to study food consumption we have added the reported value of such items as 'saving due to food stamps', 'saving due to meals purchased at school', and so on, to the basic food expenditure variable. As near as we can come, then, our dependent variable is a measure of food consumption, and our basic unit of analysis is a household's average food consumption over the period 1968–1972.⁵ Tables 1–3 at the end of section 2 provide some numerical characterization of the data used in this study. These tables should prove useful in setting a frame of reference for interpreting the results of the regression analysis.

2. The model

2.1. Household size, household composition and normalization of variables

It is difficult to proceed with the analysis of the effects of household income on the level of household consumption without taking explicit account of differences in household size. To begin with, variations in household size may very well have a larger effect on food consumption than do variations in household income. The simplest (and perhaps most common) way of accounting for variation in household size is to let *consumption per-capita* depend on *income per-capita*. The 'per-capita model', however, fails to allow either for economies of scale in consumption or for any differences in the age composition of households.

Economies of scale in consumption can be allowed for by considering an equation of the form

$$C/P = g(Y/P, P), \quad (1)$$

where C , Y , and P are food consumption, income, and household size, respectively. Differences in the age composition of various households, however, is a more difficult factor to incorporate. The best-known procedure is to measure household size not by a simple 'head count', but by a scale of 'equivalent persons', where a child, for example, is counted as some appropriate fraction of

⁴For a variable such as 'education' or 'location', we generally use the condition as of the last year of the period; for 'age of head' or the 'family size' variable (discussed below) we use the average value.

⁵Our data on income and consumption are not deflated for price changes. There is little purpose to deflating the data since all of our observations refer to the same five-year period for all households. The only purpose to deflation would have been to try and capture geographic variation in relative prices. We did not have data appropriate to such a deflation procedure. We did experiment with variables representing the effects of location on food consumption and this could potentially have picked up geographic variation in relative prices if that were of importance. The experiment, however, did not prove to be fruitful.

an adult. This approach was used by Engel and much more recently by Prais and Houthakker (1955). The main problem with this procedure has always been the establishment of a metric that would translate household members into 'equivalent persons' for clearly the appropriate transformation would be different for each component of expenditure. Fortunately for the case at hand, a convenient metric exists. The U.S. Department of Agriculture's minimum food requirement standards [Family Economics Review (1967)] provide the information necessary to construct a 'Food Needs' variable for each household that takes specific account of variations in the composition of households.⁶ We can, therefore, examine for each household the ratio of food consumption to food needs and thus effectively account for the different composition of households. In order to specify a variable more clearly related to the per-capita model and the 'equivalent person' approach of previous studies, we have transformed the Food Needs variable so as to produce a measure of the number of 'standard persons' in each household. In our data, the average number of persons per household (\bar{P}) is 3.85 and the average value of Food Needs per household (\overline{FN}) is \$1113, so that \$289 (equal to \overline{FN}/\bar{P}) can be taken as a measure of 'standard food needs per person'. If the Food Needs value calculated for a particular household, say FN , is divided by 'standard food needs per person' (\$289), the result, say N , can be viewed as the number of 'standard persons' in the particular household. Thus, in the empirical analysis, we use the following general type of model:

$$C/N = g(Y/N, N), \quad (2)$$

which allows for differences in household size, household composition, and economies of scale in consumption.⁷

2.2. *The specific functional form*

Throughout our analysis we have used the following two forms as the basis of our investigation:

$$\ln(C/N) = A_0 + A_1 \ln(N/\bar{N}) + A_2 \ln(Y/N), \quad (3)$$

$$C/N = a_0 + a_1(N - \bar{N}) + a_2(Y/N). \quad (4)$$

⁶The U.S. Department of Agriculture minimum food requirement standards in this data set were based on 1967 prices. For purposes of the current study, we have updated these standards to account for the inflation in food prices. To measure the overall inflation in food prices we used the U.S. Department of Labor's consumer price index for food.

⁷The explicit relationship between our dependent variable (C/N) and that suggested by the per-capita model (C/P) or the simple ratio of food consumption to food needs (C/FN) can be expressed as $C/N = (C/P)[(\overline{FN}/\bar{P})/(FN/P)]$, and $C/N = (C/FN)(\overline{FN}/\bar{P})$.

All of the variables in (3) and (4) have already been defined except for \bar{N} which refers to the mean number of standard persons in the sample. These are similar to the relationships used most widely by previous investigators. Both equations relate consumption per standard person to income per standard person and the household's size relative to the mean size of households in the sample. Eq. (3) assumes a constant income elasticity for food and eq. (4) a constant marginal propensity to consume (MPC) food. A possible difficulty, perhaps, is that neither of these forms permits a direct investigation of the existence of a 'saturation level' for household food consumption. This hypothesis could be investigated by a model of the following form:

$$C/N = \alpha - [\beta/(Y/N)]. \quad (5)$$

We have not investigated any such saturation hypothesis. Since one of our principal concerns is to study the effects on food consumption of alternative sources of income, we must, in some fashion, disaggregate the income variable shown in eqs. (3) or (4). In this context, the mathematical form of a saturation model becomes extremely difficult to implement statistically.

