ENGEL CURVE ESTIMATION FOR A COMPLETE SYSTEM OF AGGREGATE EXPENDITURE COMPONENTS

Wen S. Chern[1] and Horacio Soberon-Ferrer[2]
University of Maryland, College Park

ABSTRACT

This study analyzed the relationship between the Engel Curve and the complete demand systems and demonstrated that the linear form can be derived from several plausible complete demand systems. Furthermore, the inclusion of the demographic variables in the additive fashion can be justified theoretically. The study used the Box-Cox general form to validate the linear form which is theoretically plausible and the semi-log form which is empirically popular. The empirical study dealth with the estimation of the Engel functions for eight aggregate expenditure components. The results strongly suggest that the validity of various functional forms need to be assessed from the standpoint of a complete demand system. The empirical results further show that both the Box-Cox and linear forms are superior to the semi-log form.

INTRODUCTION

The purpose of this paper is to present the econometric results of estimating Engel functions for a complete system of aggregate expenditure components. The study focuses on the following aspects of the study of income-consumption relationships: (1) identification of the theoretical basis for incorporating the demographic and socioeconomic variables in the Engel function, and (2) testing of appropriate functional forms using the Box-Cox transformation. We will undertake these tasks using the framework of a complete demand system.

The main interest in estimating an Engel function is to obtain an estimate of the expenditure elasticity of consumption or demand. Expenditure elasticities provide very useful information about consumers' behavior and budget allocation. This information may be used to analyze the impacts of government income transfer or assistance programs. Knowledge about expenditure elasticities is also extremely useful to the private sector in analyzing relative market strength when household incomes change.

The classical work of Prais and Houthakker [12] remains one of the best references in the study of Engel functions. Despite the numerous studies done in the three decades since Prais and Houthakker published their work, the issue of appropriate functional forms remains unresolved. However, there have been several important developments in this area. One involves the methodology of incorporating the demographic and socioeconomic variables in Engel functions. These variables may be termed household characteristics. While it is legitimate to consider these characteristics in the Engel function, approaches for incorporating these variables are often ad hoc. Typically, these variables are simply added as explanatory variables and their forms follow whatever functional form is being used. If these are dummy variables then no transformations are made. This approach has been commonly applied to a single expenditure item or a selected group of expenditures such as food. Examples of this type of study related to food include Blaylock and Smallwood [4], Adrian and Daniel [1], Davis [7], and Basiotis, et al. [2].

The other significant area of development deals with modeling a complete demand system. The classical model is, of course, the linear expenditure system (LES). More recently developed models include the quadratic expenditure system (QES), the almost ideal demand system (AIDS), the CES demand system, the basic translog demand system (BTL), the generalized translog demand system (GTL), and many of their variations. These complete demand systems can be used to derive the specific functional forms used for the Engel curve analysis.

The recent studies by Pollak and Wales [10, 11] are most relevant to the present study. Specifically, Pollak and Wales developed and compared various methods for incorporating demographic variables in a complete demand system. In this paper we will use these methods for incorporating demographic and other socioeconomic variables in Engel functions.

Our study also takes into account other economic and statistical considerations. The Box-Cox transformation has been widely used for testing the appropriateness of functional forms. Examples of its application in the food demand area include Chang [6] and Blanciforti, et al. [3]. In this study, we also use the Box-Cox transformation for testing the functional forms.

The empirical application addresses the estimation of Engel functions with demographic and socioeconomic variables for a complete system of aggregate expenditure components. Specifically, eight expenditure categories are defined to exhaust the total expenditure. The data base used is the 1980-81 Consumer Expenditure Survey conducted by the Bureau of Labor Statistics.

---

[1]Professor, Department of Textiles and Consumer Economics

[2]Assistant professor, Department of Textiles and Consumer Economics

FUNCTIONAL FORMS

The fundamental Engel relationship can be expressed as

$$y_i = h^i(\mu),\ i = 1,\ldots,g$$

where y_i may be the expenditure (X_iP_i) or expenditure share (W_i) and μ may be the household income or total expenditure. In the literature many functional forms have been considered and tested. The most commonly tested forms include linear, double-log, semi-log, hyperbolic, and their variations.[3]

Since the Engel function expresses the effects of income on budget allocation among groups of goods and services holding prices of commodities constant, the only theoretical condition relevant for the analysis is the adding-up restriction. This restriction states that if we sum up all expenditures, the sum should be equal to total expenditure (or income). Using this criterion, the linear form would satisfy this budget identity while the double-log and the semi-log would not.

Since a demand function reduces to an Engel function when price effects are assumed constant, we should be able to derive the Engel function directly from a complete demand system. Specifically, each complete demand system suggests a specific form for the Engel relationship. As indicated earlier, many complete demand systems have been developed and estimated. It would be important to know whether these commonly used functional forms correspond to any of the well known complete demand systems. In order to derive the Engel function from a complete demand system, we can treat all price terms as constant. The results of this investigation show that (1) the linear form can be derived from the LES, CES, and the generalized CES demand systems, (2) the QES, BTL, GTL generate specific nonlinear functional forms which differ from those commonly used in Engel curve estimation, (3) the AIDS would generate a semi-log form using expenditure share as the dependent variable, (4) the double-log and semi-log (using the expenditure as the dependent variable) cannot be derived from the theoretically plausible demand systems.

Therefore among the linear, semi-log, and double-log forms, the first remains the most plausible in terms of its theoretical foundation. Specifically, let us consider the generalized CES demand system which is expressed as

$$(1)\quad W_i = \frac{P_ib_i}{\mu} + \frac{a_i^c P_i^{1-c}}{\Sigma a_k^c P_k^{1-c}}\left[1 - \frac{\Sigma P_k b_k}{\mu}\right],\ \Sigma a_k = 1$$

where a_is, b_i's and c are parameters.

If we treat prices as constants, then Eq. (1) is reduced to

$$(2)\quad W_i = \frac{k'}{\mu} + k''\left(1 - \frac{k'''}{\mu}\right)$$

or

$$(3)\quad P_iX_i = k' + k''\mu - k''k''' = k + k''\mu$$

where k, k', k'' and k''' are functions of the constant terms in Eq. (1).

Equation (3) is, of course, a linear Engel function between the expenditure for category i (i.e., P_iX_i) and the total expenditure.

The relationship to the plausible demand systems is not the only consideration in the Engel function specification. There are other important economic and statistical considerations. The various functional forms of the Engel function provide means for testing the marginal propensity to consume and income elasticity. The linear form implies (1) constant propensity to consume and (2) tendency to unitary income elasticity as income increases. The double-log form implies a constant income elasticity. On the other hand, the semi-log form implies that (1) the marginal propensity to consume varies inversely with income and, therefore, (2) income elasticity decreases as income increases. Thus, the semi-log form allows the same goods to be luxuries when income is low, and necessities when income is high. This property may be viewed as a reasonable characterization of many goods and services such as food. On statistical ground, double-log and semi-log have been found to have better statistical fit than linear and other functional forms. Consequently, double-log and semi-log forms remain the most popularly used benchmark functional forms for comparison in empirical studies.[4]

METHODS FOR INCORPORATING DEMOGRAPHIC VARIABLES

Demographic variables such as household size, age of household head, race or ethnicity are important determinants of household consumption. However, these variables were typically not included in the derivation of the theoretically plausible demand systems. If these variables are simply added into the demand system after it is derived, then one cannot be sure whether or not the original indirect or direct utility functions still hold.

Recently, Pollak and Wales [10, 11] suggested various procedures for incorporating such variables into empirical demand analysis. Demographic translating and demographic scaling are the two major procedures suggested. Under these procedures, parameters of the demand system are assumed to depend upon the demographic variables. Demographic translating introduces n "translation parameters," $d_i,\ldots\ d_n$, into each demand

---

[3] See Prais and Houthakker [12], pp. 83-85.

[4] For example, see Kakwak and Podder [8].

system with n goods and services. Suppose the original demand function is

$$Y_i = h^i(P_i, \mu).$$

The modified system with demographic translation is

$$(4) \quad Y_i = d_i + h^i(P, \mu - \Sigma P_k d_k)$$

The translation parameters, d_i, may be specified to be a linear function of the demographic variables N_r, as

$$(5) \quad d_i = \Sigma_r \delta_{ir} N_r$$

If the original demand system is generated by the indirect utility function $\psi(P, \mu)$, then it can be shown that the modified system satisfies the first order conditions corresponding to the indirect utility function $\psi(P, \mu - \Sigma P_k d_k)$.

Alternatively, demographic scaling introduces n "scaling factors" and replaces the original demand system by the following modified system:

$$(6) \quad Y_i = m_i h^i (P_1 m_1, \ldots, P_n m_n, \mu).$$

Therefore, only the m's depend upon the demographic variables and they may be expressed as a linear function as

$$m_i = 1 + \Sigma_r \varepsilon_{ir} N_r$$

Again if the original demand function is plausible, the modified system is also.

Pollak and Wales applied these two procedures to the QES, BTL, and GTL demand systems. Their results show that demographic translation results in a lower likelihood value than demographic scaling. However, the significance of these differences can not be formally tested. The demographic translating was recently applied by Kokoski [9].

Following Pollak and Wales, we can modify the Engel functions obtained from various theoretically plausible demand systems using the translating and scaling methods. The resulting Engel functions are highly nonlinear in demographic variables except those obtained from LES or generalized CES. Specifically, if we use the translation equation (4), the original generalized CES function in Eq. (3) can be rewritten as:

$$(7) \quad P_i X_i = d_i + k + k''(\mu - \Sigma P_k d_k)$$

$$= d_i + k + k''\mu - k''\Sigma P_k d_k.$$

However since P_k's are treated as constants in the Engel function and d_i's are linear functions of demographic variables as expressed in (5), Eq. (7) can be reduced to:

$$(8) \quad P_i X_i = k + k''\mu + \Sigma k''_{ir} N_r$$

where $k''_{ir} = \delta_{ir} - k''\Sigma P_k \delta_{kr}$

Eq. (8) is linear in demographic variables N_r.

BOX-COX TRANSFORMATION

The preceding analysis indicates that among the three commonly used linear, double-log, and semi-log forms, only the linear form can be derived from the theoretically plausible demand systems. Furthermore, it is theoretically plausible to include the demographic variables in the linear Engel function in the additive fashion. Both the double-log and semi-log are ad hoc functional forms. However, the semi-log form has a more plausible implication of income elasticity than both the linear and double-log specifications.

The Box-Cox transformation provides a functional form which contains linear, double-log, and semi-log as special cases. For this study, we would prefer semi-log over the double-log. Consequently, the following Box-Cox transformation is adopted:

$$(9) \quad P_{ij}X_{ij} = \alpha_o + \alpha_1\mu^* + \Sigma_k \beta_k N_{kj}^* + \Sigma_l \gamma_l N_{ij}$$

$$+ \varepsilon_{ij} \; ; \; i = 1,\ldots, g \; ; \; j = 1,\ldots, n$$

where N_{kj}'s are continous demographic variables, N_{lj}'s are dummy demographic variables, and ε_{ij} is the error term.

The variables with an asterisk are defined as

$$\mu^* = (\mu^\lambda - 1)/\lambda$$

$$N_{kj}^* = (N_{kj}^\lambda - 1)/\lambda$$

λ represents the transformation parameter to be estimated. If $\lambda = 1$, Eq. (9) is a linear form. If $\lambda = 0$, Eq. (9) is a semi-log form.

