Household Income Reporting:
An Analysis of U.S. Consumer Expenditure Survey Data

Thesia I. Garner[1] and Laura A. Blanciforti[2]

Abstract: When income data are collected from household survey respondents, incomplete responses and nonresponses frequently result. The purpose of this study is to examine the relationship between particular socioeconomic attributes of households and the probability that a household will be a complete income reporter. Data from the Interview portion of the 1987 U.S. Consumer Expenditure Survey (CEX) are analyzed. For this survey, the household does not need to provide a complete accounting of income from all sources to be considered a complete income reporter. Binomial logit analysis is used to model the probability of income response completeness. Unique to this study, in contrast to other studies of income response, is the inclusion of expenditures as an explanatory variable in the estimating equation. Accounting for expenditures is particularly important for CEX data users since published tabulations of expenditures by income are restricted to complete income reporters only. Results from examining only the primary effects reveal that the age, race, education, and occupation of the reference person, and the consumer unit's composition, region of residence, expenditures, and participation in income means tested programs are significantly related to the income reporting status of the unit. This information should be useful for data users not familiar with the implications of using the U.S. CEX's definition of complete income reporter, for statisticians developing income imputation procedures for the survey, and for data collectors devising collection procedures to improve data quality. Specific suggestions for future research directions to further explain issues related to income response are provided.

Key words: Incomplete income response; household survey data; socioeconomic characteristics; expenditures; binomial logit analysis.

1. Introduction

Income is an important variable in economic analysis. However, when we try to collect income data from survey respondents, incomplete response or nonresponse frequently results. For example, a household may report having received income from employment but may not report the amount. Missing data may also result when the interviewer mistakenly fails to ask certain income questions. The availability of large survey data bases, with extensive demographic and economic data, allows researchers to become familiar with the patterns of income response and the underlying assumptions concerning the distribution of missing values, and to examine the statistical relationships between particular reporting unit attributes and the reporting of income. Understanding these relationships is important because how one chooses to exclude cases or impute values when there are missing data will have consequences for the conclusions that are ultimately drawn from the data.

For many surveys and for data analysis, nonmissing values for those cases that do not contain valid income information are substituted. The better the determined relationships between attributes and income reporting are, the better the substituted or imputed values. Yet, depending upon the imputation procedure employed, biased and inconsistent estimates may result. Caution is warranted when a large percentage of income data is imputed.

In contrast, some data producers and researchers prefer not to substitute income or to use imputed income values in their analyses. Such is the case for the U.S. Consumer Expenditure Survey (CEX). Instead the producers of the CEX data, the Bureau of Labor Statistics (BLS), provide a variable in the data file that identifies households or consumer units as complete income reporters or not. According to the definition of this variable, a consumer unit does not need to provide a complete accounting of income from all sources to be considered a complete income reporter; complete response is only required for a subset of income sources and consumer unit members. Yet, if an analysis is restricted to complete income reporters only, estimations based on income as an independent variable in a regression may involve loss of efficiency. An analysis restricted to complete income reporters where income is the dependent variable results in coefficients of the independent variables being subject to possible bias, because the probability of having a missing value is a function of the error term.

The complete response status designation was created for the CEX to determine whether consumer units provide sufficient income information for the BLS to meet its publication requirements to produce expenditure by income class tabulations, although data users frequently use the variable to restrict their samples for multivariate analyses. The complete income reporter designation has been followed by the BLS since it was devised for use with the 1972-73 CEX (U.S. Department of Labor 1981). An early comparison of the 1972-73 CEX and the 1972 Current Population Survey (CPS) income distributions revealed that the distributions were "for all practical purposes" identical. The CEX data were restricted to complete income reporters while the CPS file included imputed income values. Based on this comparison, CEX analysts concluded that the closeness of the distributions supported use of the complete income reporter definition based on the partial income codes for income data collected in the continuing CEX (U.S. Department of. Labor 1980). Expected advantages of the sufficient

income completeness approach are that the number of "complete reporters" will be larger and more representative of the sample population that when a stricter definition of report completeness is used. Disadvantages are that incomes are likely to be underreported for "complete" households not reporting income from all sources and that complete and incomplete reporter households may be systematically different from each other. The latter implies that analytical results based on a sample of complete income reporters only may not be generalizable to the U.S. civilian noninstitutional population, the target population.

The purpose of this study is to examine the relationship between the probability that a consumer unit will be a complete income reporter, as defined for the CEX, and particular socioeconomic attributes. Results from earlier studies (Greenlees, Reece, and Zieschang 1982; Little and Samuhel 1983; Coder and Feldman 1984; Andrews and Herzog 1986; Lillard, Smith, and Welch 1986; Garner and Blanciforti 1987) were used as a basis for identifying variables to include in the probability of income response model. The unit of analysis for this study is the consumer unit, with income from members aged 14 years or older considered. This is in contrast to most other studies where the household member was the unit of analysis, with the reporting of income for particular sources examined. The samples for these studies were often restricted to working age males as well; an exception was provided by Coder and Feldman (1984).

Note that the term consumer unit, rather than household, is used throughout the remainder of this paper since the consumer unit is the basic reporting unit for the CEX. Simply, 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 arrangement; (2) a person living alone or sharing a household with others, 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 make joint expenditure decisions. Financial independence is determined by 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 1991).

Data from the 1987 CEX Interview Survey are analyzed. This sample is more representative of the U.S. civilian noninstitutional population than was the sample used in our earlier CEX study in that the earlier sample did not include consumer units living in rural areas outside Standard Metropolitan Statistical Areas. Independent variables in the present model that were not considered earlier include consumer unit composition or structure, expenditures, and whether the consumer unit participates in an income means tested program. Again, binomial logit analysis is used to model the probability of income response completeness. Results reveal that age, race, education, and occupation of the reference person, defined as the first member of the consumer unit mentioned by the respondent when asked who owns or rents the home, and the consumer unit's composition, region of residence, expenditures, and income means tested program participation status are significantly related to the income reporting status of consumer units. Results from this and the earlier CEX study are consistent for variables in both models. An additional analysis is conducted to determine the effect on the

composition of the CEX sample when a stricter definition of complete income reporter is assumed.