2.3. Further specification of model and research strategy

As noted above, we wish to investigate whether the source of household income (e.g., 'earned' income versus transfer income) has any differential effect on consumption. We approach this question by disaggregating total income. In the context of the linear model, eq. (4), we specify that the food MPC, a_2 , is given by

$$a_2 = \alpha_1 + \sum_{i=2}^k \alpha_i (Y_i/Y), \quad (6)$$

where Y_i is the i th component of total household income (Y). Eqs. (4) and (6) combined allow not only for household size, composition, and income, but also permit the food MPC to be affected by the relative importance of various income components. To understand the meaning of such a specification, suppose that total income (Y) is the sum of three components: 'labor income' (Y_1), 'transfer income' (Y_2), and the imputation 'saving due to food stamps' (Y_3).⁸

⁸If the purchase of \$10 worth of food stamps entitles the purchaser to \$15 worth of food at the food store, then the difference, \$5, is what we refer to as 'saving due to food stamps'. This is properly handled as an income imputation as long as a family's food consumption exceeds the maximum provided through the food stamp program. This is likely to be true for the vast majority of cases. In general, a food stamp subsidy lowers the *average* price of food to an eligible family and thus involves both an income and a substitution effect. But if the family purchases more than the maximum amount of food purchasable with food stamps, then the marginal cost of food to the family is the market price. Thus, in the neighborhood of its actual food purchases, the family faces a budget constraint with the same relative prices as prevail in the absence of food stamp eligibility. A pure income effect therefore suffices to describe the effect of food stamp availability.

Letting β_i ($i = 1, 2, 3$) be the food MPC's out of these three sources of income, we can write

$$C/N = a_0 + a_1(N - \bar{N}) + \beta_1(Y_1/N) + \beta_2(Y_2/N) + \beta_3(Y_3/N). \quad (7)$$

Since

$$Y_1/N = (Y/N) - (Y_2/N) - (Y_3/N),$$

we can rearrange eq. (7) to read

$$C/N = a_0 + a_1(N - \bar{N}) + \beta_1(Y/N) + (\beta_2 - \beta_1)(Y_2/N) + (\beta_3 - \beta_1)(Y_3/N). \quad (8)$$

In eq. (8) the coefficient of total income (per standard person) is the MPC out of labor income; the coefficient of transfer income is the excess of the transfer income MPC over the labor income MPC; and the coefficient of food stamp income is the excess of the food stamp MPC over the labor income MPC. Thus, eq. (8) normalizes on the labor income MPC (β_1) and highlights the degree to which the MPC's out of other income sources *differ* from the MPC out of labor income. A rather interesting phenomenon exists with respect to the coefficient of food stamp income. If the government were selling 'scrip' to low income households (say \$15 worth of scrip for \$10 in cash) and the scrip were useable for *any* purchase, one might expect to find that the MPC out of scrip income ('saving due to scrip') is about the same as that out of a cash transfer component (perhaps welfare income). But food stamps are not scrip, they can only be used to purchase (domestic) food products and indeed the food stamp program exists precisely in order to permit certain low-income households to improve their standard of food consumption. Nonetheless, it is possible for a household which becomes eligible for food stamps to 'spend' all of its food stamp income on food, spend less of its cash income on food than previously, and thereby wind up with an MPC out of food stamp income which differs little from its MPC out of some cash income source. The smaller is the coefficient of Y_3/N in eq. (8), the greater is the extent to which households manage to 'funge' their food stamp income (turn it into general scrip) by spending less cash income on food.

Finally, note that eq. (8) can also be rearranged to read

$$C/N = a_0 + a_1(N - \bar{N}) + [\beta_1 + (\beta_2 - \beta_1)(Y_2/Y) + (\beta_3 - \beta_1)(Y_3/Y)](Y/N), \quad (9)$$

which is precisely equivalent to the form given above as the pair of eqs. (4) and

(6), with the understanding

$$\begin{aligned}\alpha_1 &= \beta_1, \\ \alpha_i &= \beta_i - \beta_1, \quad \text{for } i = 2, 3, \dots, k.\end{aligned}\tag{10}$$

To estimate this model, we return to the form given in eq. (8), employ the coefficient relations in (10), and allow for an additive stochastic error term, to yield

$$C/N = a_0 + a_1(N - \bar{N}) + \alpha_1(Y/N) + \sum_{i=2}^k \alpha_i(Y_i/N) + \varepsilon.\tag{11}$$

To the logarithmic version of the model, eq. (3), we add a corresponding specification that the elasticity coefficient, A_2 , is given by

$$A_2 = \eta_1 + \sum_{i=2}^k \eta_i(Y_i/Y).\tag{12}$$

In this case the composition of household income has its direct effect on the *elasticity of demand*. Substituting (12) into (3) and allowing for a stochastic error term, yields the following 'constant elasticity' regression model:

$$\begin{aligned}\ln(C/N) &= A_0 + A_1 \ln(N/\bar{N}) + \eta_1 \ln(Y/N) \\ &+ \sum_{i=2}^k \eta_i(Y_i/Y) \ln(Y/N) + u.\end{aligned}\tag{13}$$