The income elasticity in (9) can be computed as

$$E_\mu = \alpha_1(\mu^\lambda / P_{ij}X_{ij})$$

A similar formula can be used for computing the elasticities with respect to other demographic variables.

Under the assumption that ε_i is normally and independently distributed with mean zero and variance^2, Box and Cox [5] show that, for a given λ, the maximized log likelihood of Eq. (9) is

$$(10) \quad L_{max}(\lambda) = (\tfrac{n}{2})\log \hat{\sigma}^2(\lambda) + (\lambda - 1)\Sigma_j \log (P_{ij}X_{ij})$$

where $\hat{\sigma}^2(\lambda)$ is the estimated error variance of the regression. Maximization of (10) over the entire parameter space requires selection of alternative values of λ over a reasonable range and find the λ that maximizes (10). Box and Cox suggested using the likelihood ratio method to obtain the confidence internal for λ. Specifically, $2[L_{max}(\hat{\lambda}) - L_{max}(\lambda)]$ is approximately distributed as χ^2 with one degree

Table 1. Average Annual Expenditure of Urban Consumer Units, 1980-81

Expenditure Group	Expenditure	%
Food at Home	$2,411	14
Food Away from Home	813	5
Clothing	935	6
Housing	5,051	30
Transportation	3,454	20
Health Care Services	746	4
Entertainment	762	4
Education	219	1
Reading	117	1
Personal Care Services	158	1
Alcoholic Beverages	280	2
Tobacco	175	1
Insurance and Pensions	1,264	7
Miscellaneous	760	4
Total Expenditures	17,144	100

of freedom. Therefore, the $(1 - \alpha)$ confidence interval for λ can be obtained by finding the value of λ on either side of $\hat{\lambda}$ such that $L_{max}(\hat{\lambda}) - L_{max}(\lambda) = 1/2\ X_1^2(\alpha)$

DATA SOURCES

The empirical study is based on the 1980-81 Consumer Expenditure (CE) Survey conducted by the U.S. Bureau of Labor Statistics [13]. The CE survey consists of a quarterly interview survey and a diary survey. Only the interview survey data are used in the study.

The CE public use tape provides expenditure data on a quarterly basis. These quarterly data are aggregated to obtain the annual expenditure data used in this study. Note that the 1980-81 survey used household panels with 20% of the households rotating in and out each quarter. Therefore, four continuous quarters of data were not available for all households. Consequently, households without four continuous quarters of data during 1980-81 were excluded. The sample size for model estimation includes 3,005 households.

Table 1 shows the average annual expenditures of urban consumer units for 14 aggregate components estimated from the 1980-81 surveys. Among the most important expenditure categories are food at home, housing and transportation. Since there are many households with a zero expenditure for the minor expenditure components like education, reading and personal care services, the last seven expenditure components have been further aggregated into a category of "others". Therefore, there are eight expenditure components for empirical estimation.

ECONOMETRIC VARIABLES

The econometric variables include a set of continuous variables and a set of dummy variables. There are four continuous variables defined as:

Y_i = Expenditure of i th group (Dollars),
INCOME = Total expenditure (Dollars),
SIZE = Household size (Number of Persons),
AGE = Age of reference person.

The group expenditure, Y_i, is the dependent variable in the Engel function. Income is measured by the total expenditure. SIZE and AGE are two demographic variables.

Table 2 identifies the other demographic variables and household characteristics which are dummy variables. In each set of dummy variables, one is excluded from the regression equation. The excluded variable is identified in Table 4 as the base variable. One objective of this study is to estimate the impacts of these demographic and household characteristics on the expenditure pattern.

REGRESSION RESULTS

The empirical application entails the estimation of Eq. (9) for the eight expenditure groups. The maximum likelihood estimation is implemented by selecting a set of discrete values of λ. Since one specific purpose of the study is to validate the linear and semi-log form the special cases of the Box-Cox general form, we specify λ to be within -2.0 to 2.0 with an increment of 0.1. In addition to the Box-Cox general form, we also estimated the modified Engel function with linear and semi-log. The linear form corresponds to Eq. (8) while the semi-log form is derived from taking the logarithmic transformation for the continuous explanatory variables such as total expenditure in the right hand side of Eq. (8).

The regression results for the linear and semi-log form are not presented here. It was found that both sets of regression results are very similar in terms of the sign of the coefficients and the level of statistical significance. The linear form, however, has lower mean squared error (MSE) and higher R^2 for all expenditure groups. Therefore, the linear form has a better fit than the semi-log form, a somewhat surprising result as the opposite result has typically been found in previous studies.

Table 3 presents the regression results for the Box-Cox general form. R^2's are either higher or the same as those of the linear form. Improvements in fit are observed in the equations for food away from home, clothing, transportation, and others. In order to validate the linear and semi-log forms from the Box-Cox transformation, we can compute the 95% confidence interval for the estimated λ. These confidence intervals are also presented in Table 3. As one can see, the hypotheses of λ to be equal to 1.00 (for linear form) or zero (for semilog form) are all rejected at the 0.05 significance level for all expenditure groups except housing. It is interesting to note that the estimated λ's in most cases are closer to 1.00 than to zero. The largest deviation from either linear or semi-log forms occurred for the food away from home group.

Table 2. Definition of Dummy Variables

Strata	Symbol[a]	Definition
Race/ Ethnicity	HISP	Spanish
	AFRO	Afro-Blacks
	NAFR	Non Afro-Blacks
	OTHR	Others
	(WHIT)	Whites
Region	NE	Northeast region
	NC	North Central region
	WE	West region
	(SO)	South region
Marital Status	MAR	Married
	(SIN)	Single, Widowed, Divorced and Separated
Housing Tenure	RENT	Renters
	HOMT	Homeowner with Mortgage
	HNMT	Homeowner without Mortgage
	(OCUP)	Occupied without Cash Payment
Education	HIGH	High School (1-4 yrs.)
	COLL	College (1-4 yrs. and more)
	(ELEM)	Elementary School or less
Occupation	CLER	Clerical
	PROF	Professional & Managerial
	BWOR	Blue-Collar Workers
	(UNMP)	Unemployed, retired, students, etc.
Household Head	FHEAD	Female household head
	(MHEAD)	Male household head

[a]The variables in parentheses are "base" variables which are excluded in the regression equations.

Specifically, the estimated λ is 1.90 and the 95% confidence interval is (1.75, 2.05) which is far from either 1.00 or zero. These results suggest that the appropriate functional form for food away from home should be neither linear, nor semi-log. However, for the entire set of 8 expenditure groups, the statistical results would suggest that the linear form provides a much better approximation than semi-log for the Engel relationships based on statistical fit.

The results in Table 3 show that demographic variables and household characteristics are important determinants for household expenditure patterns. Let us summarize here the most interesting results. First, age has a significant positive impact on the expenditures for food at home, and health care services while it has a significant negative impact on the expenditures for housing and entertainment. Second, household size has a significant positive impact on food at home but a negative impact on food away from home, housing, and transportation. Third, the results show that the Afro-American households spend significantly less for food at home and significantly more for clothing than white households. Fourth, households with college education spent significantly less for food at home and transportation, but more for food away from home, and clothing than those with only elementary and high school education. Finally, households in the Northeast tend to spend more for food at home, housing, and health care services

TABLE 3. Regression Results, Box-Cox Form

Variables/ Statistics	Food at Home	Food Away from Home	Housing	Clothing
CONSTANT	-342.67	448.05**	525.72	-55.50
INCOME	1.307**	0.000006**	0.27**	0.0036**
AGE	43.43**	0.0298	-20.00**	0.1018
SIZE	650.10**	-8.505**	-240.43**	11.851
HISP	116.83	-12.666	115.45	-67.53
AFRO	-163.39*	-117.977	170.96	193.55**
NAFR	-47.10	-185.30	235.09	144.73
OTHR	-29.87	-10.11	83.91	-40.27
NE	143.84**	64.06	350.16**	26.21
NC	-178.56**	43.01	239.73*	46.17
WE	-29.18	48.01	277.30*	-109.69**
MAR	105.21	-94.06*	-39.99	12.45
HIGH	17.86	157.31**	-286.93*	64.56
COLL	-169.92**	254.92**	125.79	125.31*
HOMT	-492.43**	-167.34	2042.94**	-122.59
HNMT	-514.78**	-87.06	1038.18**	-93.89
RENT	-414.06**	-184.77	1526.74**	-30.90
FHEAD	51.20	-196.02**	180.31	103.30*
CLER	71.63	147.73	-430.66**	150.36**
PROF	63.87	142.48**	-58.41	206.00**
BWOR	133.75*	41.21	-676.64**	58.44
λ	0.70	1.90	1.00	1.30
95% CI	(.58, .82)	(1.75, 2.05)	(.95, 1.05)	(1.18, 1.42)
$L_{max}(\lambda)$	-20690.12	-20069.84	-23192.30	-20121.09
R^2	.56	.45	.62	.48

Variables/ Statistics	Transportation	Health Care Services	Entertainment	Others
CONSTANT	-5186.40**	194.06	516.55**	-71.00
INCOME	42.61**	0.0006**	0.009**	0.002**
AGE	-30.83	2.651**	-2.653**	0.743
SIZE	-292.65**	-0.00	-4.44	-8.77
HISP	449.66	-38.44	-173.50	-392.50**
AFRO	-17.39	-38.84	-114.28	44.95
NAFR	-32.90	-128.30	1.29	128.83
OTHR	-50.25	4.58	72.87	-43.07
NE	-447.22**	153.07**	-18.27	22.65
NC	-10.57	-170.88**	122.60*	-102.99
WE	-115.03	-143.73**	97.46	-20.14
MAR	-99.03	179.33**	-218.70**	82.27
HIGH	-21.78	40.63	6.95	-8.68
COLL	-520.96**	89.10	90.00	16.29
HOMT	-799.64*	-181.29	-244.76	-57.30
HNMT	-228.19	-13.32	-212.18	58.26
RENT	-495.73	-150.71	-260.72	-41.68
FHEAD	14.19	-36.11	-209.61**	97.74
CLER	-172.88	-60.56	-123.71	359.89**
PROF	-394.97**	-136.52**	-30.58	215.72*
BWOR	412.76**	-211.67**	-40.56	222.68*
λ	0.50	1.40	1.20	1.40
95% CI	.47, .78)	(1.2, 1.7)	(1.03, 1.30)	(1.31, 1.55)
$L_{max}(\lambda)$	23628.90	-20143.58	-21142.67	-22176.55
R^2	.49	.24	.29	.36

*Significant at 5%
**Significant at 1%

but less for transportation than those in other regions.

The equation with the best fit is the one for housing (R^2 = 0.62). In this equation, most of the demographic and household characteristic variables are statistically significant except those for ethnicity/race, marriage, and sex of household head.