Examining the issue of income response for the CEX is of primary interest for three reasons. First, the CEX is the only survey in the U.S. that collects information on detailed categories of consumer expenditures and income. Consumer unit socioeconomic characteristics are also collected so that expenditures can be related to these characteristics. As such, the data set is a rich source of information available to researchers to conduct economic analysis. Therefore, it is important that researchers be familiar with the less restrictive definition of complete income reporter used by the BLS, and the assumptions underlying the resulting distribution of income. Second, since income is an important variable used for BLS tabulations and research, there is interest in developing an income imputation procedure for the survey. To produce unbiased and consistent imputed value estimates, it is necessary that systematic differences between complete and incomplete reporters be accounted for if they exist. And third, devising improved data collection procedures to overcome the nonreporting of income, specifically to improve data quality, is a constant goal of the BLS.

The ideas and results presented in this paper represent continued work at the BLS to examine issues related to income nonresponse in the CEX. This paper is organized into three remaining sections: Methods and Procedures, Results, and Summary and Conclusions.

2. Methods and Procedures

2.1. Data

The data used in this study were obtained from the Interview portion of the 1987

U.S. CEX. The CEX data are collected by the Bureau of the Census under the auspices of the BLS. The Interview sample, selected on a rotating panel basis since 1980, is targeted at approximately 5,000 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 is defined as all consumer units participating in a second interview during 1987 (U.S. Department of Labor 1987).

2.1.1. Outcome variable

Income reporting is 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 CEX data, is based on whether the respondent provides valid responses to questions concerning "major sources" of income and selected "other sources" of income. A valid response could be a positive or negative income value, a zero, or a blank. A valid zero or valid blank is recorded if the consumer unit member indicates that income was not received from a particular source. This is not a refusal to respond. A zero is used for a source where income was received in the past but the member acknowledges there was no income for the current reference period. A blank is used when the member acknowledges that a category is not a source of income. Major sources of income include wages and salaries, income or loss from non-farm business, partnership, or professional practice, income or loss from own farm, Social Security or Railroad Retirement, and Supplemental Security Income. Other sources of income include

unemployment compensation, worker's compensation, and veteran's payment including educational benefits but excluding military retirement, 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, the military, or government, income or loss received from roomers or boarders, income or loss received from payments from other rental units, and regular contributions received from alimony, child support, or from others outside the consumer unit. 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, fellowships, or stipends not based on working, and the cash value of food stamps. Income from these latter sources, however, is added to the income from the sources noted previously to calculate a consumer unit's total income before taxes. Income before taxes is the income variable used in CEX publications to classify consumer units (U.S. Department of Labor 1987).

In general, complete income reporters are defined as consumer units that report a non-zero amount for a major source of income or a non-zero amount for selected other sources of income, while refusals or don't know responses identify consumer units as incomplete income reporters. A consumer unit is defined as a complete income reporter if:

a. The reference person has a non-zero amount reported for a major source of income; or

b. A consumer unit member(s) other than the reference person has a non-zero

amount reported for a major source of income, and valid zeroes or valid blanks are recorded for all the major sources of income for the reference person; or

c. The consumer unit has a non-zero amount reported for at least one other source of income and valid zeroes or valid blanks are recorded from any major sources for any of the members.

Given this definition, as noted earlier, it is possible for complete income reporters not to have provided a full accounting of income from all sources. Consumer units with other combinations of entries to the income questions are considered to be incomplete income reporters. For example, a two-earner family would be designated an incomplete reporter if the spouse's wages and salaries were reported but the respondent refused or was unable to report the wages and salaries of the reference person. In contrast, a two-earner family would be a complete reporter if the reference person's wages and salaries were reported but the spouse's were not due to refusal. In the extreme case of across the board zero income, the response is considered invalid and thereby constitutes an incomplete income report.

2.1.2. Independent variables

Eleven socioeconomic factors are included in the model as independent variables. Five are characteristics of the reference person in the consumer unit: age, race, sex, education, and primary occupation. Reference person characteristics are used for the analysis since the consumer unit's characteristics are most often identified in terms of the reference person by the BLS for publication and research. The other six variables expected to be important in the

income report completeness probability model for the CEX are consumer unit characteristics: consumer unit composition, housing tenure, degree of urbanization, region, expenditures, and income means tested program participation.

2.1.3. Distribution of variables

Definitions and the distribution of variables included in the model are presented in Appendix A. All variables enter the model as categorical variables. Complete income reporting consumer units account for approximately 85% of the 1987 sample. Reference persons in the consumer unit sample were most likely to be in the 35 to 54 year age group. The majority of reference persons were white and male. More than 30% were high school graduates, and approximately one-fifth were salaried professionals or managers. Singles represented the most frequently reported consumer unit structure at 28% of the sample. Approximately 20% of the consumer units reported total quarterly expenditures of $1,500 to $3,000. The majority of consumer units in the sample were not participating in income means tested programs. The largest percentage of consumer units in the sample owned their own homes and they lived inside a large urban area. Consumer units were fairly equally distributed among regions.

2.2. Estimation procedure

The statistical analysis of the probability of complete income response is based upon a binomial logit model (Domencich and McFadden 1975; Judge, Hill, Griffiths, Lutkepohl, and Lee 1982; Maddala 1977; 1983; and Pindyck and Rubinfeld 1981). In this study the model under consideration is

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

where the $F(X_i\boldsymbol{\beta})$ is a cumulative distribu-

tion function that describes how the probability of complete income reporting is related to the socioeconomic variables, X_i. P_i is the probability that the ith consumer unit is a complete income reporter, $\boldsymbol{\beta}$ 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

$$P_i = F(X_i\boldsymbol{\beta}) = \frac{1}{1 + e^{-X_i\boldsymbol{\beta}}} \qquad (2)$$