Considering the same three income components as in the case of eqs. (7) and (8), eq. (13) specializes to

$$\begin{aligned}\ln(C/N) &= A_0 + A_1 \ln(N/\bar{N}) + \eta_1 \ln(Y/N) \\ &+ \eta_2(Y_2/Y) \ln(Y/N) + \eta_3(Y_3/Y) \ln(Y/N) + u.\end{aligned}\tag{14}$$

The elasticity of C/N with respect to Y/N , say $\eta_{C,Y}$, is then given by the function

$$\eta_{C,Y} = \eta_1 + \eta_2(Y_2/Y) + \eta_3(Y_3/Y).\tag{15}$$

In confronting the data set with these two models we decided to separate clearly the hypothesis search ('data mining') phase of the investigation from the hypothesis testing phase. To accomplish this we took advantage of the ability to separate the entire data set into two *independent* half-samples and to conduct the search phase on one half of the data, saving the remainder for the testing

phase of the investigation. During the search phase of the study we considered the following special income components (Y_i); welfare payments, social security and retirement payments, unemployment insurance benefits, miscellaneous transfer income, income subsidies due to participation in the food stamp program and/or subsidized meals at work or school, and the income generated by raising home-grown food. Further, as part of this initial stage of our investigation, we fitted the regression models to a number of different subgroups of the half-sample in order to determine whether the mechanism governing this aspect of household behavior revealed any noticeable differences between these groups. In this connection we divided our half-sample of households into a number of subgroups defined along various demographic lines, such as age, race, sex and education of the head of household. On the whole this did not prove to be a fruitful line of research as the data revealed few noticeable differences in the behavior of these groups in the context being studied here. The same was true when we divided the households into urban and rural groupings. On the whole, these search activities produced very few insights; and those which did stand out will be discussed in detail below. An alternative 'income criterion' for partitioning the sample households, however, did prove to be quite useful. In this case we identified a subgroup defined as the lowest quintile of households in our sample in terms of income per 'standard person' (Y/N). We were particularly interested in this subgroup which may be considered the 'target' population toward which the various food subsidy and income-supplement programs are aimed. Our analysis revealed very substantial differences in the allocation of household budgets to food consumption in the 'target' and 'non-target' households.

Before turning to the results of the regression analysis, we call the reader's attention to tables 1-3 which contain relevant descriptive statistics concerning

Table 1
Sample means of consumption and income variables (dollars per
'standard person').^a

	Target households	Non-target households
Total income (Y/N)	\$798.0	\$2987.0
Food consumption (C/N)	356.5	512.6
Welfare income (W/N)	172.6	36.9
Other transfer income ($YTR-W/N$)	117.5	304.9
Income from food subsidy program ($FDSA/N$)	55.5	18.6
Income generated by savings on home-grown food (SG/N)	6.7	10.2
<i>Addendum:</i>		
Number of 'standard persons' per household	5.5	3.4

^aAll calculations refer to the initial half-sample.

Table 2

The relative size of various income flows; mean ratios of indicated variables to total income.^a

	Target households	Non-target households
'Wage' income (Y_1) ^b	0.554	0.846
Welfare income (W)	0.216	0.025
Other transfer income ($YTR-W$)	0.144	0.115
Imputed items		
(i) Saving due to food subsidy programs ($FDSAIV$)	0.078	0.009
(ii) Saving on home-grown food (SG)	0.008	0.005

^aAll calculations refer to the initial half-sample.

^b'Wage' income refers to an aggregate of all those components of total income not separately treated in the table.

Table 3

Some demographic characteristics of the initial 'half-sample'.

	Target households	Non-target households	Total
Sample size (percent of total sample)	306 (18.5)	1353 (81.5)	1659 (100.0)
<i>Characteristics of head of household</i>			
(i) Male head (percent of total sample)	154 (9.3)	1045 (63.0)	1199 (72.3)
Female head (percent of total sample)	152 (9.2)	308 (18.6)	460 (27.7)
(ii) White head (percent of total sample)	40 (2.4)	979 (59.0)	1019 (61.4)
Non-white head (percent of total sample)	266 (16.0)	374 (22.5)	640 ^a (38.6)

^aAs noted above, our sample is not chosen so as to be representative of the U.S. population as a whole. For example, in 1970 only eleven percent of U.S. families had a non-white head. This compares to a figure of 38.6 percent in our sample.

the variables used in the regression analysis. These data provide useful background for interpreting the results to be presented in the succeeding sections of the paper.