ESTIMATED INCOME ELASTICITIES

Table 4 presents the elasticity estimates using alternative functional forms. Since income elasticities are dependent upon both income (total expenditure) and individual expenditure for linear and Box-Cox, and are dependent upon individual expenditure for semi-log, it is important to know how these elasticities change at different levels of income. The three sets of

Table 4. Estimated Income Elasticities[a]

Expenditure Group	Total Sample[b]			Low Income Households[c]			High Income Households[d]		
	Box-Cox	Linear	Semi-log	Box-Cox	Linear	Semi-log	Box-Cox	Linear	Semi-log
Food at Home	0.44	0.40	0.42	0.42	0.29	0.70	0.46	0.50	0.30
Food away from home	0.79	1.40	1.32	0.46	1.55	3.41	1.18	1.20	0.69
Housing	0.86	0.88	0.90	0.78	0.78	1.77	0.90	0.92	0.54
Clothing	1.11	1.30	1.18	0.99	1.48	2.95	1.20	1.40	0.63
Transportation	1.44	1.22	1.35	2.98	1.75	4.15	0.98	1.09	0.70
Health Care Service	0.66	0.81	0.83	0.35	0.58	1.31	0.99	0.97	0.57
Entertainment	1.29	1.41	1.35	1.57	2.02	4.42	1.19	1.17	0.67
Others	1.14	1.30	1.23	0.88	1.44	3.00	1.26	1.21	0.66
Adding-up Condition	0.97	1.00	1.01	1.06	1.00	2.24	0.99	1.00	0.59

[a]Estimated at sample means
[b]The sample mean of income (total expenditure) is $17,755.
[c]Total expenditures in bottom 25 percentile. The sample mean of income is $8,066.
[d]Total expenditures in upper 75 percentile. The sample mean of income is $30,764.

income elasticities presented in Table 4 are for (1) total sample, (2) low income households (total expenditures in bottom 25 percentile), and (3) high income households (total expenditures in upper 75 percentile).

The results are striking. First of all, the estimated income elasticities are very similar when computed at the means of the total sample. There is only one exception and that is the estimates for the food away from home. Specifically, the estimate of 0.79 from the Box-Cox general form is considerably smaller than those obtained from either linear or semi-log. The Box-Cox form shows that food away from home is a necessity while both linear and semi-log show it to be a luxury. This is a significant finding because it differs so greatly from those obtained from linear and semi-log forms and those obtained from previous studies. Our estimates of income elasticities for food at home, as computed at the total sample means, are very similar among the three functional forms. They range from 0.40 to 0.44, implying a very inelastic demand with respect to income. These estimates of elasticities are considerably smaller than those for subgroups of food items estimated by Kokoski [9], using the 1980-81 CE-diary survey data. Therefore, this study suggests smaller income elasticities for both food at home and food away from home than suggested by estimates from previous studies.

The adding-up condition is used to validate the econometric results. It was found that the condition is almost satisfied with both Box-Cox and the semi-log form when elasticities are computed using the total sample means. (The linear form satisfies this condition automatically). Therefore, if one looks at only the first set of elasticity estimates (for the total sample), the results are not very discriminating, therefore, it is not easy to make a choice.

The results show significant differences when income elasticities are estimated for different income groups. Consider, for example, the food away from home. The Box-Cox form shows that the income elasticity is 0.46 for low income households and is 1.18 (elastic) for high income households. On the other hand, both linear and semi-log show that income elasticities decline from low income households to high income households. Furthermore, the semi-log form produces an inelastic demand (0.69) for high income households. The results from linear and semi-log may be perceived to be a more logical Engel relationship than those of Box-Cox when looking at food away from home in isolation. However, the validity of these estimates must be evaluated from the standpoint of a complete demand system.

Further examination of the complete sets of income elasticities reveals several interesting findings. First, the semi-log form consistently produces a smaller income elasticity for low income households as compared to high income households for all expenditure groups. Second, both the Box-Cox and the linear form show mixed changes, i.e. some elasticities became larger while others became smaller. These patterns of change are, of course, more consistent with economic theory. If we hold the assumption that consumers always allocate their total budget among the goods and services (adding-up condition), then the theory requires that decreases in income elasticities for some groups must be compensated by increases in income elasticities for others. Consequently, when the adding-up condition was examined for the semi-log form it was found that the result was not consistent with the theoretical target of unity. However, this condition is satisfied for the Box-Cox form (again the linear form automatically satisfies this condition). Therefore, one must conclude that the overall results obtained from the semilog form are not consistent with economic theory.

CONCLUSIONS

This study shows that there is a one-to-one correspondence between an Engel function and a complete demand function. As such, a linear Engel function may be derived from either LES or the CES demand systems. Both are theoretically plausible. The double-log and semi-log forms cannot be derived from these well-known complete demand systems. This study also demonstrated that the inclusion of the demographic variables in the Engel function in an additive fashion can be theoretically justified if (1) the original form is linear and (2) the demographic translating is applied. For both the double-log and semi-log forms, the additive inclusion of the demographic variables should be considered as ad hoc.

The Box-Cox general form was used to validate the linear form which is theoretically plausible and the semi-log which is empirically popular. The modified Engel function was estimated for a complete system of eight aggregate expenditure components, using the 1980-81 Consumer Expenditure Survey data.

The econometric results show that the Box-Cox general form (without restriction) satisfied the adding-up condition while the semi-log form violated this theoretical condition. (The linear form automatically satisfied the condition.) The study shows that the appropriateness of the functional form should be evaluated from the standpoint of a complete demand system. Overall, the econometric results show that both the linear form and Box-Cox transformation produce reasonable sets of estimates of income elasticities, and both are clearly superior to the semi-log form.

The results show that demographic variables and household characteristics are important determinants of household budget allocation.

The Box-Cox general form produces an estimate of income elasticity considerably less than unity for food away from home. Furthermore, the results show that as income increases, the income elasticity for food away from home increases rather than decreases. The validity of the estimated patterns of income elasticities for this and other expenditure categories definitely deserves further investigation.

REFERENCES

1. Adrian, John, and Daniel, Raymond. "Impact of Socioeconomic Factors on Consumption of Selected Food Nutrients in the United States." American Journal of Agricultural Economics, 58 (February, 1976): 31-38.

2. Basiotis, P. P., Brown, M., Johnson, S. R., and Morgan, K. J. "Nutrient Availability, Food Costs, and Food Stamp." American Journal of Agricultural Economics, 65 (November 1983): 685-693.

3. Blancitorti, Laura, Green, Richard, and Sylvia Lane, "Income and Expenditure for Relatively More Versus Relatively Less Nutritious Food Over the Life Cycle," American Journal of Agricultural Economics, 63 (May 1981): 256-260.

4. Blaylock, James R., and Smallwood, David M. U. S. Demand for Food: Household Expenditures, Demographics, and Projections. U. S. Department of Agriculture, Economic Research Service, Technical Bulletin No. 1713, February 1986.

5. Box, G. E. P., and Cox, D. R. "An Analysis of Transformation." Journal of The Royal Statistical Society, Series B, 26 (1964): 211-243.

6. Chang, Hui-Shyong. "Functional Forms and the Demand for Meat in the United States." The Review of Economics and Statistics, 59 (August, 1977): 355-359.

7. Davis, Carlton G. "Linkages Between Socioeconomic Characteristics, Food Expenditure Patterns, and Nutritional Status of Low Income Household: A Critical Review." American Journal of Agricultural Economics, 64 (December 1982): 1017-1025.

8. Kakwani, N. C. and Podder, N. "Efficient Estimation of the Lorenz Curve and Associated Inequity Measures from Grouped Observations." Econometrica, 44 (January 1976): 137-148.

9. Kokoski, Mary F. "An Empirical Analysis of Intertemporal and Demographic Variations in Consumer Preferences." American Journal of Agricultural Economics, 68 (November 1986): 894-907.

10. Pollak, Robert A., and Wales, Terence J. "Comparison of the Quadratic Expenditure System and Translog Demand Systems with Alternative Specifications of Demographic Effects." Econometrica, 48 (April 1980): 595-612.

11. Pollak, Robert A., and Wales, Terence J. "Demographic Variables in Demand Analysis." Econometrica, 49 (November 1981): 1533-1551.

12. Prais, S. J., and Houthakker, A. S. The Analysis of Family Budget. Cambridge: Cambridge University Press. 1955.

13. U. S. Bureau of Labor Statistics. Consumer Expenditure Survey: Interview Survey, 1980-81. Bull. No. 2225, Washington, D.C., April 1985.

HOUSEHOLD EXPENDITURES ON CLOTHING AND TEXTILES

Pamela S. Norum, University of Missouri-Columbia[1]

ABSTRACT

This study was a preliminary analysis of household quarterly expenditures on apparel. The data were from the first quarter of the 1982 Consumer Expenditure Survey. Two different equations were estimated. Additional analysis needs to be done that examines expenditure over a one-year period, and in which the quarters correspond to the retail calendar.

INTRODUCTION

Apparel items, along with food and shelter, are considered to be basic consumer goods. Recent estimates indicate that Americans spent 147.3 billion dollars on apparel and shoes in 1985. On a per capita basis this represents $617.00 per person. In both real and current dollars expenditures for apparel and shoes have increased over time (3). Expenditures on, and consumer demand for apparel is of interest to consumer economists, apparel marketers, and other professionals interested in household consumption behavior.

Since the prices of other goods have risen faster than apparel prices, the percentage of the total family budget allocated to apparel (current dollars) has decreased. Expenditures on clothing and shoes accounted for 5.9 percent of personal consumption expenditures in 1985 as compared with 7.2 percent ten years earlier (3). Apparel goods suppliers are experiencing greater competition for the consumer dollar, and could benefit from an increased understanding of the factors influencing consumer clothing demand. The retail environment has been further affected by recent changes in both consumer behavior and sources of supply, as well as the entrance of new types of competition into the marketplace. Retailers, however, are not the only actors in the market who are facing changing conditions. Apparel manufacturers may "face extinction because of their failure to assume a marketing orientation," and their lack of understanding of the consumer (1). Knowledge of variations in expenditures for apparel items, due to differences in family characteristics, is an aid in understanding consumer demand.

Information regarding household expenditures on apparel and related services would assist consumer educators and extension specialists in evaluating family clothing needs and developing guidelines for specific family and income structures. In light of the increase in the prices of other goods, compared to clothing, there may be fewer dollars in the family budget for clothing.

Although data on the purchase and use of apparel is important to producers, retailers and professionals who work directly with consumers, research in this area has been limited. The purpose of the proposed research is to examine the effects of various socio-economic and demographic characteristics on household expenditures for apparel. This will be accomplished by the empirical estimation of a variation of the standard Engel function. The Engel function expresses household expenditures on any good or service as a function of household income (5).

PREVIOUS RESEARCH

There is a large body of work in the area of consumption and demand analysis. There has been, however, relatively little economic research done in the apparel area. Early clothing and textile researchers developed descriptive clothing budgets which lacked statistical analysis (2, 10). Increasingly, multivariate statistical techniques have been used to analyze clothing expenditures. The Consumer Expenditure Survey (CES) has been a major source of data for the analysis of clothing demand. Erickson (6), as well as Lee and Phillips (9), used the 1960 Consumer Expenditure Survey to examine clothing consumption. Erickson analyzed the clothing expenditures for individual family members, while Lee and Phillips investigated differences in farm and nonfarm consumption patterns. Dardis et al. (4) used the 1972-73 CES to examine the effect of selected socio-economic and demographic variables on annual expenditures for clothing and clothing services.

Clothing studies using data other than the CES have been conducted by Hager and Bryant (8) and Frisbee (7). Bryant and Hager investigated the winter quarter purchases of new clothing in 1970, 1971 and 1972 by participants in the Rural Income Maintenance Experiment (RIME). Frisbee used the 1978 Statistics Canada Family Expenditure Survey to estimate the relationship between household characteristics and apparel expenditures.