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

3. Results

3.1. Logit analysis

Results of the logit analysis are displayed in Table 1. All variables included in the model, their estimated coefficients, and asymptotic standard errors are presented. In Table 1 each omitted or reference group is identified in parentheses next to the variable name. Asymptotic *t*-tests are used to determine the statistical significance of individual coefficients. Only primary effects are examined since no interaction terms are included in the model. Among the socioeconomic variables included in the probability model, coefficients for the following are statistically significant: age (34 years or less and 65 years or greater), race, educated through college, having a self-employed reference person, three of the consumer unit composition variables, two of the region variables (Northeast and Midwest), all but one expenditure category (expenditures of $3,000 to $4,499),

Table 1. Estimated model coefficients and standard errors: 1987 U.S. consumer expenditure survey second interview complete income reporter (N=5,936)

Independent Variable	Estimated coefficient	Asymptotic standard error
Age of reference person (age 35 to 54 years)		
Age 34 years or less	0.2970[a]	0.1038
Age 55 to 64 years	0.1366	0.1224
Age 65 years or greater	0.5885[a]	0.1639
Race of reference person (Non-black)		
Black	−0.3777[a]	0.1256
Sex of reference person (male)		
Female	−0.1265	0.1047
Education of reference person (high school graduate)		
Elementary	0.1489	0.1385
Did not complete high school	0.1454	0.1264
Did not complete college	0.0840	0.1108
College graduate	−0.3934[a]	0.1346
Postgraduate	−0.0671	0.1508
Principal occupation of reference person (salaried professional)		
Laborer	0.2574	0.1576
Craft	0.0824	0.1791
Sales	0.0244	0.1307
Services	0.1638	0.1719
Self-employed	−0.8065[a]	0.1500
Retired	0.0542	0.1790
Not working and other	−0.0450	0.1634
Consumer unit composition (single)		
Single parent	0.1564	0.2029
Husband and wife only	−0.5210[a]	0.1325
Husband, wife, and child under 18	−0.4672[a]	0.1418
Husband, wife and other	−0.9045[a]	0.1532
Other family combinations	−0.1978	0.1534
Housing tenure (owns)		
Rents	0.1698	0.0980
Degree urban (large urban area)		
Small urban area	−0.1054	0.0893
Rural	0.0709	0.1038
Region (South)		
Northeast	−0.7116[a]	0.1056
Midwest	−0.2894[a]	0.1042
West	0.0762	0.1209
Expenditure categories (expenditures of $4,500 to $5,999)		
Expenditures of less than $1,500	−0.4848[a]	0.1869
Expenditures of $1,500 to $2,999	−0.5239[a]	0.1318
Expenditures of $3,000 to $4,499	0.0005	0.1281
Expenditures of $6,000 to $7,499	0.3635[b]	0.1452
Expenditures of $7,500 to $9,999	0.6960[a]	0.1555
Expenditures of $10,000 or more	0.5785[a]	0.1548
Income means tested program (non-participant)		
Participant	0.4450[a]	0.1362
Constant	2.1134[a]	0.2115
Likelihood ratio statistic	268.60 > 63.69[a]	
Likelihood ratio index	0.054	

[a]Statistically significant at the 0.01 level.
[b]Statistically significant at the 0.05 level.

and income means tested program participation.

3.1.1. Age

The positive coefficients for the age variables indicate that consumer units with younger (less than 35 years) and older (greater than 64 years) reference persons are more likely than those in the middle age groups to be complete reporters of income. Since they are in their peak years with respect to employment earnings, consumer units characterized as in the middle age groups are likely to have higher incomes or receive income from more numerous sources than those in the other age groups. These consumer units may be less likely to divulge their higher incomes or they may be less willing or able to report dollar amounts received from each income source. Also, since the interview questionnaire is somewhat lengthy with income questions placed at the end, it is likely that the more numerous the sources of income, the more likely respondents will provide incomplete reports.

Categorical rather than continuous age variables were entered in this model in order to directly identify where in the specific age distribution differences in response occur. This approach was followed since our earlier analysis incorporated a continuous age and age squared term (Garner and Blanciforti 1987). We found that the coefficient for age was negative while the coefficient for age squared was positive but very small. This indicated that as age increased the negative effect of age on the probability of complete income reporting diminished.

3.1.2. Race

Based on this analysis, blacks are significantly less likely than are non-blacks to be complete income reporters. This is in contrast to the results of Greenlees, Reece, and Zieschang (1982) who reported no relationship between the reporting of wage income and race. Whether the difference, with respect to the CEX, is related to differences in levels or sources of income, culture, or other factors is a question for future analysis.

3.1.3. Education

Of all the education coefficients, only the one for college graduate is statistically significant in the probability model. Consumer units with college educated reference persons are significantly less likely than those in the omitted category (high school degree) to be complete income reporters. This is consistent, in general, with studies of the probability of reporting earnings which show that increases in the years of completed education led to decreases in income response probability (Greenlees, Reece, and Zieschang 1982; Lillard, Smith, and Welch 1986). One might expect consumer units with educated reference persons to be better equipped to answer detailed income questions. Yet, more educated consumer units may have higher incomes and more varied sources of income that require more detail in their income response. In addition, more educated consumer units may place a higher value on their time and privacy than do the less educated.

3.1.4. Occupation

Self-employment of the reference person is the only occupation variable in the probability of complete income response model with a statistically significant coefficient. Consumer units with self-employed reference persons are less likely to be complete income reporters than are consumer units with salaried professional or managerial reference persons. Consumer units with reference persons in the other occupations did not differ significantly from those

represented by the omitted group in their completeness of income response. Coder and Feldman (1984) reported a similar finding for the 1983 Survey of Income and Program Participation (SIPP) nonresponse rates for individuals with self-employment income, with nonresponse rates for the self-employed exceeding the rates for individuals with income from wages and salaries. A related finding was also reported by Lillard, Smith, and Welch (1986) in their examination of the proportion of non-reporting white males by type of earnings. They noted that self-employed individuals were in occupations where nonreporting was 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" (Lillard, Smith, and Welch 1986, p. 492).

3.1.5. Consumer unit composition
Consumer unit composition is included in the model to determine whether there is differential response between single person consumer units and more complex household structure units. As expected, consumer units with more complex composition are less likely than single person consumer units to be complete income reporters. Our results reveal that consumer units in the three husband-wife consumer unit groups are significantly less likely than other units to be complete income reporters. And, among these three, the most complex (husband, wife, and other) is least likely to respond. Survey respondents in such multiple member units are expected to face difficulty in providing information about income for other members.