3. Some results

The experimentation discussed above led us to the following (linear) relationship governing the household's budget allocation to food consumption:

$$\begin{aligned}
C/N &= (A_0 + A_1T) + A_2(N - \bar{N}) + (A_3 + A_4T)(Y/N) + A_5(W/N) \\
&\quad + A_6[(YTR - W)/N] + (A_7U + A_8U')(FDSAV/N) \\
&\quad + A_9(SG/N) + \varepsilon.
\end{aligned} \tag{16}$$

The variables in eq. (16) are defined as follows:

- C = market value (\$'s) of food consumed by household excluding both restaurant meals and saving occasioned by the production of home grown food,
- N = number of 'standard persons' in the household,
- Y = household income, net of federal income taxes; in addition to labor, capital and transfer income this figure includes a number of important imputations,⁹
- W = welfare payments (\$'s) to husband and wife,
- YTR = total transfer payments (\$'s) to the household; the variable ($YTR - W$) is, therefore, the non-welfare component of the household's transfer income,
- $FDSAV$ = dollar value of subsidies from the food stamp program and from subsidized meals at work and/or school and other 'free' food,
- SG = dollar value of saving occasioned by growing own food,
- U = 1.0 if a household lives within five miles of the center of a city of at least 50,000 people, and a value of 0 otherwise ('urban' households),
- U' = 1.0 - U ('non-urban' households),
- T = 1.0 for 'target' households; 0 for 'non-target' households; target households are defined as those in the lowest quintile of the Y/N distribution,
- ε = stochastic error term.

This model represents both a sharper focussing on particular income components and a rather modest expansion, in a number of directions, of the general model proposed earlier (eq. 11). For example, this specification specifically isolates the income components which our initial experiments indicated would be useful. Thus although welfare payments are treated separately, all other transfer income is treated in one aggregate ($YTR - W$). The income component not shown separately (the analogue of Y_1 in our earlier illustrations) is the sum of labor and property income and a number of imputations. In subsequent discussion it will be useful to be able to refer to this component on which we have normalized

⁹The imputations cover the following items: (1) saving due to performing own car and house repairs, (2) imputed rent for owner-occupied housing, (3) saving generated by household participation in food subsidy programs, (4) saving from growing own food, (5) saving from rent subsidy programs, (6) other sources of income in kind.

in an economical way. For want of anything much better, we shall refer to this simply as 'wage income'. In addition, this version of the model allows for the principal differences exhibited by different subgroups of the population as revealed during a series of preliminary experiments. In particular, it allows, through the variable T , for a different intercept and marginal propensity to consume for the households in the 'target' population. Finally, allowance is also made for a differential reaction for urban and non-urban households to food subsidy programs.

The parameters of eq. (16) were estimated by least squares from the data generated by the households in our half-sample, with the following result (standard errors are shown in parentheses below the coefficients):

$$\begin{aligned} \frac{C}{N} = & \left[\begin{array}{c} 378.79 - 154.99T \\ (7.80) \quad (36.64) \end{array} \right] - 13.70 (N - \bar{N}) \\ & + \left[\begin{array}{c} 0.046 + 0.097T \\ (0.002) \quad (0.044) \end{array} \right] \frac{Y}{N} + 0.041 \frac{W}{N} \\ & - 0.0008 \frac{YTR - W}{N} + \left[\begin{array}{c} 0.204U + 0.398U' \\ (0.101) \quad (0.068) \end{array} \right] \frac{FDSA V}{N} \\ & - 1.233 \frac{SG}{N}, \end{aligned} \quad (17)$$

$$R^2 = 0.418, \quad SEE = \$135.0.$$

Table 4 presents the implied estimates of the MPC for food out of a number of different income sources for both target and non-target families. The parameter estimates permit a number of interesting, if tentative, inferences:¹⁰

- (i) As expected, there is a significant difference in the mechanism governing the budget allocation to food in 'target' and 'non-target' households. It is well known that low-income households spend a greater proportion of their income on food than do higher-income households. Our results show that this difference persists on the margin as well. In particular, the target households exhibit a substantially higher food MPC out of wage income. The 'target' households have an estimated wage income MPC of 0.143 compared to 0.046 for the 'non-target' units.

¹⁰We are very much aware of the inference problem resulting from all the data mining which preceded eq. (17) – the problem of the 'preliminary test estimator'. In section 4 below, we face this issue more squarely.

Table 4
The marginal propensity to consume food from various income sources [based on eq. (17)].

	'Wage' income (Y_1)	Welfare income (W)	Other transfer income ($YTR-W$)	Income from food subsidy program ($FDSAV$)		Aggregate income* (Y)	
				Urban households	Non-urban households	Urban households	Non-urban households
Target households	0.143	0.184	0.142	0.347	0.541	0.145	0.147
Non-target households	0.046	0.087	0.045	0.250	0.444	0.053	0.061

*The MPC out of aggregate income is computed according to eq. (6).

- (ii) The distinction between 'target' and 'non-target' households seems to be much weaker, however, when the linear model is used to approximate elasticities (evaluated at the appropriate mean values). The estimated value of the elasticity of food consumption with respect to total income averages about 0.325 for both the 'target' and 'non-target' households, with urban households exhibiting a somewhat lower figure than non-urban households. Table 5 presents these elasticity estimates in more detail. These estimates are similar to those reported by Brandow (1961), Lansing and Dickinson (1970), Tobin (1950), Girshick and Haavelmo (1947), and Hathaway (1974). They are, as would be expected, less than the *expenditure* elasticities reported in the work of Houthakker (1957) and Houthakker and Taylor (1970).