One of the major limitations of all the studies is that the data used for analysis can no longer provide a current view of clothing consumption patterns. Although each of the studies is useful in providing insight regarding possible approaches for analyzing clothing expenditures, the results do not reflect current clothing consumption patterns. The Frisbee study, although relatively current, is most useful for understanding consumption behavior in Canada, rather than in the U. S. Additionally, the generalizability of the results from the Bryant and Hager study are limited since they used a rather unique sample.

[1]Assistant Professor of Clothing and Textiles

PROCEDURE

The data used in the analysis were from the first quarter of the 1982 Consumer Expenditure Survey (CES) collected by the Bureau of Labor Statistics. The CES is the most recent and comprehensive source of consumer expenditure data. Detailed information on household expenditures, as well as socio-economic and demographic characteristics, were collected from a national sample of over 5,000 households.

The empirical relationship between expenditures and household characteristics (the modified Engel function) can be expressed as:

$$E_j = b_0 + b_1X_1 + \ldots b_nX_n$$

E_j = the dependent variable, expenditures on good j
b_i = the estimated coefficient for variable X_i (i = 1 to n)
X_i = the independent variable, a household characteristic (i = 1 to n)

Two different equations for this fall/winter quarter were estimated. The first equation (Equation I) was a duplication of the equation estimated by Dardis et al., while the second equation (Equation II) was a modification of the first that incorporated variables, or specifications of variables used by other researchers. As with previous research, ordinary least squares was used to estimate the expenditure equations.

Table 1 presents summary statistics for the sample. The average quarterly income was $3943.66. The average age of the household head was 45 years, while the average family size was 2.7. Seventy-five percent of the sample had at least a high school education. Slightly more than half of the household heads were employed in white collar jobs, while approximately one-third of the spouses were employed in white collar jobs. Approximately 75 percent of the sample was comprised of households headed by a white person.

TABLE 1. Descriptive Statistics of the Sample

VARIABLE	MEAN VALUE
Income	$3943.66
Age of head	45 years
Family size	2.7
Apparel expenditures	$253
VARIABLE	**FREQUENCY**
Education:	
Elementary	505
Some high school	518
High school	1214
Some college	989
College graduate	879
White collar job:	
Head	2104
Spouse	1093
Region:	
Northeast	901
Northcentral	1126
South	1152
West	926
Race:	
White	3588
Non-white	517

RESULTS

Equation I was estimated so that the results of this study could be compared with those of the most recent research on clothing expenditures that used the CES (4). This comparison is expected to provide insight regarding changes in the effect of various socio-economic and demographic variables on clothing expenditures over time. The dependent variable in Equation I was the sum of expenditures on apparel, including accessories, and apparel services. The independent variables included quarterly after tax income, family size, age of head, presence of children less than 6, education of head, occupation of head, region, employment status of the spouse and race.

In comparing the equations from the two time periods, it should be noted that the Dardis results were based on annual data, while the current results are based on quarterly data. Location (urban/rural) was not included in the present analysis since the 1980-81 CES was restricted to urban households. In each case, expenditures, income and family size were entered into the equation in natural log form, while the remaining variables were entered as dummy variables. All of the coefficients for the dummy variables were transformed by taking the anti-log of the coefficient as done by Dardis et al. The results are presented in Tables 2 and 3.

TABLE 2. Regression Coefficients for Income and Family Size, Logarithmic Equations, 1982 CES and 1972 CES Results[1]

Variable	B (1982)	B (1972)
Income	0.17*	0.617*
	(0.019)	(0.018)
Family size	0.544*	0.111*
	(0.05)	

1 = standard errors in parentheses
* = significant at .01

There is consistency with respect to the significance and magnitude for selected variables between 1972 and 1982. In each equation, income, family size, presence of children less than six, region, and selected age, educational, and occupational categories were significant. The interesting comparison between the equations occurs when

TABLE 3. Transformed Regression Coefficients for Dummy Variables[1], Logarithmic Equations, 1982 CES and 1972 CES Results

Characteristic	1982 Coefficient	1972 Coefficient
Age of household head (65 or older)		
Less than 25	122*	169*
25-34	123*	156*
35-44	134*	146*
45-54	112	148*
55-64	108	124*
Married (not married)	106	99
Youngest child less than six (other households)	77*	89*
Education (high school graduate)		
Some grade school	85*	82*
Some high school	85*	93*
Some college	120*	112*
College graduate	140*	114*
Occupation (craftsmen)		
Self-employed	101	120*
Salaried professionals	110	105
Salaried manager	134*	129*
Clerical workers	115	115*
Sales workers	115	125*
Operators	90	93*
Unskilled labor	92	92
Service workers	101	102
Not working	83*	78*
Retired	74*	80*
Region (West)		
Northeast	113*	107*
Northcentral	105	100
South	99	105
Working wife (non-working wife)	111	111*
White (non-white)	113	72*
Rural location (urban location)	--	84*

[1]The coefficient for the omitted category, which is given in parentheses, has a value of 100.

*Coefficient significant at 0.05 level.

there is a difference in the significance or magnitude of the coefficients.

In each equation, income and family size are significant; however, the magnitudes are markedly different. Although in each case clothing expenditures are income inelastic, the magnitude was .617 with the 1972 data, and .17 with the 1982 data. A percent change in income resulted in a much smaller percentage change in apparel expenditures in 1982. On the other hand, a one percent increase in family size resulted in a .54 percent change in apparel expenditures in 1982 compared to a .11 percent change in 1972.

The results for age of head differ somewhat between the two time periods. In the most recent analysis, there was no significant difference in apparel expenditures between households headed by a person 45-64 years old and households headed by someone over the age of 65. In 1972, households with heads ages 45-54 (55-64) spent 48 (24) percent more on apparel than households headed by a person over the age of 65. Also, for the age categories that were statistically significant in both years, the magnitudes of the coefficients were much larger in 1972 (69 percent vs. 22 percent; 56 percent vs. 23 percent and 46 percent vs. 34 percent).

Households with a child less than six spent significantly less on apparel than households without a child less than age six in both years. In 1972 the households with a younger child spent 11 percent less, whereas in 1982 they spent 23 percent less.

All of the educational categories were significant in both years. There were some interesting differences with respect to their magnitude. Those households in which the head had some high school education spent 8 percent less than high school graduates in 1972 and 15 percent less in 1982. Those with a college education spent 14 percent more in 1972 and 40 percent more in 1982 than households with a high school educated head.

Another interesting difference between the two studies was that although the magnitude of the coefficient for spouse employment was the same in both years, it was significant only in 1972.

The difference in the coefficient on the race variable also indicated differences over the 10-year period. In 1972 non-blacks spent 28 percent less on apparel than did blacks. In 1982, whites spent 13 percent more on apparel than did non-blacks. Some of the disparity may be due to the fact that the treatment of races other than white or black was handled differently in each study. However, given the limited number of these other races, it is not likely that this could explain all of the difference.

The comparison of the 1972 and 1982 equations indicates stability over time with respect to the effects of some variables, and marked changes in others. Before definite conclusions can be drawn, however, it would be necessary to extend the quarterly analysis over a period of a year, or to compare total annual expenditures.

The results for Equation II are presented in Table 4. This equation was estimated in linear form. The independent variables included in the equation are quarterly after tax income, age of head, the number of household members in various age and sex categories, the sex of head, education of the head, region, race, whether or not the head or spouse were employed in white collar jobs, and whether or not there were expenditures on sewing items.

The statistically significant variables were income, age of head, the number of children less than 2, the number of females ages 2-15 and over 16, the number of males over 16, all educational categories except elementary education, whether or not the head or spouse worked in white collar jobs, whether or not the household was located in

TABLE 4. Parameters Estimates for Household Apparel Expenditures, Linear Equation

Variable	B	S.E. of B
Intercept	11.8	23.6
Income	0.02*	0.0016
Age of head	-0.75**	0.31
Children < 2	-54.8*	15.29
Females 2-15	44.5*	7.9
Males 2-15	9.04	7.29
Females over 16	68.7*	8.2
Males over 16	22.4*	8.4
Sex of head	0.38	12.6
Eleduc	-25	16.3
Somehs	-28***	15.3
Somecoll	50.5*	12.6
Collgrad	105.4*	13.5
Whcollar	41.7*	10.1
Whcollarsp	40.7*	11.6
NE	43.2*	13.7
MW	9.4	12.95
South	31.2**	12.9
White	8.5	13.9
Sew	22.3***	12.8

* = Significant at = .01
** = Significant at = .05
*** = Significant at = .10
R^2 = .18
N = 4032

the northeast or south, and whether or not the household had any sewing expenditures.

The coefficient on quarterly income was .02. This indicates that, on average and holding other factors constant, a $1.00 change in quarterly income resulted in a two cent change in quarterly apparel expenditures in the same direction. The magnitude of the coefficient falls within the range of values (.007 to .05) found by Hager and Bryant for the effect of various types of income (negative income tax payment, wife and other family income) on winter clothing expenditures in the RIME sample.

The coefficient on age indicates that as the age of household head increases, clothing expenditures decline. This result coincides with the results of previous studies (4, 7).

With respect to the age and sex of family members, the presence of children less than two has a negative effect on apparel expenditures. On average, and ceteris paribus, the addition of an infant to a household results in a decline in quarterly apparel expenditures by $54.80. This may reflect the medical expenses and life insurance that are brought about by the addition of a child under the age of two. Infants do not demand as much in clothing relative to other goods and services since clothing is a common baby gift and there is a well developed second hand market for children's clothing. Females ages 2-15 and over 16 had a significant positive effect on apparel expenditures. The coefficients for males are relatively smaller or insignificant, supporting the idea that women and girls consume clothing to a greater extent than men and boys. (2)

Education appears to significantly affect apparel expenditures. Households in which the head had less than a high school education spent less on clothing, on average and holding other factors constant, compared to households with a high school educated head. Just the opposite was true for households with a head that had more than a high school education. These results coincide with previous research (4). Since job type is being held constant, education may reflect differences in tastes and preferences.

Occupation was proxied by whether or not a person (head or spouse) worked in a white collar job. In each case, people who worked in white collar jobs spend a little over $40 more on clothes during the quarter than those people who weren't in white collar jobs.

In terms of region, households located in the northeast and south spent significantly more on apparel than did households located in the west. The result for the northeast is not surprising given the difference in the fall/winter climate between the northeast and west. However, the result for the south is somewhat puzzling.

Finally, households in which there was some positive expenditure on sewing items spent more, on average and ceteris paribus, than households with zero sewing expenditures. This result does not support the notion that households in which sewing occurs will substitute home-made clothing for store-bought clothing. It is possible that sewing expenditures is proxying clothing interest.

CONCLUSIONS

This research has investigated the influence of selected socio-economic and demographic variables on apparel expenditures. Data from the first quarter of the 1982 Consumer Expenditure Survey were used for the analysis. The results of this study are useful in providing a general overview of current U. S. clothing demand. However, to get a fuller view of current clothing consumption, additional analysis is necessary. First, Winakor (11) has pointed out the need for measuring clothing acquisition and discard over a one-year period. This could be done using total annual expenditures or quarterly expenditures. For clothing, which is greatly influenced by seasonality, both climatic and social, as well as retailing efforts, quarterly analysis seems appropriate. Quarterly analysis would allow variations in the effect of individual variables to be measured. Additionally, selection of months within a quarter that correspond to the retail cycle would provide the most useful information to retailers.