3.1.6. Region
Region represents the only geographic variables with statistically significant coefficients. Each consumer unit is identified as living in one of four regions: Northeast, Midwest, West, or South. Consumer units living in the Northeast and those living in the Midwest are significantly less likely than those living in the South to be complete income reporters. These differences may be related to differences in the types of income received or the regional patterns of cooperation, more specifically to the willingness on the part of consumer units within certain regions to respond to income questions or to cooperate with data collectors in general. Previous researchers (Lillard, Smith, and Welch 1986; Greenlees, Reece and Zieschang 1982) using census data noted the importance of region in determining the reporting propensities of earnings. Lillard, Smith, and Welch (1986) reported that living in the South had a strong positive independent effect on reporting propensities, even after controlling for the variable's influence through earnings. Greenlees, Reece, and Zieschang (1982) found that individuals living in the South or West were most likely to report values for wages and salaries in the CPS.

3.1.7. Expenditure categories
The quarterly expenditures of consumer units are divided into seven expenditure categories defined in $1,500 increments up to $10,000. These results reveal that consumer units with lower quarterly expenditures are more likely to be incomplete income reporters than are those with expenditures in the middle range of expenditures, while those with higher expenditures are more likely to be complete income reporters. Identifying the relationship between income response completeness and expenditures is especially

important since expenditures are believed to be highly correlated with income. This result has two important implications for the CEX data. First, it is likely that consumer units with low expenditures are being excluded from BLS tabulations of expenditures by income since such tabulations are produced for complete income reporting consumer units only. Second, understanding the relationship between income report completeness and expenditures is important particularly for the development of income imputation methods for the CEX, since an aim of such methods will be to maintain the relationship between income and expenditures. If expenditures and income are positively correlated, it may be that the incomes of complete income reporters are higher than those of incomplete reporters. Alternatively, if the incomes of complete income reporters are not necessarily higher than those of incomplete income reporters, questions about the relationship between expenditures and income arise. And, it may be that respondents who do not report income completely, also do not report expenditures completely. These results suggest that additional research is needed to determine whether lower expenditures are more specifically associated with lower incomes in general or with income from particular sources.

3.1.8.   Income means tested program participation

Income from means-tested cash transfers or public assistance has been identified by others (U.S. Department of Commerce 1990; Vaughan 1989) as a source of income underreporting in the CPS and in the SIPP and, as a result, is likely to be a source of underreporting in the CEX as well. In 1983, based on a comparison with independent estimates, the CPS identified only 76% of Aid to Families with Dependent Children benefits and 85% of Supplemental Security Income benefits after income imputations. We include a variable in our response probability model to identify consumer units as participants in means tested programs to determine whether the complete income reporter identifier could serve as an indicator for the possible underreporting of income from such transfers. Our results reveal that consumer unit program participants are more likely than nonparticipants to be complete income reporters. Thus, these individuals are being included in the group of complete income reporters. Consumer units in these households are likely to report little income from other sources since income limit restrictions exist for program participation. If the CEX reports of income from means tested cash transfer programs are underreported, as they are in the CPS and SIPP, it is likely that expenditures will exceed income for some of these households.

3.2.   Chi-square tests

To test the overall significance of the model, the likelihood ratio statistic is used. The test statistic, $\chi^2 = -2(\ln \text{Likelihood}_U - \ln \text{Likelihood}_R)$, is asymptotically distributed with the degrees of freedom equal to the number of coefficients that are restricted or 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 coefficients, that is the intercept and all explanatory variables. The resulting chi-square value is significant at the 0.01 level. The null hypothesis that all of the coefficients (except the intercept) are equal to zero was rejected.

The likelihood ratio or pseudo-R^2 index, defined as $\rho^2 = 1 - (\ln \text{Likelihood}_U / \ln \text{Likelihood}_R)$, is calculated as a measure of goodness-of-fit of the binomial logit model and is analogous to the R-square in ordinary least squares regression. As such it is a measure of how well the model approximates the observed data. Generally, this likelihood ratio index has an upper bound of about 0.3; it is unlikely that an index value would approach one because that could happen only if the model predicted perfectly (Judge, Hill, Griffiths, Lutkepohl, and Lee 1982; Kinsey and Lane 1978; Pindyck and Rubinfeld 1981; Tardiff 1976). An index of 0.054 is obtained for our response model (Table 1). Although this value seems low, it may be reasonable given that values of the index between 0.2 and 0.4 are considered extremely good fits (Hensher and Johnson 1981).

Tests for the combined contribution of variables were also performed using likelihood ratio statistics. Results are presented in Table 2. All of the variables sets except sex, urbanization, and housing tenure contributed significantly to the income response model as explanatory variables.

3.3. Sample probability calculations

Sample probabilities are produced to provide a quantitative assessment of changes in independent variables on the probabilities that consumer units would be complete income reporters. These probabilities are presented in Table 3. Estimated coefficients and consumer unit characteristics are used to compute the probability of complete income reporting according to the logistic cumulative distribution function as defined in equation (2). To assess the effect of changes, a consumer unit with selected characteristics is identified for comparison. This consumer unit is characterized as having a reference person who is aged 35 to 54, non-black, male, a high school graduate, and a salaried professional. The depicted consumer unit is single, owns his own home, lives in a large urban area in the South, and does not participate in an income means tested program. His average 1987 quarterly expenditures are between $4,500 and $5,999.

On average, a consumer unit participating in a second Interview of the CEX in 1987 has a 0.8671 probability of being a complete income reporter (Table 3). The

Table 2. Chi-square tests for contributions of sets of variables: 1987 U.S. consumer expenditure survey second interview complete income reporter (N = 5,936)

Independent variable groups	Chi-square	Degrees of freedom
Age of reference person	18.694[a]	3
Race of reference person	8.713[a]	1
Sex of reference person	1.458	1
Education of reference person	15.045[b]	5
Principal occupation of reference person	47.304[a]	7
Consumer unit composition	40.827[a]	5
Housing tenure	3.027	1
Degree urban	3.055	2
Region	63.757[a]	3
Expenditures	71.805[a]	6
Income means tested program	11.199[a]	1

[a]Statistically significant at the 0.01 level.
[b]Statistically significant at the 0.05 level.