Table 5
Elasticity of food consumption with respect to total
income (Y) [based on eq. (17)].

	Urban households	Non-urban households
Target households	0.323	0.329
Non-target households	0.293	0.360

- (iii) The estimated coefficients on the $FDSA/N$ variable indicate that households respond to various food subsidy programs in the following way. Every additional dollar of subsidy income results in twenty to forty cents more food consumption than would an additional dollar of wage income. Although this MPC differential is far greater than that applying, say, to welfare income, it still implies that households succeed in using a significant share of such subsidy income to increase their general purchasing power. Thus, programs such as the Agriculture Department's food stamp program do induce increased consumption of food, but they also allow, through substitutions, increased consumption of other items as well. It is interesting to note that urban households 'funge' their subsidy incomes to a much greater extent than do non-urban households. Finally, we did test whether households responded any differently to the food stamp program than to other types of food subsidies (largely 'low priced' meals at work and/or school). The evidence indicated that the response was the same to the different programs.
- (iv) The coefficient on SG/N (saving on food produced at home) is rather difficult to interpret. Since this component was *not* included in the dependent variable, the negative sign indicates that this source of income decreases

the amount spent on other food items. The interpretation of the magnitude of the coefficient, however, is extremely difficult since it is dependent on the nature of what is grown and the percentage mark-up that is typical (for the items grown) at both the wholesale and retail level.¹¹ On the whole, this is not a very important source of income, even in the rural areas. In the rural areas this item averaged about 26 dollars per standard person in our half-sample.

The logarithmic equation corresponding to eq. (16) is specified as follows [see eq. (13)]:

$$\begin{aligned} \ln(C/N) = & (A_1 + A_2T) + \alpha_2 \ln(N/\bar{N}) + [\beta_0 + \beta_1T] \ln(Y/N) \\ & + \beta_2(W/Y) \ln(Y/N) + \beta_3[(YTR - W)/Y] \ln(Y/N) \\ & + [\beta_4U + \beta_5U'](FDSA V/Y) \ln(Y/N) \\ & + \beta_6(SG/Y) \ln(Y/N) + u. \end{aligned} \quad (18)$$

As with the linear model, the parameters of this relationship were estimated by least squares from the data generated by the households in our half-sample yielding the following results:

$$\begin{aligned} \ln(C/N) = & \left[\begin{array}{cc} 3.794 & -1.652T \\ (0.115) & (0.423) \end{array} \right] - 0.082 \ln(N/\bar{N}) \\ & + \left[\begin{array}{cc} 0.299 & +0.243T \\ (0.015) & (0.063) \end{array} \right] \ln \frac{Y}{\bar{N}} + 0.021 \left(\frac{W}{\bar{Y}} \right) \ln \frac{Y}{\bar{N}} \\ & + 0.002 \left(\frac{YTR - W}{\bar{Y}} \right) \ln \frac{Y}{\bar{N}} + \left[\begin{array}{cc} 0.109U + 0.193U' \\ (0.042) & (0.026) \end{array} \right] \\ & \times \left(\frac{FDSA V}{\bar{Y}} \right) \ln \frac{Y}{\bar{N}} - \frac{0.636}{(0.063)} \left(\frac{SG}{\bar{Y}} \right) \ln \frac{Y}{\bar{N}}, \end{aligned} \quad (19)$$

$$R^2 = 0.449, \quad SEE = 0.256.$$

¹¹The *SG* variable measures only the saving due to growing food at home, and is not, therefore, a measure of consumption. Since *no* corresponding measure of consumption was determined in the survey, we excluded *SG* itself from the dependent variable.

The inferences to be drawn from the above parameter estimates are qualitatively very similar to those generated by the linear model which was discussed above. Using the estimated coefficients from the above equation together with eq. (15) the estimated elasticity of food consumption with respect to total income is approximately 0.55 for target households and 0.30 for non-target households. Recall that the linear model generated a total income elasticity of about 0.325 for both target and non-target households (see table 5). Thus, the income elasticity estimates generated for the target households are substantially higher in the logarithmic model. The estimated logarithmic model implies a kind of 'strong Engel's Law' in which the proportion of income spent on food declines as income increases and declines more rapidly at higher income levels (i.e., the higher-income group has the lower-income elasticity).