Second, future research of a more specific nature would be useful. For example, the allocation of clothing dollars to various types of clothing (e.g., suits, sportswear) or for different family

members should be done. Additionally, analysis of expenditures by growing consumer segments, such as the elderly, is necessary. Finally, given the growth in the services industry, expenditures on apparel services would be of interest.

Besides future avenues for empirical research, other issues that should be addressed include the need for additional theoretical development. Apparel differs from other consumer goods in a variety of ways. It is a semi-durable good, and consequently consumers generally have an inventory of clothing that may influence their clothing purchasing patterns. Additionally, the fashion element embedded in clothing should be recognized. The inclusion of psychographic variables in explaining apparel purchases may prove fruitful.

REFERENCES

1. American Marketing Association. "Apparel Manufacturers Face Extinction Because of Their Failure to Assume a Marketing Orientation, Consultant Says." Marketing News 19(2):1-3.

2. Britton, Virginia. "Clothing Quantity Budgets." Family Economics Review, Fall 1974, pp. 3-7.

3. Courtless, Foan C. "Recent Trends in Clothing and Textiles." Family Economics Review, 1986, No. 2, pp. 17-26.

4. Dardis, Rachel, Derrick, Frederick, and Lehfeld, Alane. "Clothing in the United States: A Cross-Sectional Analysis." Home Economics Research Journal 10(2): 212-222.

5. Deaton, Angus, and Muellbauer, John. Economics and Consumer Behavior. New York: Cambridge University Press, 1980.

6. Erickson, A. "Clothing the Urban American Family: How Much for Whom?" Monthly Labor Review 91:14-19.

7. Frisbee, William R. "An Economic Analysis of Household Clothing Expenditures in Canada." Canadian Home Economics Journal, January 1986.

8. Hager, Christine J., and Bryant, W. Keith. "Clothing Expenditures of Low Income Rural Families." Journal of Consumer Affairs 11(2):127-132.

9. Lee, Feng-Yoo, and Phillips, K. E. "Differences in Consumption Patterns of Farm and Nonfarm Households in the United States." American Journal of Agricultural Economics 53:573-582.

10. Winakor, Geitel, et al. "Clothing Budgets." Journal of Home Economics 63:354-362.

11. Winakor, Geitel. "The Process of Clothing Consumption." Journal of Home Economics 61:629-634.

INCOME REPORTING IN THE U.S. CONSUMER EXPENDITURE SURVEY

Thesia I. Garner[1], Bureau of Labor Statistics, U.S. Department of Labor

ABSTRACT

The purpose of this study was to examine the relationship between particular socioeconomic characteristics and the probability that a household (consumer unit) reports income information. Socioeconomic variables included in the model were the age, race, sex, education, and occupation of the reference person, and the housing tenure, degree of urbanization, and region of residence of the household. Binomial logit analysis was used to model the probability of income response completeness. Data from the Interview portion of the 1983 U.S. Consumer Expenditure (CE) Survey were analyzed. Results of this study have important implications for consumer researchers conducting analyses of CE Survey data.

INTRODUCTION

Data from the U.S. Consumer Expenditure (CE) Survey are frequently used by consumer researchers to study the expenditure behavior of households or consumer units.[2] The CE Survey data are ideal for this type of analysis since they are the most comprehensive information available on the expenditures, income, and corresponding socioeconomic characteristics of American consumer units. However, when data such as these are collected, incomplete responses or nonresponses frequently result. The usefulness of survey data to economic researchers is sensitive to how missing data are treated. For example, if an analysis is restricted to complete cases only and the cases are not a random subsample of the original population, restricting the analysis would involve loss of efficiency and could yield biased results.

The focus of this paper is the nonreporting of income, a pervasive problem in household surveys. For studies in which income is included as an explanatory variable, it is imperative that researchers are familiar with the patterns of income nonresponse and the underlying assumptions concerning the distribution of missing values. The purpose of this research was to identify socieconomic factors related to the probability that a consumer unit, participating in the U.S. Consumer Expenditure Survey Interview, reports income information. Binomial logit analysis was used to model the probability of income response completeness.

It was hypothesized that the probability of income response completeness is a function of various socioeconomic characteristics. The socioeconomic variables included in the model were the age, race, sex, education, and occupation of the reference person.[3] Reference person characteristics were used for the analysis since the consumer unit's characteristics are most often identified by the reference person's. Additional variables included the housing tenure, degree of urbanization, and region of residence of the consumer unit.

Results from previous studies were used as a basis for identifying the variables to include in the probability of income response model and for specifying hypothesized relationships. These studies included examinations of the probability of reporting earnings or wages and salaries for

[1]Economist in the Division of Price and Index Number Research. Any views expressed are the author's and do not reflect policies of the Bureau of Labor Statistics (BLS) or views of other BLS staff members.

[2]The term consumer unit is used throughtout the remainder of the paper rather than household, since the consumer unit is the basic reporting unit for the CE Survey. A consumer unit is defined as one of the following: (1) all members of a particular household who are related by blood, marriage, adoption, or other legal arrangements; (2) a person living alone or sharing a household with others or living as a roomer in a private home or lodging house or in permanent living quarters in a hotel or motel, but who is financially independent; or (3) two or more persons living together who pool their income to make joint expenditure decisions. Financial independence is determined by the three major expense categories: housing, food, and other living expenses. To be considered financially independent, at least two of the three major expense categories must be provided by the respondent.

In the majority of cases, there is one consumer unit per household (U.S. Department of Labor 1986).

[3]The reference person is the first member of the consumer unit mentioned by the respondent when asked to "Start with the name of the person or one of the persons who owns or rents the home." (U.S. Department of Labor 1986).

the Current Population Survey, CPS (Greenlees, Reece, and Zieschang 1982; Lillard, Smith, and Welch 1986), income imputation procedures (Little and Samuhel 1983), nonresponse rates in the Survey of Income and Program Participation, SIPP (Coder and Feldman 1984), and factors affecting survey data quality (Andrews and Herzog 1986). Previous income reporting studies focused primarily on the reporting of earnings or wages and salaries of individuals. Often the samples were restricted to the working age population and to males. An exception was provided by Coder and Feldman (1984) in their examination of income nonresponse for SIPP; these researchers produced nonresponse rates for several sources of income for individuals age 15 and older. In contrast to these previous studies, the unit of analysis for this study was the consumer unit, with all consumer unit members aged 14 and over requested to provide income information.

Results from this study can be used to provide information concerning the relationship of income nonresponse and socioeconomic factors, which are frequently used in conducting expenditure analyses. In turn, researchers can select procedures for analysis which account for the possibility that income data are not missing at random in the Survey. The ideas and results presented in this paper represent the first stage in an overall research plan to examine issues related to income nonresponse in the CE.

This paper is organized into three remaining sections: Methods and Procedures, Results, and Summary and Conclusions.

METHODS AND PROCEDURES

Data Source

The data used in this study were from the Interview portion of the 1983 U.S. Consumer Expenditure (CE) Survey. The CE Survey data are collected by the Bureau of the Census under the auspices of the BLS. The Interview sample, selected on a rotating panel basis, is targeted at 4,800 consumer units per quarter. Each quarter one-fifth of the sample is new to the survey. After being interviewed for five consecutive quarters, each panel is dropped from the survey. Detailed income data are collected during interviews two and five only. For the purpose of this study, the sample was defined as all consumer units participating in a second interview during 1983. Because consumer units living in rural areas outside Standard Metropolitan Statistical Areas (SMSA's) were not surveyed in that year, they were not part of the study sample.

Dependent Variable

Income reporting was defined in terms of the completeness of income information obtained from consumer units. The distinction between a complete and an incomplete income reporter, used in this analysis (and by BLS in its publications of CE Survey data), was based on whether the respondent provided values for various sources of income. Sources were grouped into two categories: major sources of income and other sources of income.

Major sources of income include:
- wages and salaries
- income from non-farm business partnership or professional practice
- Social Security or Railroad Retirement
- Supplemental Security Income.

Other sources of income include[4]:
- unemployment compensation
- workmen's compensation and veteran's payments including educational benefits
- public assistance or welfare receipts including income from job training grants such as Job Corps
- interest received on savings accounts or bonds
- regular income received from dividends, royalties, estates, or trusts
- income received from pensions or annuities from private companies, military, or government
- income or loss received from roomers or boarders
- income or loss received from payments from other rental units
- regular contributions received from alimony or child support and other sources combined.

A consumer unit was defined as a complete income reporter if:

(a) the reference person had a non-zero amount reported for a major source of income, any entry for a major

[4]Additional income, not included in "Other Sources of Income" for the purpose of identifying consumer units as complete or incomplete income respondents, includes money received from the care of foster children, cash scholarships and fellowships or stipends not based on earnings and the annual value of food stamps. Income from these sources, however, is added to the income received from the sources noted above to calculate a consumer unit's income before taxes. Income before taxes is the income variable used in CE Survey publications to classify consumer units.

source of income for at least one other consumer unit member was recorded, and any entry for other sources of income was recorded; or

(b) consumer unit member(s) other than the reference person had a non-zero amount reported for a major source of income, valid blanks (recorded if a consumer unit member indicated that income was not received from a particular source) were recorded for all the major sources of income for the reference person, and any entry was recorded for other sources of income; or

(c) consumer unit member(s) had a non-zero amount reported for at least one other source of income and valid blanks were recorded for all major sources of income for all members.

Given this definition, it was possible even for complete income reporters not to have provided a full accounting of all income from all sources. Consumer units with other combinations of entries to the income questions were considered to be incomplete income respondents. In the extreme case of across-the-board zero income, the response was considered invalid and thereby constituted an incomplete income report.

Independent Variables

Eight socioeconomic factors were included in the model as independent variables. Five were characteristics of the reference person in the consumer unit: age, race, sex, education, and primary occupation. Race and sex of the reference person entered the model as control variables. An age squared term was included to account for possible nonlinear effects of age. The other three variables expected to be important in the probability model were consumer unit residence characteristics: housing tenure, degree urban, and region. Definitions of all variables included in the model are presented in Table 1. All independent variables except those for age entered the model as categorical variables. Thus, it was necessary to omit one of the dummy categories for each of the variable sets before estimating the model. The omitted category for each variable set is noted in parentheses in Table 1.

Factors hypothesized to be characteristics of consumer units least likely to be complete income reporters included: older age, more education, and self-employment of the reference person; homeownership and consumer unit residence in the Northeastern or Northcentral regions of the country. No hypotheses were made concerning race, sex, and degree of urbanization.