Table 3.   Sample probability calculations: 1987 U.S. consumer expenditure survey second interview complete income reporter (N = 5,936)

Probability at the mean	0.8671
Baseline probability[1]	0.8922
Probability with change to consumer unit with selected characteristics:	
Age 34 years or less	0.9176
Age 55 to 64 years	0.9047
Age 65 years or greater	0.9371
Black	0.8501
Female	0.8794
Elementary	0.9057
Did not complete high school	0.9054
Did not complete college	0.9000
College graduate	0.8481
Postgraduate school	0.8856
Laborer	0.9146
Craft worker	0.8999
Sales	0.8945
Services	0.9070
Self-employed	0.7870
Retired	0.8973
Not working	0.8878
Single parent	0.9063
Husband and wife only	0.8310
Husband, wife, and child under 18	0.8384
Husband, wife and other	0.7701
Other family combinations	0.8716
Rents	0.9075
Small urban area	0.8816
Rural	0.8988
Northeast	0.8025
Midwest	0.8610
West	0.8993
Expenditures of less than $1,500	0.8360
Expenditures of $1,500 to $2,999	0.8305
Expenditures of $3,000 to $4,499	0.8922
Expenditures of $6,000 to $7,499	0.9225
Expenditures of $7,500 to $9,999	0.9432
Expenditures of $10,000 or more	0.9366
Income means tested program participant	0.9281

[1] The base line probability was calculated for the consumer unit characterized as: age = 35 to 54 years; race = non-black; sex = male; education = high school graduate; occupation = salaried professional or manager; consumer unit composition = single; housing tenure = owns; degree urban = inside a large urban area; region = South; expenditure group = $4,500 to $5,999; income means tested program = non-participant.

specified consumer unit has a 0.8922 probability of being a complete income reporter. As expected, changes in the probabilities are in the same direction as the signs of the estimated parameters. The most dramatic effect on the response probabilities is recorded for the "husband, wife, and other" persons consumer unit composition. A change in the consumer unit's composition from "single" to "husband,

wife, and other" results in a decreased probability of complete income response to 0.7701. Consumer units with quarterly expenditures of between $7,500 and $9,999, but with the same other representative characteristics, are the most likely to be complete income respondents.

3.4. Comparison to stricter definition of complete income reporter

As noted earlier, the CEX definition of complete income reporter can, and most likely does, produce a sample of consumer units not reporting income from all sources. Thus, it is likely that the incomes for complete income reporters are underreported from the application of this definition alone, although other reasons for underreporting are expected as well. To examine how a stricter definition of complete income reporter could affect the composition of the sample, an additional analysis was conducted. A justification for using the stricter definition would be to reduce the effects of underreporting introduced with the less stringent CEX definition. For this examination, the consumer unit is designated as a complete income reporter if the income fields across all members within a consumer unit are coded with valid responses only, and at least one non-zero amount is included among the set of valid responses. An incomplete response results when income fields are coded invalid for any income question where a monetary amount is reported or zeroes or blanks are recorded across all sources for all members. Refer to Table 1B in Appendix B where the variable and the distribution of the stricter definition of complete income reporters are presented. Results from this analysis are displayed in Tables 2B–4B in Appendix B; these tables correspond to Tables 1–3.

Approximately 68% of the sample consumer units are complete income reporters following the stricter definition as compared to 85% using the current CEX definition. Results from the logit analysis based on the stricter definition of complete income reporter, reveal that the composition of the complete income sample is likely to change significantly with respect to certain consumer unit characteristics. In contrast to the findings presented previously, consumer units with older reference persons are more likely not to be complete income reporters. Sex of the reference person did not contribute significantly to the earlier model; however, these results imply that consumer units with female reference persons are significantly less likely to be complete income reporters than are those with male reference persons. Renters are more likely than homeowners to be complete income reporters, as are consumer units living in rural as opposed to urban areas, unlike the earlier results. The results for both education and occupation of the reference person, region of residence, and income means tested program status essentially remain the same across the two analyses. The consumer unit structure results reveal a similar pattern of nonresponse as in the previous analysis, with more complex consumer units less likely to be complete reporters. The stricter definition analysis also supports the finding that consumer units with greater expenditures are more likely to be complete income reporters than are those with lower expenditures. Race contributed significantly to the earlier model; however in this latter analysis, race was not a differentiating variable for reporters versus nonreporters.

A stricter definition results in a smaller sample with somewhat different socioeconomic attributes than the one obtained

when applying the CEX definition of complete income reporter. Researchers must individually decide whether the trade-off for less potential underreporting of income is worth the reduction in sample size and population representativeness.

4. Summary and Conclusions

The purpose of this research was to identify socioeconomic factors that are related to income reporting completeness for consumer units participating in the U.S. CEX. This is in contrast to most earlier income reporting studies where characteristics were related to individual household member response probabilities. Results of the logit analysis reveal that the age, race, education, and occupation of the reference person, consumer unit's composition, region of residence, expenditures, and participation in income means tested programs are significant variables in determining the probability that a consumer unit would be defined as 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 the sources of 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, although preliminary research indicates this is true. The results are perhaps limited in another way as well. Even for complete income fields, we cannot make a statement concerning the reliability of the income provided.

This is a separate but related issue for future research.

Results from this study do, however, have important implications for research when income data from the CEX are used. Researchers interested in using these data need to be aware that complete and incomplete reporters of income are different, and that these differences may lead to loss of efficiency in parameter estimates, or to biased estimation results if not accounted for in one's estimation procedure. As noted earlier, for the CEX, not all income must be provided for a household to be considered a complete income reporter. When conducting analyses, researchers can restrict their samples to complete income reporters without the loss in sample size that would result if a stricter definition of completeness is used. However, the total income of complete reporters are likely to be more underreported than they would be with a stricter definition. A stricter definition would mean a reduction of approximately 20% in sample size and potentially a loss in population representativeness. Providing income from certain sources for consumer unit members can be sufficient for the consumer unit to be a complete reporter; yet, some consumer units may be eliminated from the analysis because their only source of income is not included in the set of income sources to qualify them as complete reporters.