In order more thoroughly to investigate the difference between 'target' and 'non-target' households with respect to the budget allocation for food consumption, we considered the following three models:

Model A

$$\begin{aligned} C/N = & (A_0 + A_1T) + A_2(N - \bar{N}) + A_3(Y/N) + A_5(W/N) \\ & + A_6[(YTR - W)/N] + (A_7U + A_8U')(FDSA/N) \\ & + A_9(SG/N) + \varepsilon_A. \end{aligned}$$

Model B

$$\begin{aligned} C/N = & (A_0 + A_1T) + A_2(N - \bar{N}) + (A_3 + A_4T)(Y/N) + A_5(W/N) \\ & + A_6[(YTR - W)/N] + (A_7U + A_8U')(FDSA/N) \\ & + A_9(SG/N) + \varepsilon_B. \end{aligned}$$

Model C

(a) *Target households*

$$\begin{aligned} C/N = & \alpha_0 + \alpha_1(N - \bar{N}) + \alpha_2(Y/N) + (\alpha_3U + \alpha_4U')(W/N) \\ & + \alpha_5[(YTR - W)/N] + (\alpha_6U + \alpha_7U')(FDSA/N) \\ & + (\alpha_8U + \alpha_9U')(SG/N) + \varepsilon_{C,1}. \end{aligned}$$

(b) *Non-target households*

$$\begin{aligned} C/N = & \beta_0 + \beta_1(N - \bar{N}) + \beta_2(Y/N) + \beta_3(W/N) + \beta_5[(YTR - W)/N] \\ & + (\beta_6U + \beta_7U')(FDSA/N) + \beta_8(SG/N) + \varepsilon_{C,2}. \end{aligned}$$

These comprise a series of 'nested' models in the sense that model B is a special case of model C (where: $\alpha_3 = \alpha_4 = \beta_3$, $\alpha_5 = \beta_5$, $\alpha_6 = \beta_6$, $\alpha_7 = \beta_7$, $\alpha_8 = \alpha_9 = \beta_8$), and model A is a special case of model B (where $A_4 = 0$). Model A allows only for a separate intercept term for target households, while model B allows in addition, for separate MPC's out of wage income, and model C relaxes even more constraints. In this context, therefore, we can perform an *F*-test to decide which of the restrictions should be relaxed.

We proceeded in a sequential manner, beginning by comparing models A and B. The appropriate *F*-test is defined as follows:

$$F_{1,1649} = [(ESS_A - ESS_B)/ESS_B]/[df_B/(df_A - df_B)],$$

ESS_i = Error sum of squares for model *i*,

df_i = degrees of freedom for error in model *i*.

This test statistic is appropriate for testing the hypothesis that models A and B are the same (i.e., that the data contain the restrictions implicit in model A). If the computed *F*-value from the above statistic is greater than the critical *F*-value (for the appropriate degrees of freedom) then we reject the null or 'no difference' hypothesis in favor of model B. The computed *F*-value, in this case 4.826, enables us to reject the null hypothesis (that both models are the same) at the 0.025 level. Thus, model B with separate marginal propensities to consume out of wage income for target and non-target households is preferred to model A. In a perfectly analogous fashion we tested the restrictions implied by model B relative to model C. In this case, however, we were unable to reject the null hypothesis. Model B – which is eq. (16) as well – therefore remains our preferred specification.

4. A test of the model

As outlined above, one aspect of our research strategy was to separate the 'hypothesis search' phase of our investigation from the hypothesis testing phase. Our procedure was to divide our entire sample into two independent 'half-samples', saving the second of these samples for a test of the model (or models) isolated by preliminary data analysis on the first sample. The parameter estimates presented in section 3 above were all derived in the hypothesis search phase of our study on the basis of data analysis on the initial half-sample. In order to test these models [eqs. (16) and (18)] we then estimated the parameters of these relationships from the data generated by the second sample. The following parameter estimates were derived when the linear model was applied to the second half-sample:

$$\begin{aligned} \frac{C}{N} = & \left[\begin{array}{cc} 405.19 & -226.32T \\ (7.31) & (32.85) \end{array} \right] - 16.19 (N - \bar{N}) \\ & + \left[\begin{array}{cc} 0.036 & + 0.134T \\ (0.002) & (0.035) \end{array} \right] \left(\frac{Y}{N} \right) + 0.052 \frac{W}{N} - 0.013 \frac{(YTR - W)}{N} \\ & + [0.474U + 0.595U'] \frac{FDSA V}{N} - 1.19 \frac{SG}{N}, \end{aligned} \quad (20)$$

$$R^2 = 0.407, \quad SEE = \$136.3.$$

The above parameter estimates together with their associated standard errors confirm all the major hypotheses isolated during the hypothesis search phase of our investigation. In particular, we note the continued strong evidence in favor of allowing both a separate intercept and a separate MPC for 'wage' income for the target households. Further, the roles discussed above with respect to the influence of welfare income, income generated by food subsidy programs, etc., on the budget allocation process are fully supported by these estimates. Table 6 presents the MPC estimates by source of income implied by eq. (20) and compares them to those implied by our data analysis on the initial half-sample (see table 4).

Although in the framework of the linear model the evidence from the second half-sample qualitatively confirms the conclusions derived in the initial phase of our investigation, table 6 does reveal a number of quantitative differences in the point estimates of a number of the MPC's. In terms of the MPC for total income (Y), the second half-sample generates an average estimate of 0.175 and 0.038 for the target and non-target families, respectively. This compares to the initial estimates of 0.146 and 0.056 derived from the first half-sample. In order to test formally the hypothesis of equal coefficients in the two half-samples we carried out the F -test suggested by Chow (1960). The results of the Chow test led us to reject the hypothesis of equality. That is, despite the fact that both half-samples produce qualitatively equivalent conclusions, the data do not support the hypothesis of equivalent parameter values in the two half-samples. Alternatively, one may view the Chow test as casting suspicion on the linear form of the model.