Table 1. Definitions of Variables

Dependent Variable	
Response:	Unity if consumer unit responded to income questions which identified the unit as a complete income reporter; zero if incomplete income reporter.
Independent Variables	
Age of Reference Person:	Age in years.
Age Squared:	Age of reference person squared.
Race of Reference Person:	Unity if black; zero if white or other.
Sex of Reference Person:	Unity if male; zero if female.
Education of Reference Person	
No school:	Unity if never attended school; zero otherwise.
Elementary:	Unity if 1-8 years of schooling completed or less than a high school graduate; zero otherwise.
College:	Unity if a college graduate (4 years); zero otherwise.
Postgraduate:	Unity if more than 4 years of college completed; zero otherwise.
(Omitted category was High School: reference person is a high school graduate or has completed some years of college.)	
Primary Occupation of Reference Person	
Sales:	Unity if person received the most earnings in the past 12 months from employment in a technical, sales, or administrative support occupation; zero otherwise.
Services:	Unity if person received the most earnings in the past 12 months from employment in a services, farming, forestry, or fishing occupation or in the armed forces; zero otherwise.
Laborer:	Unity if person received the most earnings in the past 12 months from employment as an operator, fabricator, or laborer; zero otherwise.
Craft:	Unity if person received the most earnings in the past 12 months from employment in a precision production, craft, or repair occupation; zero otherwise.
Self-employed:	Unity if person received the most earnings in the past 12 months from self-employment; zero otherwise.
Retired:	Unity if person was retired in the past 12 months; zero otherwise.
Not Working or other:	Unity if person was not working in the past 12 months or did not respond to occupation question; zero otherwise.
(Omitted category was Salaried Professional or Manager: reference person received the most earnings in the past 12 months from employment in a managerial or professional specialty occupation.)	
Housing Tenure:	Unity if consumer unit owned home; zero if consumer unit rented.
Degree Urban:	Unity if consumer unit resided in the central city of a SMSA; zero if consumer unit resided inside a SMSA in other places of 50,000 or over, places less than 50,000 and other urban territories, urban places of 2,500 to 50,000 outside an urbanized area, rural non-farm areas, or on a rural farm, or outside a SMSA in urbanized areas and urban places of 2,500 to 50,000 outside an urbanized area.
Region	
Northeast:	Unity if consumer unit resided in the Northeast region; zero otherwise.
Northcentral:	Unity if consumer unit resided in the Northcentral region; zero otherwise.
West:	Unity if consumer unit resided in the West region; zero otherwise.
(Omitted category was South: consumer unit resided in the South region.)	
Constant:	Unity for all observations.

A variable could influence income response completeness independently of the value of income or the value of income could indirectly affect response completeness through its effect on the independent variable. However, no attempt was made in this study to isolate the direct effects of the independent variables on response completeness probabilities.

Estimation Procedure

The statistical analysis of the probability of complete income response was based upon a binomial logit model.[5] In this study the model under consideration was

$$P_i = \text{Prob}(Y_i = 1) = F(X_i\beta), \text{ and} \qquad (1)$$

$$1-P_i = \text{Prob}(Y_i = 0) = 1 - F(X_i\beta) \qquad (2)$$

where the $F(X_i\beta)$ is a cumulative distribution function that describes how the probabilities of complete income reporting and incomplete income reporting, respectively, are related to the socioeconomic variables. Pi is the probability that the i-th consumer unit is a complete income reporter, Xi is the vector of characteristics of the i-th consumer unit, and β is the vector of unknown parameters. The binomial logit model assumes a cumulative logistic probability distribution for the underlying function. The probability of a complete income response is defined mathematically as

$$Pi = F(Xi\beta) = \frac{1}{1+e^{-(X_i\beta)}} \qquad (3)$$

The X_i's are considered to be observations on nonstochastic variables which are independent of each other. The error terms implicit in the model are assumed to follow the Weibull or extreme value distribution and are assumed to be independent.

Logit parameter estimates were obtained through maximum likelihood estimation, using the interative Newton-Raphson optimization procedure. The computer software package (LOGIT) employed for the analysis was developed by Antos (1983).

RESULTS

In 1983, 4,611 consumer units participated in a second interview of the CE Survey. The mean values and percent distribution of independent variables used in the logit analysis are presented in Table 2.

Table 2. Mean Values and Percent Distribution of Variables in Logit Analysis

Independent Variable	Full Sample (n=4611)	Complete Income Reporter (n=4018)	Incomplete Income Reporter (n=593)
Age of Reference Person			
Age	45.91	45.59	48.03
Age squared	2428.23	2404.79	2587.07
Race of Reference Person			
Black	11.08	11.03	11.47
White and other	88.92	88.97	88.53
Sex of Reference Person			
Male	67.73	67.10	72.01
Female	32.27	32.90	27.99
Education of Reference Person			
No school	0.54	0.47	1.01
Elementary	25.53	25.91	22.93
High School	52.11	52.27	51.10
College	11.00	10.55	14.00
Postgraduate	10.82	10.80	10.96
Principal Occupation of Reference Person			
Salaried Professional or Manager	21.66	21.73	21.24
Sales	17.76	18.09	15.51
Services	8.72	9.23	5.23
Laborer	11.88	12.10	10.46
Craft	6.68	6.72	6.41
Self-employed	6.12	5.00	13.66
Retired	15.25	15.43	14.00
Not working and other	11.93	11.70	13.49
Housing Tenure			
Owned	59.57	58.19	68.97
Rented	40.43	41.81	31.03
Degree Urban			
Inside a central city	34.85	35.09	33.22
Outside a central city	65.15	64.91	66.78
Region			
Northeast	22.53	20.81	34.23
South	27.83	28.70	21.92
Northcentral	26.15	25.73	29.01
West	23.49	24.76	14.84

Results of the logit analysis are displayed in Table 3. To test the significance over all of the coefficients in the model, the likelihood

[5] There are numerous references in the literature to logit analysis. See Domencich and McFadden (1975), Judge et al. (1982), Maddala (1977), Maddala (1983), Pindyck and Rubinfeld (1981).

ratio statistic was used.[6] The resulting Chi square value was significant at the $\alpha = 0.01$ level. The null hypothesis that all of the coefficients (except the intercept) are equal to zero was rejected.

Table 3. Estimated Model Parameters and Standard Errors

Independent Variable	Estimated Parameter	Asymptotic Standard Error
Age of Reference Person		
Age	-0.0463*	0.0174
Age squared	0.0004**	0.0002
Race of Reference Person (White and other)		
Black	-0.2709	0.1538
Sex of Reference Person (Female)		
Male	-0.1257	0.1092
Education of Reference Person (High school)		
No school	-0.7433	0.4910
Elementary	0.1479	0.1220
College	-0.3684**	0.1457
Postgraduate	0.0185	0.1636
Principal Occupation of Reference Person (Salaried Professional or Manager)		
Sales	0.0180	0.1568
Services	0.3513	0.2226
Laborer	0.0644	0.1833
Craft	-0.0749	0.2123
Self-employed	-1.0839*	0.1737
Retired	0.0810	0.2048
Not working and other	-0.2467	0.1826
Housing Tenure (Rented)		
Owned	-0.2601**	0.1109
Degree Urban (Outside a central city)		
Inside a central city	-0.0831	0.1016
Region (South)		
Northeast	-0.8287*	0.1265
Northcentral	-0.3702*	0.1276
West	0.2166	0.1491
Constant	3.7568*	0.4143
Likelihood Ratio Statistic	182.253> 37.57	
Likelihood Ratio Index	0.052	

*Statistically significant at the 0.01 level.
**Statistically significant at the 0.05 level.

The likelihood ratio[7] index was calculated as a measure of goodness-of-fit of the logit model. A relatively low index value (0.052) was obtained, although it may be reasonable given the cross-sectional nature of the data.

Among the socioeconomic variables included in the probability model, age, age squared, one of the education dummy variables (college), one of the occupation variables (self-employed), housing tenure, and two of the region variables (Northeast and Northcentral) had statistically significant coefficient values. For the most part, these results were consistent with the hypotheses and with the findings of previous researchers.

The negative coefficient for the age variable indicates that as age increased, the probability of complete income response decreased. If consumer units with older reference persons have high incomes and/or receive income from numerous sources, they may be less likely to divulge their incomes or to note dollar amounts received from each income source. Or, they may react to the questionnaire design or interview procedure. However, the coefficient for age squared was positive, although small. This indicates that, for most of the sample, as age increased, the effect of age on the probability of complete income reporting (while negative) diminished. This change could occur once the reference person enters his or her retirement years. The prime earning years would be considered past and perhaps one's willingness to participate in long surveys, such as the CE, increases after retiring from the labor market.

Of all the education coefficients, only the one for college was significant in the probability model. Consumer units with college educated reference persons were significantly less likely than those in the omitted category (high school degree or some college) to be complete income

[6] The test statistic is $\chi^2 = -2(\log \text{Likelihood}_R - \log \text{LikelihoodU})$. The statistic is asymptotically Chi-square distributed with the degrees of freedom equal to the number of coefficients set equal to zero. The log likelihood function for the restricted model, represented by R, is obtained when the function is maximized with respect to the intercept only. The log likelihood of the unrestricted model, U, is obtained when the function is maximized with respect to all the coefficient estimates corresponding to the intercept and all explanatory variables.

[7] The index is

$$\rho^2 = 1 - \frac{\log \text{LikelihoodU}}{\log \text{LikelihoodR}}.$$

Generally the likelihood ratio index has an upper bound of about 0.3; it is unlikely that an index value would approach one because that could only happen if all individuals' predicted probabilities were exactly zero or one (Kinsey and Lane 1978; Pindyck and Rubinfeld 1981; Tardiff 1976).

reporters. Consumer units in the college educated group may have both higher incomes and more varied sources of income which require more detail in their income response than do the less educated. They may also place a higher value on their time and privacy (Lillard, Smith, and Welch 1986). Although the other education coefficients were not significant, the bimodal result for the signs relative to the probability of complete income response may indicate that the effects of education on the probability are nonlinear or that there is an interaction of education and other variables in the model.

Self-employment of the reference person was an important variable in the probability of complete reporting model. Consumer units with self-employed reference persons were significantly less likely to be complete income reporters than were consumer units with salaried professional or manager reference persons. Coder and Feldman (1984) and Lillard, Smith, and Welch (1986) reported similar findings for SIPP and the CPS, respectively. Researchers (Lillard, Smith, and Welch 1986) have noted that self-employed individuals are in occupations in which nonreporting is considerably higher than the average, and that these occupations share one or both of the following characteristics: "They are among the highest income occupations, or considerable ambiguity surrounds the calculation of net income from receipts and expenses for income tax purposes" (p.492).

Homeownership was negatively related to the probability of complete income response. It is likely that high income consumer units are more likely to be homeowners than to be renters, and that high income consumer units are likely to be incomplete income reporters. Thus, it is unlikely that homeownership exerts an independent effect on the propensity to respond. Whether the relationship is exerted rather through income is a subject for future research.

Coefficents for two of the region variables, Northeast and Northcentral, were statistically significant. Consumer units living in the Northeast and those living in the Northcentral regions of the country were significantly less likely than those living in the South to be complete income reporters. A ranking of the coefficients indicates that consumer units living in the Northeast were less likely to be complete income reporters than were those living in the Northcentral region. These differences may be related to regional patterns of cooperation or to differences in the value of income received or types of income received.

Tests for the combined contribution of variables represented by more than two dummy variables (as in the case of education, occupation, and region) or by more than one continuous variable (as in the case of age) were performed using the likelihood ratio statistic. Chi-square results are presented in Table 4. All of the variable sets contributed significantly to the income response model as explanatory variables.

Table 4. Chi Square Tests for Contribution of Sets of Variables

Independent Variable Set	Chi Square Statistic	Degrees of Freedom
Age of Reference Person	9.20**	2
Education of Reference Person	11.52**	4
Principal Occupation of Reference Person	59.18*	7
Region	72.74*	3

*Statistically significant at the 0.01 level.
**Statistically significant at the 0.05 level.