Analyses are underway at the BLS currently to examine the extent of these problems in the CEX data. When considering an income imputation procedure for the CEX, results from this study can serve as a basis for evaluating model-based imputation procedures that account for cases where the incompleteness of income is related to income itself (David, Little, Samuhel, and Triest 1983; Fay 1986). In addition, 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 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 outcome variable could be defined to represent the three types of complete income reporting situations noted in the outcome variable section, plus incomplete reporting categories for combinations of "refusals" and "don't knows." Or, response completeness 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 variable could be defined in terms of the income reporting of individuals within consumer units; this would provide results more comparable to previous studies. Examination of the relationship between income report completeness and income level and the interaction of income with the socioeconomic characteristics of the consumer unit also need to be conducted. In lieu of, or in addition to, the reference person's characteristics, those of the survey respondent, if different, could be considered as explanatory variables in the model. One might expect persons with particular characteristics to be better respondents than others. For example, Andrews and Herzog (1986), conducting research on survey measures, found that respondent age is the demographic factor associated with the largest differences in survey data quality, and that

elderly individuals may answer with less precision than younger individuals. Additional socioeconomic variables that might be related to income response completeness include the 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, race, education, and occupation.

To test hypotheses concerning data collection features of the survey, administrative variables should be considered. These might include the total number of minutes for the interview, month when the interview was conducted, general survey nonresponse relative to specific income nonresponse, whether records were used in answering an interviewer's questions, and information about the general willingness of the respondent to cooperate with the interviewer, if available.

For whatever reason, consumer units that are identified as incomplete income reporters are not providing sufficient income responses for the purposes of BLS to be included in the expenditure by income tabulations. Reasons for incomplete income response are varied and complex. Cognitive research can be used to identify reasons for incomplete responses; subsequently, revisions to the questionnaire and data collection procedures can be introduced to increase income response completeness.

In conclusion, this study must be considered one of exploration. However, we think that the results are sufficiently promising to warrant future research on the completeness of income reporting and the missingness of income in the U.S. CEX.

Appendix A

Table 1A. Variables, distribution, and definitions: 1987 U.S. consumer expenditure survey second interview complete income reporter (N = 5,936)

Variable	Mean	Standard deviation	Definition
REP	0.8524	0.3547	Unity if consumer unit identified as a complete income reporter; zero otherwise.
NONREP*	0.1476	0.3547	Unity if consumer unit not identified as a complete reporter; zero otherwise
AGE34	0.3037	0.4599	Unity if reference person age 34 years or younger; zero otherwise.
AGE54*	0.3526	0.4778	Unity if reference person age 45 to 54 years; zero otherwise.
AGE64	0.1402	0.3472	Unity if reference person age 55 to 64 years; zero otherwise.
AGE65	0.2035	0.4026	Unity if reference person age 65 years and over; zero otherwise.
BLACK	0.1108	0.3140	Unity if race of reference person is black; zero otherwise.
NONBLK*	0.8892	0.3140	Unity if race of reference person is white or other non-black race; zero otherwise.
FEMALE	0.3349	0.4720	Unity if sex of reference person is female; zero otherwise.
MALE*	0.6651	0.4720	Unity if sex of reference person is male; zero otherwise.
ELEM	0.1200	0.3249	Unity if reference person completed eight years or less (including none) of education; zero otherwise.
LTHISCH	0.1331	0.3400	Unity if reference person completed more than eight years of education but did not complete high school; zero otherwise.
HISCHG*	0.3117	0.4632	Unity if reference person graduated high school; zero otherwise.
LTCOLL	0.2220	0.4156	Unity if reference person attended college but did not complete a college degree (4 years completed); zero otherwise.
COLLGR	0.1085	0.3110	Unity if reference person graduated college (4 years); zero otherwise.
POSTGRAD	0.1048	0.3063	Unity if reference person completed more than 4 years of college; zero otherwise.
LABORER	0.1181	0.3227	Unity if reference person received the most earnings in the past 12 months from salaried employment as an extraction, metal, or wood worker, an operative, or general laborer; zero otherwise.
CRAFT	0.0713	0.2573	Unity if reference person received the most earnings in the past 12 months from salaried employment in a precision production, craft, or repair occupation (e.g., mechanic, repairer, construction worker, textile or apparel worker); zero otherwise.

Table 1A. *Continued*

Variable	Mean	Standard deviation	Definition
SALES	0.1717	0.3771	Unity if reference person received the most earnings in the past 12 months from salaried employment in a technical occupation, in sales (business goods and services or retail), in a supervisory position, or in an administrative support or clerical occupation; zero otherwise.
SERVICES	0.0819	0.2742	Unity if reference person received the most earnings in the past 12 months from salaried employment as a private household worker, a farmer or farm laborer, a protective or other service employee in forestry, fishing, and groundskeeping, or in the armed forces; zero otherwise.
SELFEM	0.0649	0.2463	Unity if reference person received the most earnings in the past 12 months from self-employment (all occupations); zero otherwise.
RETIRED	0.1739	0.3789	Unity if reference person was retired in the past 12 months; zero otherwise.
NOTWKO	0.1080	0.3104	Unity if reference person was not working in the past 12 months (e.g., unemployed, working without pay, not working because ill or disabled, taking care of home/family, going to school, or doing something else) or member did not respond to the occupation question; zero otherwise.
PROFESS*	0.2106	0.4078	Unity if reference person received the most earnings in the past 12 months from employment as a salaried professional or manager, teacher, or educational or vocational counselor; zero otherwise.
SINGLE*	0.2783	0.4482	Unity if consumer unit includes one person only; zero otherwise.
SINPAR	0.0601	0.2378	Unity if consumer unit includes one parent with at least one child less than 18 years of age; zero otherwise.
HWONLY	0.2106	0.4078	Unity if consumer unit includes only the husband and wife; zero otherwise.
HWCHI	0.2238	0.4169	Unity if consumer unit includes a husband, wife, and a child or children with the oldest child less than 18 years of age; zero otherwise.
HWOTH	0.1223	0.3278	Unity if consumer unit includes a husband, wife, and at least one other person with the latter being a child who is 18 years of age or older or who is a grandchild, in-law, brother or sister of the reference person, mother or father of the reference person, or other related and unrelated persons; zero otherwise.