When the logarithmic model was applied to the second half-sample we obtained the following result:

Table 6
The marginal propensity to consume food from various income sources.

	'Wage' income (Y_1)	Welfare income (W)	Other transfer income ($YTR-W$)	Income from food subsidy program ($FDSAV$)		Aggregate income (Y)	
				Urban households	Non-urban households	Urban households	Non-urban households
<i>(a) 'First' half-sample (hypothesis search phase)</i>							
Target households	0.143	0.184	0.142	0.247	0.541	0.145	0.147
Non-target households	0.046	0.087	0.045	0.250	0.446	0.053	0.061
<i>(b) 'Second' half-sample (hypothesis test phase)</i>							
Target households	0.170	0.222	0.157	0.644	0.765	0.174	0.175
Non-target households	0.036	0.088	0.023	0.510	0.631	0.037	0.038

$$\begin{aligned}
 \ln \frac{C}{N} = & \left[\begin{array}{cc} 3.83 & -1.61T \\ (0.13) & (0.40) \end{array} \right] - \frac{0.071}{(0.016)} \ln(N/\bar{N}) + \left[\begin{array}{cc} 0.297 & +0.223T \\ (0.016) & (0.058) \end{array} \right] \\
 & \times \ln\left(\frac{Y}{N}\right) + \frac{0.030}{(0.007)} \frac{W}{Y} \ln\left(\frac{Y}{N}\right) + \frac{0.004}{(0.005)} \frac{(YTR-W)}{Y} \ln\left(\frac{Y}{N}\right) \\
 & + \left[\begin{array}{cc} 0.162U & +0.209U' \\ (0.055) & (0.033) \end{array} \right] \frac{FDSA V}{Y} \ln\left(\frac{Y}{N}\right) \\
 & - \frac{0.669}{(0.078)} \frac{SG}{Y} \ln\left(\frac{Y}{N}\right), \tag{21}
 \end{aligned}$$

$$R^2 = 0.425, \quad SEE = 0.275.$$

Table 7 compares the estimated coefficients derived in the framework of the logarithmic model from the first and second half-sample.

Table 7
A comparison of the estimated coefficients of the logarithmic model
in the 'first' and 'second' half-samples.

Variable	Estimated coefficients	
	First half-sample	Second half-sample
Constant	3.794	3.833
T	-1.652	-1.614
$\ln(N/\bar{N})$	-0.082	-0.071
$\ln(Y/N)$	0.299	0.297
$T\ln(Y/N)$	0.243	0.223
$W/Y * \ln(Y/N)$	0.021	0.030
$(YTR-W)/Y * \ln(Y/N)$	0.002	0.004
$UFDSA V/Y * \ln(Y/N)$	0.109	0.162
$U'FDSA V/Y * \ln(Y/N)$	0.193	0.209
$SG/Y * \ln(Y/N)$	-0.636	-0.669

In this case the results are extraordinarily similar *both* qualitatively *and* quantitatively. The total income elasticity of food consumption estimated from the second half-sample averaged 0.53 and 0.29 for the target and non-target households, respectively. This compares to estimates of 0.55 and 0.30 derived in the hypothesis search phase of our study. Finally, we again employed the Chow procedure to test the hypothesis that the half-samples imply equal

coefficient values in the logarithmic model. As one would expect from the evidence presented in table 7 we were unable to reject the hypothesis of equality in this case. In view of those results we would be inclined to prefer the logarithmic over the linear version of the model, and it would seem to be appropriate to employ the entire data set to estimate the parameters of the logarithmic model so as to gain the advantages of the larger sample size. The estimated coefficients based on the entire data set are as follows:¹²

$$\begin{aligned} \ln \frac{C}{N} = & \left[\begin{array}{cc} 3.845 & -1.567T \\ (0.087) & (0.284) \end{array} \right] - \frac{0.078}{(0.010)} \ln(N/\bar{N}) + \left[\begin{array}{cc} 0.296 & +0.220T \\ (0.011) & (0.042) \end{array} \right] \\ & \times \ln\left(\frac{Y}{N}\right) + \frac{0.026}{(0.004)} \frac{W}{Y} \ln\left(\frac{Y}{N}\right) + \frac{0.003}{(0.003)} \frac{YTR-W}{Y} \ln\left(\frac{Y}{N}\right) \\ & + [0.132U + 0.200U'] \frac{FDSA V}{Y} \ln\left(\frac{Y}{N}\right) \\ & - \frac{0.651}{(0.019)} \frac{SG}{Y} \ln\left(\frac{Y}{N}\right), \end{aligned} \quad (22)$$

$$R^2 = 0.435, \quad SEE = 0.266.$$

These coefficients imply the elasticities for food consumption with respect to total income which are shown in table 8.

Table 8
Estimated elasticities of food consumption with respect to total income [pooled sample, eq. (22)].