Since logit coefficients cannot directly provide a quantitative assessment of an independent variable's impact on the probability of an event occurring, an alternative procedure was followed. First, the estimated coefficients and consumer unit characteristics were used to compute the probability of complete income reporting according to the logistic cumulative distribution function as defined in equation (3). Characteristics at the mean were used to calculate the "average" response completeness probability. The impact of changes in the independent variables on the probability of complete income response was examined by calculating the probability at particular values of independent variables and then changing one value while holding all others constant. This procedure was employed since the majority of independent variables in the model are categorical and thus only non-marginal changes are meaningful. Only variables found to be significant in individual t-tests or those belonging to sets of variables that were significant were changed in order to observe the effect on the probabilities.

To assess the impact of changes, a representative consumer unit was identified for comparison. Charactersitics of the representative consumer unit were selected as the mean values of the continuous variables and as the categorical variables with the highest frequency of

occurrence for the sample[8]; these characteristics are noted in Table 5.

Table 5. Sample Probability Calculations for Complete and Incomplete Income Reporters

Case Description	Probabilities	
	Complete Income Reporter	Incomplete Income Reporter
Sample (at the mean)	0.8849	0.1151
Representative Consumer Unit[a]	0.9018	0.0982
Changes to the Representative Consumer Unit		
Age (increase of one standard deviation)	0.8918	0.1082
No school	0.8136	0.1864
Elementary	0.9141	0.0859
College	0.8640	0.1360
Postgraduate	0.9034	0.0966
Sales	0.9034	0.0966
Services	0.9288	0.0712
Laborer	0.9073	0.0927
Craft	0.8949	0.1051
Self-employed	0.7564	0.2436
Retired	0.9087	0.0913
Not working and other	0.8777	0.1223
Rented	0.9225	0.0775
Northeast	0.8003	0.1997
Northcentral	0.8638	0.1362
West	0.9194	0.0806

[a]In all calculations, the "representative consumer unit" was characterized as follows: age = 45.9 years; (mean of age squared = 2428.23); race = white or other; sex = male; education = high school; occupation = salaried professional or manager; housing tenure = owned; degree urban = outside a central city; region = South.

On average, consumer units participating in a second interview of the CE Survey in 1983 had a 0.8849 probability of being complete income reporters (Table 5). The representative consumer unit had a 0.9018 probability of being a complete income respondent. As expected, changes in the probabilities were in the same direction as the signs of the estimated parameters.

The most dramatic impact on the response probabilities was recorded for the self-employed. A change in the reference person's occupation from salaried professional or manager to self-employed resulted in a decreased probability of complete income response to 0.7564. Consumer units with reference persons in service occupations, but with the same other representative characteristics, were most likely to be complete income respondents.

[8]Due to the procedure used to identify characteristics, the representative consumer unit for this analysis would be expected to differ from a representative consumer unit for another analysis if definitions for the categorical variables differed. For example, the percentage of consumer units with a reference person in the high school group would decrease if reference persons with some college, but not a four-year degree, were reclassified into the college category which currently refers to four-year college graduates only.

SUMMARY AND CONCLUSIONS

The purpose of this research was to identify consumer unit socioeconomic factors which are related to income reporting completeness. This is in contrast to earlier income/earnings reporting studies in which individual characteristics were related to response probabilities. Results of the logit analysis indicate that the age, education, and occupation of the reference person, housing tenure, and region in which the consumer unit resided were significant variables in determining the probability of being a complete income reporter. For the analysis, no attempt was made to test whether the socioeconomic variables influenced income completeness through their effect on income or whether the variables independently influenced income completeness. Based on the results from this study complete income reporting consumer units and incomplete income reporting consumer units are different in terms of several socioeconomic variables. However, it would be premature to say at this time that the pattern of incomplete income reporting is related to the missing income itself.

Results from this study have important implications for research. Consumer researchers interested in using income from the CE Survey need to be aware that complete and incomplete reporters of income are different; these differences may lead to biased estimation results if not accounted for in one's estimation procedure. We at the BLS are interested in developing an income imputation procedure for the CE Survey; results from this study can serve as a basis for evaluating model-based imputation procedures which can account for cases in which the incompleteness of income is related to income itself (for references see David, et al. 1983; Fay 1986). Focusing on factors related to income report completeness is important when revising data collection procedures to improve data quality.

Future research is needed to more specifically identify factors which are related to incomplete income response, and to determine whether the pattern of incomplete reporting is related to the missing income itself. Future research could include the testing of various specifications of the probability of income response completeness model. For example, the dependent variable could be defined to represent the three types of complete income reporting situations plus

incomplete reporting categories for combinations of refusals and "don't knows." Or, the dependent variable could be defined in terms of the reporting of income by source (e.g., wages and salaries, self-employment income, retirement, and other). Or, the dependent variable could be defined in terms of the reporting of individuals within consumer units. In lieu of the reference person's characteristics, those of the survey respondent could be included as explanatory variables in the model. Additional socioeconomic variables which might be related to income response completeness include marital status of the survey respondent or reference person, number of children, number of persons within the consumer unit with an income source, work status of consumer unit members in terms of fulltime and parttime, and interaction terms for age, education, and occupation. In order to test hypotheses concerning data collection features of the survey, administrative variables could be added to the model. These might include the total number of minutes for the interview, month in which the interview was conducted, general survey nonresponse relative to specific income nonresponse, and whether records were used in answering an interviewer's questions.

For whatever reason, consumer units which are identified as incomplete income reporters are unwilling or reluctant to provide requested income information. Reasons for incomplete income response are varied and complex. Cognitive research could be used to identify reasons for incomplete responses; subsequently, revisions to the questionnaire and data collection procedures could be introduced to increase income response completeness.

In conclusion, I repeat, this study must be considered as one of exploration. However, I think that the results are sufficiently promising to warrant future research on the incompleteness of income reporting and the missingness of income in the U.S. Consumer Expenditure Survey.

REFERENCES

Andrews, Frank M., and Herzog, A.R. (1986), "The Quality of Survey Data as Related to Age of Respondent," Journal of the American Statistical Association, 81 (394), 403-410.

Antos, Joseph (1983), LOGIT package version 3/83, available at the Bureau of Labor Statistics.

Coder, John F., and Feldman, Angela M. (1984), "Early Indications of Item Nonresponse on the Survey of Income and Program Participation," American Statistical Association, 1984 Proceedings of the Section on Survey Research Methods.

David, Martin H., Little, Roderick, Samuhel, Michael, and Triest, Robert (1983), "Imputation Models Based on the Propensity to Respond," American Statistical Association, 1983 Proceedings of the Section on Business and Economics Statistics.

Domencich, T., and McFadden, D. (1975), Urban Travel Demand: A Behavioral Analysis, Amsterdam: North Holland Press.

Fay, Robert E. (1986), "Causal Models for Patterns of Nonresponse," Journal of the American Statistical Association, 81 (394), 354-365.

Greenlees, John S., Reece, William S., and Zieschang, Kimberly D., (1982), "Imputation of Missing Values When the Probability of Response Depends on the Variable Being Imputed," Journal of the American Statistical Association, (Applications Section) 77 (378), 251-261.

Judge, G.G., Hill, R.C., Griffiths, W., Lutkepohl, H., and Lee, T. (1982), Introduction to the Theory and Practice of Econometrics, New York: John Wiley and Sons.

Kinsey, Jean, and Lane, Sylvia (1978), "The Effect of Debt on Perceived Household Welfare," Journal of Consumer Affairs, 12 (1), 48-62.

Lillard, Lee, Smith, James P., and Welch, Finis (1986), "What Do We Really Know about Wages? The Importance of Nonreporting and Census Imputation," Journal of Political Economy, 94

Little, Roderick J.A., and Samuhel, Michael E. (1983), "Alternative Models for CPS Income Imputation," American Statistical Association, 1983 Proceedings of the Section on Survey Research Methods.

Maddala, G.S. (1977), Econometrics, New York: McGraw Hill Book Company.

Maddala, G.S. (1983), Limited-dependent and Qualitative Variables in Econometrics, Cambridge: Cambridge University Press.

Pindyck, R.S., and Rubinfeld, D.L. (1981), Econometric Models and Economic Forecasts (2nd ed.), New York: McGraw-Hill Book Company.

Tardiff, T. (1976), "A Note on Goodness of Fit Statistics for Probit and Logit Models," Transportation, 5(4), 377-388.

U.S. Department of Labor, Bureau of Labor Statistics, (1986), Consumer Expenditure Survey: Interview Survey 1984, Bulletin 2267, Washington, D.C.: U.S. Government Printing Office.

ACKNOWLEDGEMENTS

Thanks are extended to Laura A. Blanciforti, Paul Hsen, Eva Jacobs, Brent Moulton, Julie Nelson, Diane Woodbury, and Kimberly D. Zieschang for providing comments on an earlier draft of this paper. Thanks are also extended to Herb Cover for computer assistance, to Anna Sanders for expert typing, to Adriana Silberstein for providing input into the development of the paper, and to Kirk Kaneer for preparing presentation materials. Any errors which remain in the paper are the sole responsibility of the author. A more detailed version of this paper was presented at the Third Annual Census Research Conference, March 29 - April 1, 1987, Baltimore, Maryland.

RESEARCH AND DEVELOPMENT--COSTS AND BENEFITS OF NONFARM EMPLOYMENT FOR FARM FAMILIES

Jeanne L. Hafstrom, University of Illinois--Urbana-Champaign[1]
MaryAnn Paynter, University of Illinois--Urbana-Champaign[2]

SOFTWARE PROJECT BACKGROUND

The North Central Region has the largest percentage of farm population of the four geographical regions in the United States. Because of the current farm crisis, many farm families are finding it impossible to live on current farm incomes and wonder if the solution is to secure nonfarm employment for either the husband or wife. Farm families do not know where to turn to get help. When they seek counsel, professionals do not have data to use in helping them to make their decision.

A search of software uncovered no programs designed to provide effective help for farm families in decision making concerning cost and benefits of nonfarm employment of one or both spouses. Data from Illinois, Nebraska, North Dakota, and Wisconsin can illustrate advantages and disadvantages of various alternatives by predicting outcomes of several alternatives. These data can be supplemented with data available from other sources (e.g., income tax records, consumer expenditure data).

Our research team submitted a Proposal, "Development of Basic Self-Instructional Modules in Analysis of Benefits of Nonfarm Employment for Members of Farm Families," to the North Central Computing Institute (NCCI). The project, first proposed by Paynter [1], included personnel from Illinois, Nebraska, North Dakota, and Wisconsin. As requested, partial funding was received from NCCI for a Chicago meeting and for help with software programming. Inputs to the spreadsheet were agreed upon at that meeting. Hafstrom combined Team inputs into a computer spreadsheet program which included Notes on Supporting Data, the Worksheet, Questions to be Asked at Various Parts of the Worksheet, and Instructions for Programmer. A preliminary handbook, "Does a 2nd Job Pay Off?--Help for Family Decision-Making," also was developed.

WORKSHEET

Income

The program is designed to estimate net income for second earner employment. It includes calculation of gross income from employment plus other income related to employment (tips, etc.). Benefits (insurance, profit sharing, pension contributions, etc.) related to employment are calculated. Subtotals are calculated and entered into a total income cell.