Table 1A. Continued

Variable	Mean	Standard deviation	Definition
OTHFAM	0.1048	0.3063	Unity if consumer unit includes other combinations of individuals than those noted previously; zero otherwise.
RENTS	0.3763	0.4845	Unity if consumer unit rents its home; zero otherwise.
OWNS*	0.6237	0.4845	Unity if consumer owns its home; zero otherwise.
SMAURB	0.3073	0.4614	Unity if consumer unit resides in an urban place with a population of less than 25,000, in another urban territory, or in an urbanized area or urban place of 2,500 to 25,000 outside an urbanized area (metropolitan statistical area (MSA) and non-MSA); zero otherwise.
LARURB*	0.4660	0.4990	Unity if consumer unit resides in the central city of a metropolitan statistical area or in another place with a population of 25,000 or over; zero otherwise.
RURAL	0.2268	0.4188	Unity if consumer unit resides in a rural area; zero otherwise.
NOREAST	0.2070	0.4052	Unity if consumer unit resides in the Northeast region; zero otherwise.
SOUTH*	0.2953	0.4562	Unity if consumer unit resides in the South region; zero otherwise.
MIDWEST	0.2672	0.4425	Unity if consumer unit resides in the Midwest region; zero otherwise.
WEST	0.2305	0.4212	Unity if consumer unit resides in the West region; zero otherwise.
EXP1	0.0915	0.2883	Unity if consumer unit expenditures for the quarter were less than $1,500; zero otherwise.
EXP2	0.2111	0.4081	Unity if consumer unit expenditures for the quarter were greater than or equal to $1,500 but less than $3,000; zero otherwise.
EXP3	0.1878	0.3906	Unity if consumer unit expenditures for the quarter were greater than or equal to $3,000 but less than $4,500; zero otherwise.
EXP4*	0.1580	0.3648	Unity if consumer unit expenditures for the quarter were greater than or equal to $4,500 but less than $6,000; zero otherwise.
EXP5	0.1213	0.3265	Unity if consumer unit expenditures for the quarter were greater than or equal to $6,000 but less than $7,500; zero otherwise.
EXP6	0.1188	0.3235	Unity if consumer unit expenditures for the quarter were greater than or equal to $7,500 but less than $10,000; zero otherwise.
EXP7	0.1115	0.3148	Unity if consumer unit expenditures for the quarter were greater than or equal to $10,000; zero otherwise.

Table 1A. *Continued*

Variable	Mean	Standard deviation	Definition
MEANST	0.1449	0.3520	Unity if consumer unit or an individual within the consumer unit participated in an income means tested program (i.e. received school lunch free or at a reduced price, lived in public housing and/or the government is paying part of the housing cost, received food stamps, received supplemental security income (SSI) checks from the government during the past twelve months, or received income from public assistance or welfare, e.g. Aid to Families with Dependent Children (AFDC) or money received from job training programs such as Job Corps); zero otherwise.
NONMEANST*	0.8551	0.3520	Unity if consumer unit did not participate in an income means tested program; zero otherwise.

*Denotes reference group.

Appendix B

Table 1B. *Variables, distributions and definitions: 1987 U.S. consumer expenditure survey second interview stricter definition of complete income reporter (N = 5,936)*

Variable	Mean	Standard deviation	Definition
INCCOM	0.6754	0.4683	Unity if income fields included valid responses, including at least one non-zero amount; zero otherwise.
NONINCCOM*	0.3246	0.4683	Unity if income fields were coded invalid for any income question where a monetary amount is reported or across the board blanks and/or zeroes resulted; zero otherwise.

*Denotes reference group.

Table 2B. *Estimated model parameters and standard errors: 1987 U.S. consumer expenditure survey second interview stricter definition of complete income reporter (N = 5,936)*

Independent variable	Estimated coefficient	Asymptotic standard error
Age of reference person (age 35 to 54 years)		
Aged 34 years or less	0.3442[a]	0.0812
Aged 55 to 64 years	−0.2156[b]	0.0971
Aged 65 years or greater	−0.3294[a]	0.1226
Race of reference person (non-black)		
Black	−0.1168	0.1026
Sex of reference person (male)		
Female	−0.1895[b]	0.0812

Table 2B.   Continued

Independent variable	Estimated coefficient	Asymptotic standard error
Education of reference person (high school graduate)		
Elementary	0.2896[a]	0.1071
Did not complete high school	0.0322	0.0971
Did not complete college	−0.0489	0.0845
College graduate	−0.2940[a]	0.1084
Postgraduate	−0.1432	0.1152
Principal occupation of reference person (salaried professional)		
Laborer	0.1930	0.1208
Craft	0.0869	0.1385
Sales	−0.0249	0.1006
Services	0.0021	0.1326
Self-employed	−0.5750[a]	0.1287
Retired	0.0046	0.1350
Not working and other	−0.1125	0.1281
Consumer unit composition (single)		
Single parent	0.0319	0.1583
Husband and wife only	−0.4769[a]	0.1018
Husband, wife, and child under 18	−0.4154[a]	0.1107
Husband, wife, and other	−1.1478[a]	0.1215
Other family combinations	−0.4740[a]	0.1136
Housing tenure (owns)		
Rents	0.3252[a]	0.0745
Degree urban (large urban area)		
Small urban area	0.0282	0.0692
Rural	0.2618[a]	0.0799
Region (South)		
Northeast	−0.7024[a]	0.0835
Midwest	−0.4805[a]	0.0789
West	0.1054	0.0886
Expenditure categories (expenditures of $4,500 to $5,999)		
Expenditures of less than $1,500	0.0561	0.1466
Expenditures of $1,500 to $2,999	−0.0350	0.1055
Expenditures of $3,000 to $4,499	0.1442	0.1004
Expenditures of $6,000 to $7,499	0.3057[a]	0.1113
Expenditures of $7,500 to $9,999	0.3992[a]	0.1133
Expenditures of $10,000 or more	0.2187	0.1169
Income means tested program (non-participant)		
Participant	0.3720[a]	0.1033
Constant	1.1677[a]	0.1630
Likelihood ratio statistic	528.32 > 63.69[a]	
Likelihood ratio index	0.070	

[a]Statistically significant at the 0.01 level.
[b]Statistically significant at the 0.05 level.