	Urban households	Non-urban households
Target households	0.526	0.532
Non-target households	0.301	0.303

¹²While it does seem to be appropriate to pool the two data sets to improve the efficiency of the parameter estimates, it does not seem to be appropriate to assume that the first half-sample is providing as many degrees of freedom as the second half-sample, an assumption implicit in the results shown in eq. (22). It is obvious from table 7 that any reasonable scheme for giving more weight to the second half-sample would produce only trivial differences from the coefficient estimates shown in eq. (22). The reader might prefer, however, to use eq. (21) to determine the standard error of estimate for the logarithmic equation. In that case the coefficient standard errors shown in eq. (22) should be multiplied by 1.034 (= 0.275/0.266) to account for the 'spuriously low' standard error of estimate in eq. (22).

Finally, we computed a new set of estimated income elasticities under the assumption that no welfare or food subsidy programs existed (i.e., $W = FDSA V = 0.0$). In this situation the only supplemental income programs available would be the various social insurance programs. The effect of this assumption on the aggregate income elasticity of food consumption was marginal. For both target and non-target households the estimated elasticity was lowered a minor amount (0.02 at most). Thus, although these programs permit a higher *level* of food consumption, they do not have any substantive effect on the income elasticity of food consumption.

5. Conclusions

The present study has attempted to isolate, from a body of data on household budgets, certain key parameters governing the household budget allocation to food consumption. In this initial study we have concentrated on the equilibrium or steady state properties of this mechanism. Within this context we have investigated the differential roles of various types of household income and have dealt separately with the budget allocation process in the case of the 'neediest' households in our sample. To summarize our principal findings:

- (i) In the allocation of resources to food consumption, the major distinguishing element is membership in the target group (i.e., low income per standard person). Households in the target group display an income elasticity of food consumption slightly in excess of 1/2; for non-target households the corresponding figure is about 3/10.
- (ii) Welfare income is devoted to food consumption to a greater extent than is true of wage income, but the difference is small at the margin with the result that the overall income elasticity is barely affected by empirically relevant variations in the share of income accounted for by welfare payments. We found no evidence to indicate that non-welfare transfers should be treated differently from wage income in the analysis of food consumption.
- (iii) Income from food subsidy programs (e.g., food stamps) proved to be even more strongly devoted to food consumption than did welfare payments. However, the food subsidy programs are sufficiently small as a source of income (even for the target group) that their effect on the income elasticity of food consumption is small.
- (iv) The evidence clearly implies that the various income-supplement programs aimed at the poor permit a higher standard of food consumption. But they have minimal effects on either the overall marginal propensity to consume food or the income elasticity of food consumption.

Although a number of interesting and useful results were obtained in this study, the basic data set remains rich in possibilities for further analysis.

Perhaps one of the most important aspects of the budget allocation process that remains to be investigated with these data is the dynamics of household response to short-run variations in income, income sources, and household composition. A pooling of the Panel data in the manner suggested by Nerlove (1971) would provide an unusually good opportunity to study the dynamic aspects of the budget allocation process.¹³

¹³Such a study is now being undertaken by Shapiro in conjunction with our colleague Professor Jan Kmenta and Dr. Jacob Benus.

References

- Brandow, G.E., 1961, Interrelations among demands for farm products and implications for control of market supply (Pennsylvania State University College of Agriculture, University Park, Penn.).
- Chow, G.C., 1960, Tests of equality between sets of coefficients in two linear regressions, *Econometrica* 3, 591–605.
- Family Economics Review, 1967, United States Department of Agriculture.
- Girshick, M.A. and T. Haavelmo, 1947, Statistical analysis of the demand for food: Examples of simultaneous estimation of structural equations, *Econometrica* 2, 79–110.
- Hathaway, D.E., 1974, Food prices and inflation. Brookings Papers on Economic Activity (The Brookings Institution, Washington, D.C.) 63–109.
- Houthakker, H.S., 1957, An international comparison of household expenditure patterns, commemorating the centenary of Engel's Law, *Econometrica*, 532–552.
- Houthakker, H.S. and L.D. Taylor, 1970, Consumer demand in the United States, 2nd ed. (Harvard University Press, Cambridge, Mass.).
- Lansing, J. and K. Dickinson, 1970, Consumption patterns of the poor, mimeo. (Survey Research Center, University of Michigan, Ann Arbor, Mich.).
- Morgan, J., ed., 1974, Five thousand American families – Patterns of economic progress, I, II (Institute for Social Research, University of Michigan, Ann Arbor, Mich.).
- Nerlove, M., 1971, Further evidence on the estimation of dynamic economic relations from a time series of cross sections, *Econometrica* 39, 359–382.
- Prais, S.J. and H.S. Houthakker, 1955, The analysis of family budgets (Cambridge University Press, Cambridge, Mass.).
- Stigler, G., 1954, The early history of empirical studies of consumer behavior, *Journal of Political Economy*, 95–113.
- Tobin, J., 1950, A statistical demand function for food in the United States, *Journal of the Royal Statistical Society A* 113, 113–141.