Expenses

Some expenses are to be entered for a specific job or position (e.g., work clothing, dues/licenses, tools/supplies). For transportation, respondent is asked about vehicle purchase and/or replacement, license/taxes/fees, insurance, tolls, carpool, etc. To illustrate some family expenses rise with employment, categories estimate child care, hired household help, extra dry cleaning, meals out, etc. Personal expenses, dinners/lunches at work, clothing, personal care, are included. Other possible expenses incurred as a result of employment are hired farm help, gifts/parties. Totals for each major expense category and a Subtotal of expenses related to employment are calculated.

Totals

The expense subtotal is summed with any cost of benefits related to employment. Gross Income minus Total Expenses are calculated as net income from employment for one job/position. Totals are printed out for the family.

Noneconomic Considerations

Before the program will print out, participants are asked if they have discussed noneconomic considerations that will influence decisions with household members (e.g., sharing of household workload, child care, seasonal variation in farm workloads, volunteer commitments).

LOOP AND HELP SCREENS

This program can be repeated for different job types as many times as necessary. Then printouts can be compared for families to use in decision making about external employment. It is designed to use Help screens to raise questions at each step. Research findings and expertise by professionals are used. Data from income-expenditure studies, income tax records, and local sources are included or suggested. A handbook for use by counselors is being prepared. Completion is expected in Summer, 1987.

REFERENCE

1. Paynter, M. A Proposal for a Computer Program: Does Off-farm Work Pay Off?" Proceedings of Family Financial Management Professionals NCCI Workshop. St. Paul, MN, 1985, pp. 119-120.

[1]Associate Professor of Family and Consumer Economics
[2]Assistant Professor and Extension Specialist of Family Economics

PROGRAMS FOR FARM FAMILIES AND CONSUMERS

Karen P. Goebel, University of Wisconsin--Madison[1]
Nancy C. Hook[2]

SITUATION

Seek off-farm income has become the message to farm families as research has indicated the increasing proportion of income coming from nonfarm employment. In Wisconsin, programming efforts have been expanded to assist farm families in decision making to increase income. Efforts are comprehensive and include decision making in relation to both business operation and family system. The goal of the farm operation is to make a profit. Family goals are unique for each family, changing through the life cycle. Therefore, programs include more than just costs and benefits of off-farm income.

ANALYSIS OF FARM OPERATION

Using past farm records assists in determining profitability of the total operation and of each enterprise. Basic record keeping seminars can assist families in preparing for this analysis. Seminar topics include collecting, sorting/filing information and preparing statements of net worth, earnings, and farm/family expenses. The intertwined nature of family and farm financial affairs makes analysis complex. Analysis may suggest the need for more efficient operation of various enterprises, reorganization, expansion, diversification, or phasing out of the farm operation as well as supplemental income from off-farm employment.

A major decision is whether to stay in or give up farming. If the decision is to stay in farming, then consideration of off-farm income or alternative sources of income becomes feasible. Off-farm employment may be full- or part-time for either or both spouses. Programming for family members would desirably include employment opportunities and job search skills, résumé writing, interviewing, and related topics. Because off-farm income in Wisconsin is provided more frequently by farm women with higher educational levels, programming should include opportunities for upgrading and retraining to expand skills and abilities that will match the demand for employment.

If the decision is to leave farming, career development becomes paramount. Programming efforts in education and employment training are provided in tandem with the Vocational, Technical and Adult Education (VTAE) system, and other colleges and agencies in the community.

ANALYSIS OF OFF-FARM EMPLOYMENT

In analyzing costs and benefits of off-farm income, anticipated changes in economic, social, and psychological environment of the family need to be considered. The economic gain of fringe benefits is often an incentive for working off-farm, particularly if adequate health insurance is included. A manual to accompany software being developed includes questions for families to consider on costs and benefits of off-farm employment and role and responsibility changes within the family system. Farmers have personal and satisfaction needs to fulfill as well as a need to contribute to income. Women may find nonfarm employment causes increased emotional/physical stress and conflict because of ingrained traditional values. Recognizing these costs as well as the obvious economic costs is important and may call for innovative treatment and family/neighborhood support.

A "day of learning" is planned to incorporate: analysis of farm operation; employment opportunities and job search skills; development of alternative sources of income at home; educational opportunities; analysis of off-farm employment; household time and stress management. This model program will include exhibits, seminars and workshops. In-depth workshops will precede, follow, and be part of the "day of learning." Particularly important are workshops on records management and to enhance production and marketing skills of women. University of Wisconsin Extension will provide speakers, discussion leaders, and coordination of community resources. Agricultural, community development, and home economics faculty are working on this integrated program. VTAE system and other nonprofit community agencies have major responsibility for career development and employment skills. Extension is assuming major responsibility for analysis of the farm/household operation and alternative sources of income.

Various programs in the Midwest to support farm families in coping with the farm crisis have revealed a number of problems and assisted in formulating new programs. In Wisconsin, the UWEX SOS (Strategies On Survival) program included one-on-one counseling with individuals and families and workshops. County staff found counseling labor intensive but very effective in motivating families to act. Workshops were efficient but tended to be less effective in assisting families in decision making. Consequently, a volunteer farm financial counselor program is being piloted in three counties. Volunteers receive 30 hours of instruction and will assist families in management, financial analysis, and stress reduction. Referrals will be made to appropriate cooperating agencies.

[1]Associate Professor and Extension Specialist, University of Wisconsin, Madison, WI
[2]Financial Counselor/Analyst, Cambridge, WI

FINANCIAL MANAGEMENT: APPLICATION FOR THE CLASSROOM

M. M. (Peggy) Whan,[1] Sheila Mammen,[2] Judi Hornseth[3]
North Dakota State University

SITUATION OF FARM FAMILIES

According to a recent survey, the majority of the farm families in one state have reduced their present consumption level due to the current farm economic situation. The random survey of North Dakota farm families (N = 155) conducted in Summer 1986 indicated that:

- income, savings, and standard of living are important to 95% of the sample, however two-thirds of them are dissatisfied with their economic situation.
- it is important to 61% of the respondents to meet sudden financial demands, yet 70% of the sample is unable to meet a sudden financial demand.
- the current farm economic situation has caused three-fourths of them to reduce their present consumption of goods and services.
- more than one-half of the sample feel they have little or no control over their life.
- dissatisfaction with their financial situation contributes greatly to this perceived lack of control.

Budget counseling and better money management skills would help these families utilize their financial resources more effectively and therefore better cope with financial stress. This is especially true because farm families cannot control the market price of their farm commodities. Their recourse is to either change their expenditure level or to increase their income.

In most North Dakota rural communities there are few, if any, new off-farm employment opportunities available. Therefore, better financial management is the most viable alternative for these families.

CLASSROOM SOLUTIONS

Students in Family Economics classes at North Dakota University are interested in helping economically stressed farm families and other households get back on their feet financially. In order to work professionally with such families, college students must understand the principles of money management. To that end a reality-based, computerized case study incorporating money management principles was developed. Family Economics students are taught application of financial management principles using this computerized case study.

The major advantage of the case study method is that it enables all students in the personal and family finance course to work within a common framework, yet it allows for individual decision making. The microcomputer is an extremely effective teaching tool when spreadsheet programs are utilized because students can see the total financial plan immediately, examine alternatives, and obtain instant feedback. Student accuracy is greatly improved due to fewer mathematical errors. For the teacher, the major consideration is finding software that is inexpensive, that can be copied without infringement of copyright laws, is user friendly, and requires little time to learn.[4]

Each student is required to develop a complete case history of a family including its financial goals and philosophy toward money. The assignment emphasizes the theory and principles of money management taught throughout the course. The assignment is handed to the students at the beginning of the course and is to be completed at the end of the ten-week quarter. Students make value and goal judgments of their own through the planning process. Therefore, the assignment is personalized to meet the financial goals as cited by the student for that family.

The lack of complexity of the software and the ease with which the software can be applied relieves the apprehension of students who have not previously used the microcomputer. Student evaluations indicate they grasp concepts associated with budgeting and financial planning much more readily when they do not use the computer.

The knowledge of the theory of money management and its computer application will enhance the professional skills of Family Economics students who wish to counsel families experiencing financial difficulties. With this type of counseling, families will be helped to reassess and revise their financial goals so that they can better cope financially with their debt load. As farm families become more satisfied with their economics situation, it is hoped that they will regain control over their life.

---

[1]Associate Professor, Family Economics
[2]Associate Professor, Family Economics
[3]Extension Associate, Family Resource Management

[4]PC-Calc(tm), A Spreadsheet Program for the IBM Personal Computer, Jim Button, ButtonWard, P.O. Box 5786, Bellevue, WA 98006 met the criteria of this assignment.

INSERVICE EDUCATION FOR COMMUNITY WORKERS

Kathy Prochaska-Cue, University of Nebraska--Lincoln[1]

FARM FINANCIAL CRISIS

American rural families and communities are experiencing major stress as a result of the farm financial crisis. As most economists and public policy makers agree, the rural crisis will not be just a short-term problem. For many rural communities, a transition period of 5 to 10 years is expected [1].

Community services and groups such as churches, mental health personnel, social service and community action agency personnel, teachers, and those employed by financial institutions work directly with rural family members experiencing financial distress. Such "helpers" may not understand the nature of the rural problem, having had no actual experience with irregular, self-employed income. On the surface, the addition of off-farm employment seems to be a possible, temporary solution for rural financial problems, but the small actual financial benefits of such employment because of a lack of employment opportunities in the local community, low wage rates, and the additional financial costs of new employment may not be recognized.

The effects on the individual employed and other family members also need to be considered. Lack of time can lead to a change in roles, responsibilities, lifestyles, priorities, routines, and goals. When such changes are in conflict with the expectations and values of the worker and other family members, additional stress results.

CES RESPONSE

For many years the Cooperative Extension Service (CES) in most states has worked with various "helping" agencies and organizations in communities providing education for clientele and education inservice opportunities for the professionals and para-professionals who are employed or who voluntarily serve these community groups. An example of such cooperation is a continuing education program for clergy and spouses planned by the CES in Nebraska and Interchurch Ministries of Nebraska.

Because of the irregular nature of some clergy incomes, the workshop was planned to provide help for clergy and their spouses in their personal situations and when helping parishioners. The day-long workshop includes training in the use of CES family financial and stress management materials; information on community referral sources; discussion of the nature of farm income; alternatives available to families during times of financial problems; methods to creatively cope with stresses and problems of others; and the opportunity for clergy to share and spend time with colleagues. Workshop resource persons include: the CES specialists in family economics and management, and in family life; the local extension agent in agriculture; and a representative of Interchurch Ministries of Nebraska.

In Spring of 1987 the workshop is being piloted in the southeast quarter of the state. Planning was done with a committee composed of representatives of the two sponsoring organizations and clergy from the pilot area. Because of relatively low registration, future workshops will be tied to the district conferences of each religious denomination instead of maintaining the inter-denominational nature of the pilot.

FUTURE USE OF COMPUTER TEMPLATE

The computer template developed by our group for helping rural families decide whether an off-farm job is beneficial is a welcome addition for this type of community education.

Community workers could better understand the irregular nature of self-employed income and the "trade-offs" of taking off-farm employment by using the template to "model" case studies. In addition, they could refer clients and parishioners to CES for use of the computer template in an actual decision making process.

REFERENCE

1. Monroe, M. N. "At Issue: Rural Family Stress." Journal of Home Economics 78, (Winter 1986). pp. 57-58.

[1]Assistant Professor, and Extension Family Economics and Management Specialist