Table 3B.   Chi-square tests for contributions of sets of variables: 1987 U.S. consumer expenditure survey second interview stricter definition of complete income reporter (N=5,936)

Independent variable groups	Chi-square	Degrees of freedom
Age of reference person	37.216[a]	3
Race of reference person	1.286	1
Sex of reference person	5.454[b]	1
Education of reference person	18.450[a]	5
Principal occupation of reference person	33.571[a]	7
Consumer unit composition	103.142[a]	5
Housing tenure	19.179[a]	1
Degree urban	11.849	2
Region	122.350[a]	3
Expenditures	19.773[a]	6
Income means tested program	13.287[a]	1

[a]Statistically significant at the 0.01 level.
[b]Statistically significant at the 0.05 level.

Table 4B.   Sample probability calculations: 1987 U.S. consumer expenditure survey second interview stricter definition of complete income reporter (N=5,936)

Probability at the mean	0.6906
Baseline probability[1]	0.7627
Probability with change to consumer unit with selected characteristics:	
Age 34 years or less	0.8193
Age 55 to 64 years	0.7215
Age 65 years or greater	0.6981
Black	0.7409
Female	0.7267
Elementary	0.8111
Did not complete high school	0.7685
Did not complete college	0.7538
College graduate	0.7055
Postgraduate school	0.7358
Laborer	0.7959
Craft worker	0.7781
Sales	0.7582
Services	0.7631
Self-employed	0.6440
Retired	0.7636
Not working	0.7418
Single parent	0.7685
Husband and wife only	0.6661
Husband, wife, and child under 18	0.6797
Husband, wife and other	0.5050
Other family combinations	0.6668
Rents	0.8165
Small urban area	0.7678
Rural	0.8068
Northeast	0.6143
Midwest	0.6653
West	0.7813

Table 4B.   Continued

Expenditures of less than $1,500	0.7727
Expenditures of $1,500 to $2,999	0.7563
Expenditures of $3,000 to $4,499	0.7873
Expenditures of $6,000 to $7,499	0.8136
Expenditures of $7,500 to $9,999	0.8273
Expenditures of $10,000 or more	0.8000
Income means tested program participant	0.8234

[1]The base line probability was calculated for this consumer unit in the same way as was done when the less strict definition was assumed. Refer to Table 3.

5.   References

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

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

Coder, J.F. and Feldman, A.M. (1984). Early Indications of Item Nonresponse on the Survey of Income and Program Participation. Proceedings of the Section on Survey Research Methods, American Statistical Association, 693–697.

David, M.H., Little, R., Samuhel, M., and Triest, R. (1983). Imputation Models Based on the Propensity to Respond. Proceedings of the Section on Business and Economics Statistics, American Statistical Association, 168–173.

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

Fay, R.E. (1986). Causal Models for Patterns of Nonresponse. Journal of the American Statistical Association, 81, 354–365.

Garner, T.I. and Blanciforti, L.A. (1987). Reporting of Household Income: Complete Versus Incomplete Response. Proceedings of the Third Annual Research Conference, Bureau of the Census, 644–661.

Garner, T.I. and Blanciforti, L.A. (1991). Household Income Report Completeness: An Analysis of U.S. Consumer Expenditure Survey Data. Proceedings of the Section on Business and Economics Statistics, American Statistical Association, 304–309.

Greenlees, J.S., Reece, W.S., and Zieschang, K.D. (1982). Imputation of Missing Values When the Probability of Response Depends on the Variable Being Imputed. Journal of the American Statistical Association, 77, 251–261.

Hensher, D.A. and Johnson, L.W. (1981). Applied Discrete-Choice Modelling. London: Croom Helm.

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, J. and Lane, S. (1978). The Effects of Debt on Perceived Household Welfare. Journal of Consumer Affairs, 12, 48–62.

Lillard, L., Smith, J.P., and Welch, F. (1986). What Do We Really Know about Wages? The Importance of Nonreporting and Census Imputation. Journal of Political Economy, 94, 489–506.

Little, R.J.A. and Samuhel, M.E. (1983). Alternative Models for CPS Income Imputation. Proceedings of the Section on Survey Research Methods, American Statistical Association, 168–173.

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

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.

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

U.S. Department of Commerce, Bureau of the Census (1990). Money Income and Poverty Status in the United State 1988. Current Population Reports, Consumer Income Series P60 no. 166, Washington, D.C.: U.S. Government Printing Office.

U.S. Department of Labor, Bureau of Labor Statistics (1980). Partial Income Response Coding for the CEX. Memorandum from Charles M. Mason to Stephen Baer, April 10.

U.S. Department of Labor, Bureau of Labor Statistics (1981). Specifications for Coding of Income Response Status-Diary Survey. Memorandum from Eva E. Jacobs to John Marcoot, March 24.

U.S. Department of Labor, Bureau of Labor Statistics (1987). Unpublished Microdata and Documentation from the Quarterly Interview, Consumer Expenditure Survey, Washington, D.C.

U.S. Department of Labor, Bureau of Labor Statistics (1991). Consumer Expenditure Survey, 1988–89, Bulletin 2383, Washington, D.C.: U.S. Government Printing Office.

Vaughan, D.R. (1989). Reflections on the Income Estimates from the Initial Panel of the Survey of Income and Program Participation (SIPP). SIPP Working Paper no. 8906, U.S. Department of Commerce, Bureau of the Census, Washington, D.C.